

# UC San Diego

## UC San Diego Electronic Theses and Dissertations

### Title

The nature and time course of motor programming in apraxia of speech

### Permalink

<https://escholarship.org/uc/item/3vb0z9bx>

### Author

Maas, Edwin

### Publication Date

2006

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO  
SAN DIEGO STATE UNIVERSITY

The Nature and Time Course of Motor Programming in Apraxia of Speech

A Dissertation submitted in partial satisfaction of the requirement for the degree  
Doctor of Philosophy

in

Language and Communicative Disorders

by

Edwin Maas

Committee in charge:

University of California, San Diego  
Professor Victor Ferreira

University of California, Los Angeles  
Professor Richard A. Schmidt

San Diego State University  
Professor Donald A. Robin, Chair  
Professor Tracy E. Love  
Professor Lewis P. Shapiro

2006

Copyright

Edwin Maas, 2006

All rights reserved.

The Dissertation of Edwin Maas is approved, and it is acceptable in quality and form  
for publication on microfilm:

---

---

---

---

---

---

Chair

University of California, San Diego

San Diego State University

2006

## TABLE OF CONTENTS

Signature Page .....	iii
Table of Contents .....	iv
List of Abbreviations .....	ix
List of Figures .....	xi
List of Tables .....	xii
List of Graphs .....	xiv
Acknowledgements .....	xv
Vita .....	xviii
Abstract .....	xxv
1.0. Introduction .....	1
1.1. Overview .....	3
2.0. Motor Programming .....	6
2.1. Degrees of Freedom .....	6
2.2. Serial Order .....	8
2.3. Units and Processes .....	10
2.4. Speech as a Special Motor Skill .....	11
3.0. Models of Speech Motor Programming .....	21
3.1. Spoken Language Production .....	21
3.1.1. The Nijmegen Model .....	23
3.2. Units of Motor Control .....	26
3.2.1. Schema Theory .....	27

3.2.2.	Schema Theory and Speech Production .....	31
3.3.	Processes of Motor Control .....	34
3.3.1.	The Sternberg Model .....	35
3.3.2.	The INT/SEQ Model .....	40
3.3.3.	The INT/SEQ Model and Speech Production .....	45
3.4.	Summary .....	53
4.0.	Apraxia of Speech (AOS) .....	57
4.1.	Characteristics of AOS .....	57
4.2.	AOS in Relation to Models of Speech Motor Programming .....	64
4.2.1.	The Nijmegen Model and AOS .....	64
4.2.2.	Schema Theory and AOS .....	67
4.2.3.	The INT/SEQ Model and AOS .....	69
4.3.	Summary .....	73
5.0.	Speech Motor Programming in AOS and Aphasia .....	75
5.1.	Methods .....	77
5.1.1.	Participants .....	77
5.1.2.	Materials and Equipment .....	83
5.1.3.	Task and Procedures .....	84
5.1.4.	Design and Analysis .....	89
5.1.5.	Predictions .....	94
5.2.	Results .....	95
5.2.1.	Accuracy: Number of Error Trials .....	95

5.2.2.	INT: Study Time (ST) .....	99
5.2.3.	SEQ: Reaction Time (RT) .....	106
5.2.4.	Execution: Duration, and Absolute and Relative Timing Error .....	112
5.3.	Discussion .....	116
5.3.1.	Speech Motor Programming in Younger Control Speakers .....	116
5.3.2.	Speech Motor Programming in Apraxia of Speech and Aphasia .....	120
5.4.	Conclusions .....	133
6.0.	Nonspeech Motor Programming in AOS .....	135
6.1.	Methods .....	138
6.1.1.	Participants .....	138
6.1.2.	Task and procedures .....	139
6.1.3.	Design, Analysis, and Predictions .....	143
6.2.	Results .....	144
6.2.1.	Accuracy: Number of Error Trials .....	144
6.2.2.	INT: Study Time (ST) .....	146
6.2.3.	SEQ: Reaction Time (RT) .....	151
6.2.4.	Execution: Duration, and Absolute and Relative Timing Error .....	156
6.3.	Discussion .....	162

6.3.1.	Motor Programming in Younger Control Participants ...	162
6.3.2.	Motor Programming in Apraxia of Speech .....	164
6.4.	Conclusions .....	174
7.0.	The Role of Rhythm in Speech Motor Programming .....	176
7.1.	Experiment 3: INT and SEQ (Self-Selection Paradigm) .....	180
7.1.2.	Methods .....	184
7.1.2.1.	Participants .....	184
7.1.2.2.	Materials and Design .....	184
7.1.2.3.	Task and Procedures .....	187
7.1.2.4.	Design, Analysis, and Predictions .....	191
7.1.3.	Results .....	192
7.1.3.1.	Matching Task .....	192
7.1.3.2.	Self-selection Task .....	193
7.1.4.	Discussion .....	195
7.2.	Experiment 4: INT (Choice RT) .....	200
7.2.1.	Methods .....	201
7.2.1.1.	Participants .....	201
7.2.1.2.	Materials and Design .....	201
7.2.1.3.	Task and Procedures .....	201
7.2.1.4.	Design, Analysis, and Predictions .....	202
7.2.2.	Results .....	203
7.2.2.1.	Matching Task .....	203



7.2.2.2.	Choice RT Paradigm .....	204
7.2.3.	Discussion .....	205
7.3.	General Discussion .....	206
8.0.	Discussion and Conclusions .....	210
8.1.	Speech Motor Programming in the Intact System .....	213
8.2.	Speech Motor Programming in Apraxia of Speech .....	224
8.3.	Limitations .....	230
8.4.	Conclusion and Future Directions .....	233
Appendix A	.....	238
Appendix B	.....	240
Appendix C	.....	245
References	.....	247

## LIST OF ABBREVIATIONS

ABA-2 = Apraxia Battery for Adults – 2 (Dabul, 2000)

AMC = age-matched control group

APH = individual(s) with aphasia

AOS = Apraxia of Speech

BDAE = Boston Diagnostic Aphasia Examination (3<sup>rd</sup> edition) (Goodglass, Kaplan, & Barresi, 2001)

BNT = Boston Naming Test (Kaplan, Goodglass, & Weintraub, 1983)

CH = composition hypothesis

CON = control group

GMP = generalized motor program (Schmidt, 1975)

IM = integrative model (Ballard, Robin, & Folkins, 2003)

INT = preprogramming process of motor programming (Klapp, 1995)

MT = movement time

PALPA = Psycholinguistic Assessment of Language Processing in Aphasia (Kay, Lesser, & Coltheart, 1992)

RH = rhythm hypothesis

RT = reaction time

RSDT = revised standardized difference test (Crawford & Garthwaite, 2005)

SEQ = sequencing process of motor programming (Klapp, 1995)

SRS = scalable response structure (Shea & Wulf, 2005)

ST = study time

TDM = task-dependent model (Ziegler, 2003a)

WAB = Western Aphasia Battery (Kertesz, 1982)

YCON = younger control group

## LIST OF FIGURES

Figure 3-1. Self-selection paradigm (e.g., Immink & Wright, 2001) .....	43
Figure 5-1. Self-selection paradigm adapted to represent trial events as used in Experiment 1 .....	86
Figure 6-1. Self-selection paradigm adapted to represent trial events as used in Experiment 2 .....	141

## LIST OF TABLES

Table 5-1. Participant information for Experiment 1 .....	79
Table 5-2. Study Time (ST) comparisons for each patient against age-matched controls for single syllables .....	103
Table 5-3. Study Time (ST) comparisons for each patient against age-matched controls for sequence length .....	104
Table 5-4. Reaction Time (RT) comparisons for each patient against age-matched controls for single syllables .....	110
Table 5-5. Reaction Time (RT) comparisons for each patient against age-matched controls for sequence length .....	111
Table 5-6. Absolute timing error (E) for single syllables for each patient (collapsed across blocks), compared to young controls.....	115
Table 5-7. Relative (AE-prop) and absolute timing error (E) for sequences (collapsed across 4S and 4L, and across blocks) .....	115
Table 6-1. Participant information for Experiment 2 .....	140
Table 6-2. ST comparisons for each patient against the control sample for duration of single presses .....	150
Table 6-3. ST comparisons for Sequence Length (collapsed across blocks) for each patient .....	151
Table 6-4. RT comparisons for Duration (collapsed across blocks) for each patient .....	154
Table 6-5. RT comparisons for Sequence Length (collapsed across blocks) for each patient .....	155
Table 6-6. Single press duration for each patient (collapsed across blocks), compared to controls (N=13) .....	158
Table 6-7. Absolute timing error (E) for single presses for each patient (collapsed across blocks), compared to controls .....	159
Table 6-8. Relative timing error (AE-prop) and absolute timing error (E) for sequences (collapsed across 4S and 4L, and across blocks).....	160

Table 7-1. Overview of materials for Experiment 3 and 4.....	185
Table 7-2. Means (SDs) from Experiment 3.....	194
Table 7-3. Means (SDs) from Experiment 4.....	203
Table 8-1. Overview of the results from the four experiments.....	211
Table A-1. Results from the Apraxia Battery for Adults – 2 (ABA-2; Dabul, 2000) .....	238
Table A-2. Results from the Western Aphasia Battery (WAB; Kertesz, 1982) ..	239
Table B-1. Percent correct by response type (collapsed across blocks) for each patient against age-matched control group.....	243
Table B-2. Syllable duration for single syllables for each patient (collapsed across blocks), compared to young controls .....	244
Table C-1. Percent correct by response type (collapsed across blocks) for each patient against the control group .....	246

## LIST OF GRAPHS

Graph 5-1. Percent correct for each patient and both control groups, Experiment 1 (collapsed across blocks) .....	96
Graph 5-2. Study Time (msec) by block, for single syllable duration, 1S vs. 1L (A) and sequence length, 1 vs. 4 syllables (B) .....	100
Graph 5-3. Reaction Time across blocks for single syllable duration (A) and sequence length (B) .....	108
Graph 5-4. Single syllable durations by block for speakers with AOS (N=4) and young control speakers (N=8) .....	113
Graph 6-1. Percent correct by response type (collapsed across blocks), Experiment 2 .....	145
Graph 6-2. Finger movement Study Time for duration, 1S vs. 1L (A), and for sequence length, 1 vs. 4 (B) .....	147
Graph 6-3. Finger movement RT for duration, 1S vs. 1L (A) and for sequence length, 1 vs. 4 (B).....	152
Graph 6-5. Single button press duration by group and block.....	157
Graph B-1. Total percentage correct by group and block, collapsed across response type .....	240
Graph B-2. Distribution of error types as a percentage of all errors, for younger controls (N=12), collapsed across response types .....	241
Graph B-3. Distribution of error types as a percentage of all errors, for age- matched controls (N=5), collapsed across response types .....	241
Graph B-4. Distribution of error types as a percentage of all errors, for AOS group (N=4), collapsed across response types .....	242
Graph B-5. Distribution of error types as a percentage of all errors, for APH group (N=3), collapsed across response types .....	242
Graph C-1. Accuracy (percent correct) by block for Experiment 2 (collapsed across response types) .....	245

## ACKNOWLEDGEMENTS

Chapters 5 and 6 are being prepared for publication in a single manuscript entitled “Motor Programming in Apraxia of Speech”, by Edwin Maas, Donald A. Robin, David L. Wright, and Kirrie J. Ballard. The dissertation author is the primary investigator and author of the manuscript.

Chapter 7 is being prepared for publication, under the running title “The Role of Rhythm in Speech Motor Programming, by Edwin Maas and Donald A. Robin. The dissertation author is the primary investigator and author of the manuscript.

Although writing a dissertation can be a lonely process, this dissertation would not have been possible without the support of many others. I would like to acknowledge a number of people whose support has been invaluable throughout my doctoral career. First, my deepest thanks go to Don, the chair of my committee, for providing me with a broad and thorough education. I truly appreciate not only his sharp intellect, his outstanding mentorship, his unwavering support, and his frank honesty, as well as his sincere friendship over the years.

I also want to extend thanks to my committee members, Vic Ferreira, Tracy Love, Dick Schmidt, and Lew Shapiro, for all their insightful, constructive input, and their kind support throughout these years. My education has been enriched by having had the opportunity to interact with and learn from each of their individual mentoring styles. A special warm thanks also goes to David Swinney, with whom my journey to San Diego and the JDP started, and who sadly did not get to witness the end of this



journey. I will always remember him fondly as a kind, wise, respectful mentor and friend.

Throughout the years, I have benefited from the support and wisdom of my academic community, in particular Jessica Barlow, Marilyn Newhoff, Judy Reilly, Kirrie Ballard, Mike Caligiuri, David Wright, Gaby Wulf, and Mick McNeil. Thanks also to the members of the Laboratory for Speech and Language Neuroscience for their help and stimulating discussions, and to Carmen for helping navigate the bureaucracy. I am also grateful to Roelien Bastiaanse, for introducing me to the idea of a doctorate degree and the possibility of extending basic sciences to clinically relevant research. I would not be where I am today without you.

It goes without saying that I am indebted to the individuals who participated in these studies. Without their participation, this dissertation would not exist, and their enthusiasm made the experience all the more enjoyable. I especially appreciate the willingness of the individuals with speech and/or language problems to welcome me into their homes and to share their experiences with me.

I am extremely fortunate to have in my corner a number of wonderful friends, who have been there for me through thick and thin, who have supported me, fed me, entertained me, expanded my horizons in many different ways, made me laugh, and made me a better person. You know who you are. I also want to thank the House of Nine and the Language Science Lab at Boston University for providing a safe haven whenever I was in Boston.

Endless gratitude also goes to my family. My parents and my brother have always supported me in every possible way, and enabled me to pursue my dreams even if it meant that I would not live next door. Bedankt pa, ma, en Martin. Ik ben jullie eindelijk dankbaar voor alles wat jullie voor me doen, en voor het feit dat jullie altijd voor me klaar staan. Thanks also to the DeDe family for their support and for welcoming me into their family as their own.

Finally, extra special thanks go to Gayle, for her neverending support, her incisive intellect, her kindness, her wisdom, her sense of humor, and for always knowing how to make me feel good. You are the best, Gayle, and I could not have done this without you.

## VITA

### **Education**

- 1995 Bachelor of Arts, University of Groningen, The Netherlands (Dutch Literature and Linguistics)
- 1998 Master of Arts (cum laude), University of Groningen, The Netherlands (Neurolinguistics: Aphasiology)
- 2006 Doctor of Philosophy, University of California, San Diego, and San Diego State University (Language and Communicative Disorders)

### **Professional Experience**

Research Assistant 2000 to 2002  
Department of Communicative Disorders, San Diego State University  
Aphasia treatment research; comparing two treatment approaches to sentence production deficits in aphasia

Lab Manager / Research Assistant 1999  
Department of Linguistics, University of Groningen  
Psycholinguistic research; Dutch sentence processing, verb activation

Clinical Linguist 1999  
Rotterdam Aphasia Foundation (Rehabilitation Center Rijndam, Rotterdam)  
Assessment of language impairments secondary to neurological damage; design and evaluation of treatment progress in a multidisciplinary setting

Visiting Scholar 1998 to 1999  
Department of Communicative Disorders, San Diego State University  
Department of Psychology, University of California, San Diego  
Psycholinguistic research; on-line sentence processing, pronominal reference

Research Assistant 1997 to 1998  
Department of Linguistics, University of Groningen  
Development and norming of test battery for aphasic sentence-level deficits

Research Assistant 1997 to 1998  
Department of Linguistics, University of Groningen  
Psycholinguistic research; real-time sentence processing

Internship Clinical Linguist 1997  
Rotterdam Aphasia Foundation (Rehabilitation Center Rijndam, Rotterdam)  
Assessment of neurogenic language disorders; design and evaluation of treatment progress in a multidisciplinary setting.

## **Teaching Experience**

### Lecturer

*Phonetics* (undergraduate) 2003-2004  
School of Speech, Language, and Hearing Sciences (SDSU)

### Teaching Assistantships

*Phonetics* (undergraduate) 2000-2002  
School of Speech, Language, and Hearing Sciences (SDSU)

*Minds and Brains: Introduction to Cognitive Science* (undergraduate) 2002  
Department of Cognitive Science (UCSD)

*Speech Instrumentation* (graduate) 2004  
School of Speech, Language, and Hearing Sciences (SDSU)

*Language as a Cognitive System* (undergraduate) 2006  
Department of Linguistics (UCSD)  
Dr. Mieko Ueno, e-mail: ueno@ling.ucsd.edu

### Guest Lecturer

*Neurological Foundations of Cognition* (undergraduate) 2004-2005  
Department of Cognitive Science (UCSD)

## **Publications**

Ballard, K.J., Maas, E., & Robin, D.A. (in press). Treating control of voicing in apraxia of speech with variable practice. *Aphasiology*.

Freedman, S.E., Maas, E., Caligiuri, M.P., Wulf, G., & Robin, D.A. (in press). Internal vs. external: Oral-motor performance as a function of attentional focus. *Journal of Speech, Language, and Hearing Research*.

Robin, D.A., Maas, E., Sandberg, Y., & Schmidt, R.A. (in press). Motor control and learning and childhood apraxia of speech. To appear in: P. Hall & L. Jordan (Eds.), *Childhood Apraxia of Speech*.

Maas, E., Freedman, S., & Robin, D.A. (in press). Traumatic brain injury. To appear in M.L. McNeil (Ed.), *Clinical Management of Sensorimotor Speech Disorders (2nd edition)*. Thieme.

- Freedman, S., Maas, E., & Robin, D.A. (in press). Neurofibromatosis. To appear in M.L. McNeil (Ed.), *Clinical Management of Sensorimotor Speech Disorders (2nd edition)*. Thieme.
- Maas, E., Robin, D.A., Steinhauer, V., Ballard, K.J., Magnuson, C.E., & Wright, D.L. (2006). Motor programming in apraxia of speech: A reaction time approach. *Stem-, Spraak- en Taalpathologie, 14 (supplement), 69*.
- Maas, E. & Robin, D.A. (2005). Motor speech disorders: Apraxia and dysarthria. In L. Justice (Ed.), *Communication Sciences and Disorders: An Introduction*. Columbus, OH: Merrill Prentice Hall.
- Love, T., Maas, E., & Swinney, D. (2003). The influence of language exposure on lexical and syntactic language processing. *Experimental Psychology, 50 (3)*, 204-216.
- Bastiaanse, R., Edwards, S., Maas, E., & Rispens, J. (2003). Assessing comprehension and production of verbs and sentences: the Verb and Sentence Test (VAST). *Aphasiology, 17 (1)*, 49-73.
- Maas, E., Barlow, J., Robin, D., & Shapiro, L. (2002). Treatment of sound errors in aphasia and apraxia of speech: effects of phonological complexity. *Aphasiology, 16 (4/5/6)*, 609-622.
- Bastiaanse, R., Rispens, J., & Maas, E. (2000). De Werkwoorden- en Zinnentest (WEZT). Een nieuw diagnostisch instrument voor afasie. *Logopedie en Foniatrie, 72 (12)*, 259-267. [ The Verb and Sentence Test (VAST). A new diagnostic instrument for aphasia. ]
- Bastiaanse, R., Maas, E., & Rispens, J. (2000). *De Werkwoorden en Zinsverwerkingstest (WEZT)*. Lisse: Swets & Zeitlinger. [The Verb and Sentence Test (VAST). ]
- Maas, E. (1998). Sentence comprehension and sentence anagrams in Broca's aphasics and Wernicke's aphasics. Unpublished Masters thesis, University of Groningen, Groningen, The Netherlands.
- Maas, E. (1998). Zinsbegrip en zinsanagrammen bij Broca-patiënten en Wernicke-patiënten. *Toegepaste Taalwetenschap in Artikelen, 3*, 65-76. Proceedings of the ANéLA/VKL/WAP conference, May 15th 1998, Amsterdam). [ Sentence comprehension and sentence anagrams in Broca's aphasia patients and Wernicke's aphasia patients. ]

### **Papers submitted / in preparation**

Maas, E., Robin, D.A., Austermann Hula, S., Freedman, S.E., Wulf, G., Ballard, K.J., & Schmidt, R.A. (in preparation). Principles of motor learning in treatment of motor speech disorders. *American Journal of Speech-Language Pathology*.

Maas, E., Robin, D.A., Wright, D.L., & Ballard, K.J. (in preparation). Motor programming in apraxia of speech and aphasia.

Austermann Hula, S., Robin, D.A., Maas, E., Ballard, K.J., & Schmidt, R.A. (submitted). Principles of motor learning applied to the treatment of acquired apraxia of speech: Effects of feedback frequency and timing. *Journal of Speech, Language, and Hearing Research*.

### **Conference Presentations**

#### **Oral presentations**

Maas, E., Robin, D., Magnuson, C., Wright, D., & Ballard, K. (2005). Motor programming in apraxia of speech: central or modality-specific deficit? American Speech, Language, and Hearing Association Convention (San Diego, CA, November 2005).

Ballard, K.J., Maas, E., Robin, D.A., Guo, L.-Y. (2005). Treating interarticulator coordination in apraxia of speech with variable practice. Paper presented at the 35th Clinical Aphasiology Conference (Sanibel Island, FL, May-June 2005).

Wright, D.L., Magnuson, C.E., Robin, D., Maas, E., & Ballard, K. (2005). Programming timing during movement sequences: an investigation of Apraxia of Speech. Paper presented at the European Workshop on Movement Science (Vienna, Austria, June 2005).

Austermann, S., Robin, D., Ballard, K., Maas, E., & Schmidt, R. (2004). Low versus high frequency feedback in apraxia of speech. Paper presented at the Conference on Motor Speech: Motor Speech Disorders (Albuquerque, NM, March 2004).

Shapiro, L., Swinney, D., & Maas, E. (2003). The contributions of lexical and structural information in language comprehension. Symposium presented at the 33rd Clinical Aphasiology Conference (Orca's Island, WA, May 2003).

Maas, E. & Barlow, J. A. (2003). Complexity in language and learning: implications for planning treatment . Invited presentation at the First Annual Southern

California Conference on Speech, Language and Hearing Sciences (San Diego, CA, February 2003).

Friedmann, N. & Shapiro, L. (2002). Agrammatic comprehension of Hebrew OSV and OVS sentences. Paper presented on behalf of the authors, neither of whom was able to attend the 32nd Clinical Aphasiology Conference (Ridgedale, MO, June 2002).

Maas, E., Barlow, J., Robin, D., & Shapiro, L. (2001). Treatment of phonological errors in aphasia and apraxia of speech: Effects of phonological complexity. Paper presented at the 31st Clinical Aphasiology Conference (Santa Fe, NM, June 2001).

Maas, E., Barlow, J., Morris, D., Robin, D., & Shapiro, L. (2000). Treatment of phonological errors in aphasia: Effects of complexity. Paper presented at the 9th Annual International Aphasia Rehabilitation Conference (Rotterdam, the Netherlands, August 2000).

Maas, E. (1998). Zinsbegrip en zinsanagrammen in Broca-patiënten en Wernicke-patiënten. Paper presented at the TABU-day, Lectures in Linguistics (Groningen, the Netherlands, June 1998).

Maas, E. (1998). Zinsbegrip en zinsanagrammen in Broca-patiënten en Wernicke-patiënten. Paper presented at the Conference of the ANéLA/VKL/WAP (Amsterdam, the Netherlands, May 1998).

#### Poster presentations

Maas, E., Robin, D.A., Steinhauer, V., Ballard, K.J., Magnuson, C.E., & Wright, D.L. (2006). Motor programming in apraxia of speech: A reaction time approach. Poster presented at the 5th International Conference on Speech Motor Control, Nijmegen (the Netherlands), June 7-10, 2006. (Award for 2nd best poster of conference).

Maas, E., Robin, D.A., Steinhauer, V., Magnuson, C.E., Wright, D.L., & Ballard, K.J. (2006). Apraxia of speech as an impairment of preprogramming. Poster presented at the 36th Clinical Aphasiology Conference, Ghent (Belgium), May 29 – June 2, 2006.

Ballard, K.J., Maas, E., & Robin, D.A. (2006). Treating control of voicing in apraxia of speech with variable practice. Poster presented at the Western Australia 2006 Speech Pathology Australia National Conference, Frontiers (Fremantle, Australia, May 2006).

- Maas, E., Robin, D.A., Ballard, K.J., Magnuson, C.E., & Wright, D.L. (2006). Speech motor programming in apraxia of speech: A reaction time approach. Poster presented at the Conference on Motor Speech: Motor Speech Disorders (Austin, TX, March 2006).
- Austermann, S.N., Robin, D.A., Maas, E., Ballard, K.J., & Schmidt, R.A. (2006). Principles of motor learning applied to the treatment of acquired apraxia of speech: Feedback frequency and timing. Poster presented at the Conference on Motor Speech: Motor Speech Disorders (Austin, TX, March 2006).
- Freedman, S.E., Maas, E., Caligiuri, M.P., Wulf, G., & Robin, D.A. (2006). Oral-motor learning, retention, and transfer as a function of attentional focus. Poster presented at the Conference on Motor Speech: Speech Motor Control (Austin, TX, March 2006).
- Freedman, S. Maas, E., Wulf, G., Caligiuri, M., & Robin, D. (2005). Effects of attentional focus on oral motor control and learning. American Speech, Language, and Hearing Association Convention (San Diego, CA, November 2005).
- Maas, E., Robin, D.A., Magnuson, C.E., Wright, D.L., & Ballard, K. (2005). Motor programming in apraxia of speech and aphasia. Clinical Aphasiology Conference, Sanibel Island, FL, May-June 2005.
- Austermann, S., Robin, D.A., Maas, E., Ballard, K., & Schmidt, R. (2005). The effect of feedback delay on the treatment of acquired apraxia of speech. Clinical Aphasiology Conference, Sanibel Island, FL, May-June 2005.
- Freedman, S., Maas, E., Wulf, G., Caligiuri, & Robin, D.A. (2005). Effects of attentional focus on oral-motor control and learning. Clinical Aphasiology Conference, Sanibel Island, FL, May-June 2005.
- Austermann, S., Maas, E., Robin, D., Ballard, K., & Schmidt, R. (2004). The role of feedback frequency in the treatment of acquired apraxia of speech. Poster to be presented at the 34th Clinical Aphasiology Conference (Park City, UT, May 2004).
- De Goede, D., Wester, F., Bastiaanse, R., Maas, E., Den Ouden, D.B., Shapiro, L., & Swinney, D. (2003). Verb movement and gap-filling in Dutch sentence processing. Poster presented at the Verb and Sentence Processing in Aphasia Group conference on Verbs: Properties, Processes and Problems (London, UK, April 2003).



Wester, F., De Goede, D., Bastiaanse, R., Maas, E., Den Ouden, D.B., Shapiro, L., & Swinney, D. (2003). Verb activation in on-line sentence processing. Poster presented at the 16th Annual CUNY Conference on Human Sentence Processing (Cambridge, MA, March 2003).

Maas, E., Shapiro, L., & Swinney, D. (2003). The nature of gap-filling: syntactic or semantic order? Poster presented at the 16th Annual CUNY Conference on Human Sentence Processing (Cambridge, MA, March 2003).

Bastiaanse, R. & Maas, E. (1998). Verb movement in agrammatic production. Poster presented at the Academy of Aphasia (Santa Fe, NM, November 1998).

### **Professional Service**

Ad Hoc Reviewer, Journal of Speech, Language, and Hearing Research	2006
Ad Hoc Reviewer, American Journal of Speech-Language Pathology	2005
Member of subcommittee on "Motor Speech in Adults and Children" for ASHA Convention, 2005 (San Diego, CA)	2005
UCSD Graduate Student Representative	2000-2005
UCSD GSA Finance Committee	2001-2005
UCSD GSA Finance Reform Committee	2001
UCSD Office for Students with Disabilities Advisory Committee	2001-2005
UCSD Academic Dishonesty Hearing Board	2001-2005
UCSD Student Regulations and Revisions Committee	2002-2003
UCSD Campus-Wide Judicial Board	2003-2005
Graduate Seminar Series Coordinator	2002-2005

### **Awards and Honors**

2nd Best Poster Award at the 5th International Speech Motor Control conference, 2006 (Nijmegen, Netherlands)

NIDCD Research Symposium in Clinical Aphasiology Student Fellowship for presentation at the Clinical Aphasiology Conference, 2005 (Sanibel Island, FL)

American Speech-Language Hearing Foundation Graduate Student Scholarship, 2004

Kala Singh Memorial Scholarship, 2004

NIDCD Research Symposium in Clinical Aphasiology Student Fellowship, for presentation at the Clinical Aphasiology Conference, 2004 (Park City, UT)

Friends of the International Center at UCSD Scholarship, 2003

Friends of the International Center at UCSD Scholarship, 2001

Phi Beta Delta International Student Scholarship, 2001

VSB Scholarship, 1998

### **Professional Affiliations**

National Student Speech Language and Hearing Association, since 2003

Vereniging voor Klinische Linguïstiek (Clinical Linguistics Association), since 1998

ABSTRACT OF THE DISSERTATION

The Nature and Time Course of Motor Programming in Apraxia of Speech

by

Edwin Maas

Doctor of Philosophy in Language and Communicative Disorders

University of California, San Diego, 2006

San Diego State University, 2006

Professor Donald A. Robin, Chair

Speech production involves many different muscles at very high speeds, and yet we can speak fluently with very few errors. Speech motor programming refers to the processes that transform an abstract linguistic code into coordinated patterns of muscle activity. Speech motor programming can be impaired secondary to

neurological damage, resulting in the disorder called Apraxia of Speech (AOS). This disorder remains poorly understood. In this dissertation, the nature of speech motor programming and AOS is investigated in a model-driven approach to address the following questions: 1) What are the units and processes of speech motor control?; 2) Is speech controlled by a separate motor control system?; 3) Does AOS involve a process-specific deficit?; and 4) Does AOS involve a speech-specific deficit?

The research in this dissertation is framed in the context of a recent model of motor programming that delineates two processes occurring at different points in time, namely preprogramming and sequencing. In Experiment 1, a reaction time methodology that has been shown to capture these two processes was used to examine the programming of repeated syllables in unimpaired speakers and in individuals with AOS or aphasia. The results provided support for the hypothesis that AOS involves a deficit in preprogramming but not in sequencing or movement initiation; both processes appeared to be intact in individuals with aphasia. In addition, the results from unimpaired speakers suggested that syllable sequences are programmed as single units. Experiment 2 extended these findings to the programming of finger movement sequences, where it was found that individuals with AOS showed similar preprogramming deficits. Experiments 3 and 4 were designed to examine the influence

of rhythm on speech motor programming in unimpaired speakers. The results suggested that rhythmic sequences are preprogrammed faster than non-rhythmic sequences, although methodological concerns preclude strong conclusions.

In conclusion, this dissertation contributes to our understanding of speech motor programming in terms of its time course, its units, and its component processes. In addition, the nature of the motor programming deficit in AOS was further delineated to represent a non-speech specific deficit in preprogramming, while sequencing and initiation appear to be intact.

## **1.0. Introduction**

Speech production, one of our most effective means of communication, involves a highly complex spatio-temporal coordination of the muscles of respiration, phonation, resonance, and articulation. Even within a given speech production subsystem (e.g., articulation), the coordinative demands required to produce fluent, effortless speech are extremely high. One aspect of speech production that has received relatively little attention is the central planning and programming of speech movements. Speech motor programming is defined here as the set of processes responsible for transforming an abstract linguistic (phonological) code into spatially and temporally coordinated patterns of muscle contractions that produce the audible movements we recognize as speech (cf. Spencer & Rogers, 2005). This dissertation represents the initial development of a research program focused on post-lexical stages of speech production, and has the following two main aims:

- 1) To increase our understanding of the central control involved in movement preparation and execution in normal speech production. Specifically, a contemporary model of motor programming is applied to speech production using a reaction time approach, to address the following questions:
  - i. What is the nature of units of speech motor programming?
  - ii. Are there separate motor programming subprocesses?

2) To increase our understanding of disorders of speech motor programming. Specifically, apraxia of speech (AOS) is examined in the context of a recent model of motor programming, to address the following questions:

- i. Does AOS involve a deficit in a specific motor programming subprocess?
- ii. Does AOS involve a central (not speech-specific) deficit in motor programming?

The long-term objectives of my research program include understanding the cognitive and neural mechanisms involved in speech motor control and learning, the relations among phonological planning, motor planning and programming, sequencing of speech movements, and the treatment of motor speech disorders such as AOS. Apraxia of speech is considered a disorder of speech motor programming, and has been defined as an impairment of transforming an intact phonological code into kinematic patterns (McNeil, Robin, & Schmidt, 1997). A more in-depth discussion of AOS is provided in Chapter 4.

The four studies that comprise this dissertation have both theoretical and clinical significance, and serve as the basis of a translational approach to speech motor control and learning for my long-term goals. Theoretically, understanding the various processes and units involved in the preparation and execution of speech movements is critical to developing cognitive models of speech production. Given the relatively underspecified nature of speech motor programming and learning in current models of spoken language production (e.g., Dell, 1986; Levelt, Roelofs, & Meyer, 1999), this

dissertation starts from the premise that applying the methodologies and general conceptual framework of motor programming to the study of speech production can provide insight into the nature of units of speech motor control, the on-line programming of speech, and the mechanisms that underlie speech breakdown and (re)learning.

Clinically, improving treatment outcomes for individuals with motor speech disorders requires an understanding both of the nature of the disorder and of the mechanisms and conditions of learning. Regarding the nature of the disorder, study and understanding of speech disorders should be driven by current models of speech production, which may allow for a more precise localization of the deficit(s). There are indications that speakers with AOS also exhibit difficulties in nonspeech oral motor tasks involving controlled spatiotemporal coordination of articulators (e.g., Ballard & Robin, in press; Hageman, Robin, Moon, & Folkins, 1994; McNeil, Weismer, Adams, & Mulligan, 1990), which strongly suggests that an account of AOS should be grounded in a theory of motor control, rather than (exclusively) in a theory of speech production.

### **1.1. Overview**

The experiments reported in this dissertation were designed to address the questions stated above. Chapter 2 provides a discussion of several issues related to motor programming and speech motor programming in particular. One major issue in the study of motor control and speech production (and cognition in general), pertains

to the nature of the representations and processes that produce a particular action (or percept). With respect to speech motor control, an additional issue that continues to be subject of debate in the literature is the extent to which speech and nonspeech motor skills share motor control systems. This issue is discussed in greater detail in Chapter 2 (Section 2.4), followed by a discussion of various models of motor programming. Included in this discussion are models that were originally developed on the basis of nonspeech motor behavior, but whose constructs and methods are potentially relevant to understanding the representations or units and the processes involved in controlling speech movements. As well, these models are applied to understanding disorders of speech motor programming, in particular AOS, in order to provide a model-driven approach to the delineation of the nature of the deficit in AOS.

The present work adopts a reaction time approach, based on the assumption that motor programming typically occurs prior to movement initiation. The paradigm used in this work further contributes to our understanding of the time course of motor programming by allowing the separation of preprogramming and processes related to movement initiation. This paradigm, which will be further discussed in Chapter 3 (Section 3.3.2), also has been shown to be sensitive to the number of units that a sequence comprises.

With respect to the first aim of increasing our understanding of the central control involved in movement preparation and execution in normal speech production, Experiment 1 (Chapter 5) examines the differential demands on separate motor programming stages for the production of single syllables or syllable sequences with



specific timing requirements, to determine whether in such sequences the syllable or the sequence functions as a unit.

In addition, the performance of individuals with AOS on this task is compared to the performance of a control group, in order to determine whether the deficit in AOS can be localized to a specific speech motor programming process. Such a finding would not only further delineate the nature of AOS, but would also provide neuropsychological evidence to support the existence of separate processes of motor programming. Experiment 2 (Chapter 6) examines the programming of finger movement sequences with identical temporal patterns as the syllable sequences in Experiment 1, to determine whether deficits in speech motor programming observed in individuals with AOS represent a speech-specific motor programming problem or a more domain-general problem. Such a finding would further suggest that motor programming for speech and motor programming for nonspeech movements may share parts of a motor control system.

In Experiments 3 and 4 (Chapter 7), a new set of syllable sequences, with different rhythmic patterns, is examined in unimpaired speakers, to determine how rhythm affects the units and processes of speech motor programming. Finally, Chapter 8 provides a discussion of the findings from the four experiments, in relation to the aims and questions posed above.

## **2.0. Motor Programming**

In general, motor programming involves the specification of time, space, and effector systems to be used for a particular action, in order to achieve a specific goal. People perform skilled actions every day, and many of these movements are quite complex and fast (e.g., writing, washing our hands), involving the coordination in space and time of many different muscles and muscle groups. This is especially true of speech production which requires the coordination of approximately 140,000 neuromuscular events per second during conversation (Darley, Aronson, & Brown, 1975). During speech production, discrete and invariant language units (e.g., phonemes, syllables) must be mapped onto space and time in order to produce fluent and context-dependent speech. Each abstract phonological unit maps onto a different set of articulators and muscle groups that have to be coordinated with respect to each other. How we manage to control such actions is a fundamental question to be answered by any model of motor control.

### **2.1. Degrees of Freedom**

The large number of muscles and muscle groups involved in many actions means that there are a large number of degrees of freedom (independently controllable structures), which may lead to a processing overload for the motor system if each of these must be controlled independently (the degrees of freedom problem; Kent, Adams, & Turner, 1996). One common solution to this problem is to assume that the neuromotor control system for speech production (and other cognitive motor systems

as well) reduces the number of degrees of freedom (and thus, processing load) by controlling certain aspects of a movement as a single unit (e.g., Abbs, Gracco, & Cole, 1984; Keele, 1968; Varley & Whiteside, 2001a). An often-made assumption following from the idea that certain aspects of a movement are controlled together is that such units can be specified in advance (programmed) from internal activation or central drive (e.g., linguistic input, goal direction), without reference to feedback from the environment (see for example, Abbs et al., 1984, relative to speech motor control).

However, an alternative approach is to assume that there are not preconceived units that are programmed, but rather that each action emerges as part of a dynamic system, where muscles form stable groupings, called coordinative structures (Kelso, Saltzman, & Tuller, 1986), on-line based on the input from the environment in order to achieve the demands of a particular situation. Thus, in this view, the degrees of freedom are also reduced by formation of groupings of muscle that act as one unit; the critical difference is that such coordinative structures do not become stored in memory as motor programs that can be used to specify the action in advance. However, the utility of the dynamic systems approach to understanding motor control and learning has substantial limitations (e.g., Schmidt, 2003; Schmidt & Lee, 2005), and there is strong evidence to support the notion that a certain pattern of muscle contractions is specified in the absence of sensory feedback (see Section 3.2.1). Moreover, programming models have clear predictions relative to motor learning, while dynamic approaches have not considered learning in a systematic manner to date. For these reasons, this theoretical perspective is not considered in the work presented here.

Within the motor programming approach to motor control, it is generally assumed that with practice, a number of smaller movement elements (e.g., gestures) can be integrated into a larger unit, a process referred to as chunking (e.g., Bullock, 2004; Klapp, 1995; Sakai, Hikosaka, & Nakamura, 2004; Verwey, 1999, 2003; Wright, Black, Immink, Brueckner, & Magnuson, 2004). In a sense, formation of larger units out of smaller ones with practice can be seen as a shift from serial processing (serial activation of successive elements) to more parallel processing (simultaneous activation of successive elements/motoneuron groups) (Verwey, 1999).

## **2.2. Serial Order**

A second fundamental problem for serial behaviors such as speech production is how the serial order of different aspects of an action is controlled (Lashley, 1951; Keele, Cohen, & Ivry, 1990; Kent et al., 1996). Lashley was the first to pose this problem, and he argued against an associative chaining view, according to which the feedback produced by one element of a sequence provides the cue to initiate the next element. Lashley pointed out several problems with this view, such as the fact that movement elements typically occur in many different serial positions, rendering the cue of a given element ambiguous as to the identity of the next element, and the fact that delays associated with the feedback loops are too long to function as triggers for the next element.

Instead, Lashley proposed that all elements of a sequence are co-activated and that their order is specified by a plan at a higher level. Support for this idea comes

from several sources, including 1) the fact that serial order errors sometimes occur (e.g. big dipper → dig bipper), suggesting that later-occurring elements are simultaneously available for production, 2) the fact that a movement may be carried out with different effector systems, suggesting the existence of a more abstract representation of the movement that does not include muscle-specific information (e.g., Keele, Jennings, Jones, Caulton, & Cohen, 1995), 3) the fact that sequence accuracy and element accuracy are independent (e.g., Lai, Shea, Wulf, & Wright, 2001; Shea & Park, 2003), and 4) the fact that movement initiation times are longer before more complex (longer) movements than before simple movements (e.g., Henry & Rogers, 1960; Sternberg, Monsell, Knoll, & Wright, 1978; Klapp, 1977, 1995; Klapp & Erwin, 1976).

Ever since Lashley's seminal paper, the assumption that serial behavior involves a hierarchical control structure in which a higher-level plan activates and orders units at a lower level has been widely accepted in the motor behavior literature (e.g., Bullock, 2004; Keele, Cohen, & Ivry, 1990; Klapp, 1995; Schmidt & Lee, 2005; Semjen & Gottsdanker, 1990; Sternberg et al., 1978; see Rhodes, Bullock, Verwey, Averbeck, & Page, 2004, for review) as well as in the speech production literature as the distinction between frames and fillers (Dell, 1986; Levelt, 1989; MacNeilage, 1998; Shattuck-Hufnagel, 1979, 1992). In this dissertation, I will focus on models in this general class.

### **2.3. Units and Processes**

The notion that control of serial motor behavior such as speech involves units of action in a hierarchical architecture raises several important issues. First, it must define, specify, and identify the presumed units or primitives (size, detail of representation). Second, it must define, specify, and identify the putative processing stages, their independence (modular vs. interactive), and their temporal organization (serial vs. parallel). With respect to speech, this second issue involves the question of how the language system activates the speech motor system. This issue has received relatively little attention in the literature (Kent et al., 1996), and it has been noted that the putative separation between phonological encoding and motor programming levels is difficult to verify empirically, due to the cumulative and potentially interactive nature of the system (Rogers & Storkel, 1998). Thus, factors that affect processing at a phonological level (e.g., syllable structure) will also exert an influence on subsequent motor programming levels (Maas, Barlow, Robin, & Shapiro, 2002).

Models of speech production can be distinguished from each other on the basis of the assumptions they make with respect to these issues. In this dissertation, the focus is largely limited to those models that 1) include a motor programming component, 2) have been used to account for disorders of speech motor programming, i.e. AOS, 3) make predictions about reaction times, and 4) include mechanisms of speech motor learning. Specifically, the models to be discussed include the Nijmegen model of spoken language production (Levelt, 1989; Levelt, Roelofs, & Meyer, 1999), the Schema Theory of motor control and learning (Schmidt, 1975, 1976, 2003;

Schmidt & Lee, 2005), the Sternberg model of rapid serial movement production (Sternberg et al., 1978), and the recent INT/SEQ model of motor programming proposed by Klapp (1995, 2003). These models will be discussed in Chapter 3.

#### **2.4. Speech as a Special Motor Skill**

An important question with respect to speech motor programming is whether speech is controlled by a specialized motor control system or whether speech shares a control system with other motor skills (e.g., Ballard, Robin, & Folkins, 2003; Folkins, Moon, Luschei, Robin, Tye-Murray, & Moll, 1995; Robin, Solomon, Moon, & Folkins, 1997; Weismer, 2006; Weismer & Liss, 1991; Ziegler, 2002, 2003a, b). Two positions have emerged from recent debate in the literature (e.g., Ballard, Robin, & Folkins, 2003; Ziegler, 2003a, b). The *task-dependent model* (TDM; Ziegler, 2003a,b) assumes that the oral motor system is partitioned along the lines of different behavioral goals or tasks, including emotional expression (e.g., crying), vegetative functions (e.g., chewing, swallowing), speech, and novel volitional motor tasks (e.g., visuomotor tracking, diadochokinetic tasks, imitation). The *integrative model* (IM; Ballard, Robin & Folkins, 2003) on the other hand assumes that the oral motor system is partitioned along the lines of the specific demands of the movement to be performed (e.g., the temporal and spatial coordination of different muscles), regardless of the intended behavioral goal. In other words, the TDM proposes a specialized, independent neuromotor control system for speech, whereas the IM proposes a more

general neuromotor control system that overlaps between tasks to the extent that the tasks share properties.

One source of support for the TDM comes from neuropsychology, specifically from double dissociations between vegetative breathing and breathing for speech, swallowing and speech, emotional expression and volitional movements, speech and cranial nerve functioning, and apraxia of speech and oral apraxia (see Ziegler, 2003a, for review). Most of these double dissociations have involved rather crude measures, such as comparing tasks in terms of presence or absence of an impairment (e.g., finding normal swallowing in the presence of dysarthria), and are not double dissociations in a statistical sense (e.g., Bates, Appelbaum, Salcedo, Saygin, & Pizzamiglio, 2003; Crawford, Garthwaite, & Gray, 2003). With respect to the relation between speech and other volitional movements, there have been a few studies that have attempted to find correlations between severity of the impairments, or between acoustic measures in speech and in nonspeech tasks.

An example of the latter approach is a study by Ziegler (2002), in which syllable duration in diadochokinesis (DDK; rapid syllable repetition) was compared to the duration of the same syllables in a sentence repetition task, for a range of speech disorders. On the assumption that repeating the same syllable was somewhat speech-like (Ziegler, 2003a: 20), correlations were expected between the duration measures of the same syllables. Although Ziegler (2002) did obtain correlations (ranging from .59 to .92 in patients, .44 in controls) between speaking rate in DDK and in sentence repetition, he argued that speech and DDK were subserved by different neuromotor



control systems based on the finding that when subgroups were matched for articulation rate in sentence repetition, patients with ataxic dysarthria showed greater slowing of articulation in DDK relative to sentence repetition, whereas patients with AOS did not show a difference in articulation rate between DDK and sentence repetition tasks. In other words, the patients with ataxic dysarthria showed that DDK and sentence repetition dissociate.

Another line of evidence that has been used to support the notion of separate subsystems for different oral motor tasks comes from comparisons of kinematics for different tasks. For example, Moore, Caulfield, and Green (2001) examined respiratory kinematics in toddlers during rest breathing, speech, and nonspeech vocalizations. These authors observed that the respiratory kinematics differed between rest breathing and the other tasks, which suggests that speech breathing does not emerge from rest breathing and thus may be subserved by a separate speech motor control system.

Alternatively, proponents of the integrative model have sought to find associations between impairments, arguing that a lack of associations may be due to insensitivity of the task or measure to reveal impairments, and thus, shared functions (Ballard et al., 2003). Robin and others have used the association approach to find correlations between speech and nonspeech tasks that may reliably predict a speech disorder. One of the most promising tasks in this regard is the visuomotor tracking (VMT) task (e.g., Ballard & Robin, in press; Hageman et al., 1994; Robin, Hageman, Moon, Clark, Woodworth, & Folkins, submitted).

In a visuomotor tracking task, participants track a predictable or unpredictable visual pattern on a screen with a cursor that moves in response to movement of an articulator (e.g., the jaw, lip) or modulation of voice pitch. Hageman et al. (1994) studied both in three tracking conditions, namely tracking with the jaw, lower lip, or voice fundamental frequency. The predictable patterns involved sinusoidal targets that varied in frequency (0.3, 0.6, and 0.9 Hz). The unpredictable pattern was a composite of ten sine waves with different frequencies but equal amplitudes. Importantly, the sinusoidal movement captures a critical aspect of speech in that both gestures have a peak velocity that occurs in the middle of the movement (Gracco & Abbs, 1986). Analyses in the Hageman et al. study focused on phase-shift and accuracy. Hageman et al. found that patients with apraxia of speech (AOS) showed a significant phase-lag for the predictable patterns, whereas the control participants and patients with conduction aphasia did not. However, for the unpredictable patterns both groups showed a significant phase-lag. As for tracking accuracy, cross-correlations between target and cursor for individuals with AOS were lower and more variable than those for the control participants in the predictable conditions, although both groups displayed decreases in performance with increasing target frequency. Interestingly, the performance for both groups was similar with respect to the unpredictable pattern. Thus, participants with AOS showed impaired intra-articulator coordination in a dynamic nonspeech motor task, which may suggest that the participants with AOS were less efficient in developing and implementing a motor program for the predictable movement (Hageman et al., 1994). The implication is that AOS is not a

disorder of speech motor control per se, but affects other motor control tasks as well, and thus that speech and nonspeech oral motor tasks share an overlapping neuromotor control system.

This conclusion is strengthened by the observation that all patients with AOS that have been tested on this paradigm have demonstrated a problem in tracking predictable patterns (Ballard, Granier, & Robin, 2000: 986). Moreover, a recent study not only replicated these findings but also observed that visuomotor tracking performance for predictable (but not unpredictable) targets correlates highly with perceptual ratings of intelligibility (Robin et al., submitted), suggesting that tracking of predictable targets has predictive value of speech impairment in AOS.

According to the IM, the dissociations observed between different tasks should be explained in terms of the properties of the tasks rather than by an a priori partition by behavioral goals. Dissociations between chewing and speaking, for example, may reflect dissociations along dimensions such as temporal requirements (speech tends to involve high velocities, unlike chewing), force requirements (greater force is required in chewing than in speech) (Ballard et al., 2003), and control of breathing (speaking occurs only on the exhalation cycle and thus respiration must be coordinated with oral movements, whereas one can inhale and exhale during chewing).

Similarly, the findings by Moore and colleagues (2001) of different respiratory kinematics during rest breathing and speech breathing merely suggest that these tasks differ; they do not necessitate the conclusion that the relevant difference is that of speech versus nonspeech. Interestingly, Moore et al. did not observe any differences

between speech and nonspeech vocalizations in terms of respiratory kinematics, suggesting that indeed some nonspeech tasks share properties with speech motor control. Moreover, Boliek, Hixon, Watson, and Morgan (1996, 1997) have demonstrated similar respiratory kinematic patterns in speech and a variety of nonspeech behaviors (e.g., crying, whimpering, grunting) in infants and toddlers.

The division of labor between the different putative motor systems in the TDM is not always clear. For example, which system is in control of the oral musculature when we speak while crying or eating? Which system is responsible for repeating nonwords (by definition a novel volitional oral motor task that involves producing speech sounds)? Which system is responsible for producing words in a second language learned in adulthood? And, if proficiency is achieved in a second language, does this indicate a shift to the speech motor control system, or the development of a separate speech motor control system for the other language, or merely improved ability to use the nonspeech volitional motor system? These questions essentially lead to a fundamental challenge for the TDM, which is to provide a clear and universally accepted definition of “speech” (e.g., Ballard et al., 2003). However, defining speech turns out to be more difficult than it seems. For example, Ziegler (2006; see also Weismer, 2006) defines speech as the production of oral movements that generate sound in the service of communication. Such a definition appears to exclude word repetition (imitative task, absence of true communicative goal) and production of nonwords (absence of meaning), and may include communicative oral motor

behaviors not typically viewed as speech, such as blowing someone a kiss, alveolar clicks indicating disapproval, or whistling to get someone's attention.

The resolution of the issue of modularity of the speech motor control system has important methodological implications for how speech can be studied. Many laboratory tasks used to investigate speech production can be viewed as nonspeech tasks, in that they involve tasks without a communicative goal (e.g., picture naming to a computer screen; word or sentence repetition; nonword production). According to the above definition of speech in the TDM, such tasks cannot in principle speak to the nature of speech motor control since they are controlled by the novel volitional oral motor control system (although recall that the speech task Ziegler utilized in his 2002 study involves sentence repetition). Presumably, the definition of speech can be modified to more accurately capture those behaviors that most people would consider speech, for example by saying that communicative intent or semantic content is not a necessary condition and that as long as the phonological system is engaged, one can call the behavior speech. However, this would then include the DDK tasks that Ziegler (2002, 2003a) explicitly excludes as nonspeech. Ziegler (2002, 2003a) discusses several differences between DDK and "speech", which includes the fact that DDK tasks are maximal performance tasks. If rate is the relevant factor, then repeating syllables at a normal rate might be considered speech. However, if one slows speech down further, the speaker may enter a nonspeech mode of control (Ziegler, 2003: 24). This also implies that a definition in terms of engaging the phonological system is insufficient, and furthermore implies that many speakers with motor speech disorders,

who often show significantly reduced speech rates, may in fact not be producing speech at all, but rather produce sounds using their novel volitional motor system (which would then of course make the use of nonspeech tasks relevant again). The relevance of these comments is to demonstrate the point that defining speech is not a trivial matter, and that this approach ultimately boils down to listing a set of properties that speech possesses, which is essentially the approach advocated by proponents of the IM.

There are also important clinical implications of these opposing views. In terms of diagnostics, the TDM implies that no useful information about motor speech disorders (esp. AOS but also dysarthria) can be derived from using nonspeech tasks such as visuomotor tracking or DDK (Ziegler, 2002, 2003a), since such tasks can only speak to an entirely different oral motor system. In contrast, the IM suggests that carefully designed nonspeech tasks may reveal important information about the nature of motor speech disorders by examining the abilities and limitations of the oral motor system independent from the linguistic input to this system (e.g., Robin et al., 1997).

In terms of treatment, a definition of speech that excludes nonwords implies that treatment for motor speech disorders should only use real word targets, since no transfer is to be expected from targeting nonwords, based on the generally accepted notion that transfer of learning only occurs to behaviors that share properties with the treated behavior (e.g., Ballard et al., 2003; Schmidt & Lee, 2005). There is evidence to suggest that treatment involving repetition of nonwords does in fact generalize to production of real words (e.g., Maas et al., 2002; Schneider & Frens, 2005). Such

findings suggest that a definition of speech should include nonwords, and that semantic meaning and a communicative goal are not necessary conditions for a behavior to constitute speech. Moreover, there is preliminary evidence that even practice on a nonspeech task involving a complex sequence of oral movements and generating intraoral pressure results in changes in generation of intraoral pressure during speech (Shaiman, McNeil, Szuminsky, Meigh, & Botler, 2006), suggesting overlapping motor control systems for speech and nonspeech tasks to the extent they share properties.

This dissertation is framed within the IM theory for several reasons. First, evidence for the TDM comes from dissociations between behaviors, which have been described in terms of the speech-nonspeech distinction. However, as noted, the absence of a clear a priori definition that captures speech and only speech substantially weakens the TDM as it essentially reduces the model to the IM approach of conceiving of speech not as a unitary concept but as a set of certain properties that combine or integrate to define speech, and that may occur in different combinations in different motor tasks. Thus, dissociations may best be understood in terms of these properties, and our understanding of speech motor control will be increased if we can identify those properties (e.g., by examining these properties in isolation from linguistic input). In contrast, the observed associations between speech and nonspeech tasks cannot be easily explained by the TDM. In this model, additional mechanisms must be postulated to account for such associations.

Second, and relatedly, the IM forces one to widen the scope of research to other, nonspeech tasks, an approach that in principle allows one to make more accurate generalizations than a focus restricted to speech production only, and by extension may lead to a deeper understanding of the nature of motor control for speech as well as nonspeech behavior, and of motor speech disorders. In contrast, the TDM may lead to a more restricted research focus, including only speech tasks as the proper object of inquiry, or including only nonspeech tasks that are radically different both in terms of behavioral goals and in terms of motor demands. For instance, suppose AOS involves a deficit in the control of high-velocity movements, which affects speech but also certain other motor tasks. Excluding examination of nonspeech tasks as irrelevant to speech, or comparing speech to oral movements typically used to test for oral apraxia (e.g., puffing one's cheeks; Dabul, 2000), the conclusion might be that AOS represents a speech-specific deficit, rather than a deficit in control of high-velocity movements. One can only arrive at the latter conclusion by recognizing this property as relevant and consequently matching nonspeech tasks on this property. In this light it is important to consider that some influential models of motor control and learning propose units of action that specify relative timing requirements independent of the specific effector system that executes the movement (e.g., Schema Theory; Schmidt, 1975; Schmidt & Lee, 2005; see Section 3.2.1 for more details).

From this position, it is logical to apply the constructs and methods developed in the general motor behavior literature to the study and conceptualization of speech motor programming, which will be undertaken in the next chapter.



### **3.0. Models of Speech Motor Programming**

In this chapter, I will first present a model of spoken language production in order to ground the subsequent discussion of models of motor programming. Various models of (speech) motor programming in the hierarchical tradition will be discussed, as will their potential significance to understanding AOS. Specifically, the emphasis will be on a model that has focused on the units of action, namely Schema Theory (Schmidt, 1975, 1976, 2003; see Section 3.2.1), and a model that has focused on the processes involved in controlling serial order of movement sequences, namely the INT/SEQ model (Klapp, 1995, 2003; see Section 3.3.2).

#### **3.1. Spoken Language Production**

In production of speech, it is generally assumed that the intended message undergoes a number of transformations from more abstract to less abstract representations (e.g., Levelt et al., 1999; Dell, 1986), ultimately resulting in the coordinated sequencing of muscle contractions that generate speech. Several levels of representation have been proposed in this process (e.g., word form retrieval, phonological encoding), and although there is fairly wide consensus on the existence of these levels, as well as some of the details such as the distinction between frames and fillers and the phoneme as a fundamental processing unit (Dell, 1986; Garrett, 1982; Levelt, 1989; Shattuck-Hufnagel, 1979, 1992), the details of the representations and computations at those levels are subject of debate (e.g., Levelt et al., 1999 and commentaries; Rapp & Goldrick, 2000, 2004; Roelofs, 2004).

Major issues distinguishing the various models are the degree of interactivity (the degree to which two processes influence each other), both between processing levels and within processing levels, the cascaded vs. discrete nature of processing (the degree to which activation can spread to a subsequent level before a decision is made at the previous level), and the nature of representations (Levelt, 1999; Rapp & Goldrick, 2000, 2004). For example, Dell's (1986) spreading activation model of speech production involves semantic, morphologic, and phonologic processing levels with continuous interaction (feedback plus feedforward spread of activation) between levels, and includes explicit representation of syllable structure. In contrast, the Nijmegen model (Levelt et al., 1999) consists of a strictly serial (feedforward only), discrete model (only a selected lexical item activates its phonological form), in which syllable structure is not explicitly represented.

Any speech production model, and importantly any model that attempts to explain speech motor learning and breakdown, needs to make contact with other cognitive functions such as memory, perception, and motor programming. Unfortunately, current models are often restricted in scope (e.g., Croot, 2001; Kent et al., 1996; Ziegler, 2001). For example, Dell's (1986) model of speech production focuses almost exclusively on semantic, morphological and phonological stages of processing, involving units such as morphemes, syllables and phonemes, and largely ignores other cognitive and motor processing constraints. For this reason, I will discuss the Nijmegen model (Levelt et al., 1999) in order to ground the present work, since this model includes a post-phonological component.

### 3.1.1. The Nijmegen Model

Arguably the most comprehensive and detailed model of spoken language production is the Nijmegen model (Levelt, 1989; Levelt et al., 1999). The model is based largely on evidence from reaction time studies, unlike most other models of spoken language production, which rely heavily on speech error data (e.g., Dell, 1986; Garrett, 1982). This serial, strictly-feedforward model is relatively detailed with respect to stages such as lexical retrieval (response selection) and phonological and phonetic encoding (response programming).

Phonological encoding involves the simultaneous but separate retrieval of the segments (phonemes) and metrical frames (specifying the number of syllables, and for infrequent stress patterns, the stressed syllable), and the process of their integration (prosodification, Levelt et al., 1999: 22). Prosodification proceeds serially (left-to-right) to generate an output in terms of phonological words consisting of phonological syllables. Since prosodification is a serial process, the model requires temporary processing buffers (Levelt et al., 1999; Rogers & Storkel, 1998, 1999), in which material is stored until the last (right-most) segment has been inserted into the frame. The next stage can only begin once the preceding stage has produced a (minimal) output, in this case a phonological word.

The next stage is phonetic encoding (the model's motor programming stage, Levelt, 1989: 327), which involves the transformation of the abstract phonological words into a specification of the articulatory gestures (phonetic syllables); these gestures are stored in an articulatory buffer until all gestures for a particular

phonological word have been computed. These articulatory gestures are considered to be context-dependent and continuous, and gestures may overlap in time (cf. coarticulation), as opposed to the context-independent, discrete and static nature of phonological representations (Levelt et al., 1999). Gestures specify the goals to be achieved but they do not specify movement trajectories or positions per se. Phonetic representations are built up either by concatenating parts of the phonological representations (i.e. segments), or by retrieving stored phonetic syllables from a memory store called the syllabary (Levelt et al., 1999). It is assumed that frequent phonetic syllables become unitized, minimizing the need to create every phonetic syllable from the individual segments in the phonological representation each time it is needed. These two routes are thought to operate in parallel for all syllables, with the direct (syllabary) route being faster in most cases except for infrequent syllables (Levelt & Wheeldon, 1994: 265). It is further assumed that each segment activates all articulatory gestures it is associated with, and thus that there is competition between articulatory gestures (Levelt, 1999). Once all phonetic syllables are retrieved or constructed, the articulation system can execute the programs.

One primary source of evidence for the syllabary was a study by Levelt and Wheeldon (1994), in which they reasoned that the greater computational load for infrequent syllables should result in longer reaction times for these syllables than for frequent syllables. Crossing word frequency with syllable frequency in bisyllabic words, Levelt and Wheeldon observed independent and additive effects of both factors, suggesting that high-frequency syllables were indeed prepared faster than low-

frequency syllables, for both high-frequency and low-frequency words. However, syllable frequency was confounded with segment frequency, and some later studies in which this confound was removed failed to replicate the syllable frequency effects (Levelt et al., 1999), although another recent study did produce syllable frequency effects on naming latency (Cholin, Levelt, & Schiller, 2005).

Two other predictions from the notion of a mental syllabary have also been investigated. One is the prediction that for low-frequency syllables, naming latencies should be longer for more complex syllables (e.g., CCCVC) than for simpler syllables (e.g., CVC). Levelt and Wheeldon (1994) failed to find syllable complexity effects, whereas Santiago, MacKay, Palma, and Rho (2000) did observe such effects for syllable onsets. The other prediction is that coarticulation is predicted to be greater in syllables that are stored as fixed units than in syllables that are assembled from individual segments (e.g., Whiteside & Varley, 1998). Nijland, Maassen, Van der Meulen, Gabreëls, Kraaimaat, and Schreuder (2003) did not observe any differences in coarticulation between high- and low-frequency syllables in a group of typically developing children; however, it is possible that in adults, who have had more practice with speech, such differences would be found.

In short, empirical evidence for the syllabary is rather limited at present. In addition, there is an ongoing debate in the literature regarding the size of the units stored in such a memory store (e.g., Aichert & Ziegler, 2004; Varley, Whiteside, Windsor, & Fisher, 2006). Specifically, Varley et al. have argued that once frequency is assumed to be the determining factor for entry into the memory store, there is no

principled reason to restrict the store to syllable-sized units, but that words or even entire phrases may become stored as fixed units. Finally, although the model defines syllable motor programs in terms of articulatory goals such as lip rounding, the further details of subsequent spatiotemporal coordination of muscles are left to a relatively underspecified “articulatory network” (Levelt et al., 1999).

### **3.2. Units of Motor Control**

In connecting language production to the motor control system, one fundamental challenge is the problem of the units of motor control – that is, what are the primitives, or units, on which the motor control system operates, and how do these relate to units at preceding levels of processing? One goal of this dissertation is to further determine the nature and size of units of speech motor programming. In order to do so, I take the approach of borrowing constructs and methods from the general motor behavior literature.

An early definition of a unit of action or motor program was proposed by Keele (1968), who defined a motor program as “a set of muscle commands that are structured before a movement sequence begins, and that allows the entire sequence to be carried out uninfluenced by peripheral feedback” (p. 387). This definition implies that motor programs are specific to the muscles used for its execution, and faces what Schmidt (1975) termed the storage problem and the novelty problem. The storage problem refers to the fact that for each action, one would need a different program since no action is ever performed exactly the same, and thus the number of programs

that must be stored would have to be extremely large. The novelty problem refers to the fact that we can generate the same action in different contexts, with different effector systems, and that in this sense, each action is novel. It is difficult to see how a novel movement can be produced based on a fixed, highly specific motor program.

### 3.2.1. Schema Theory

To address the storage and novelty problems, Schmidt (1975, 1976) proposed the Schema Theory of motor control and learning, in which the *generalized motor program* (GMP) is the basic unit of motor behavior. A GMP is an abstract movement structure that specifies relative timing and relative force of muscle contractions, whereas the absolute timing, absolute force, and the specific structures or muscles to be used in the movement are specified by *parameters* (Schmidt & Lee, 2005). A general class of movements is governed by a single GMP, which can be scaled according to the current task demands, thus reducing the number of programs that need to be stored in memory.

Within Schema Theory, it is assumed that practice using a particular GMP with different absolute parameters strengthens the GMP by virtue of the development of so-called schemas, which are relationships between various sources of information (Schmidt, 1976, Schmidt & Lee, 2005). There are two types of schemas, namely recall schemas and recognition schemas. A recall schema is defined as the relationship between the motor specifications for reaching a particular desired outcome, the initial state of the muscles or structures, and the actual outcome of the movement. This

relationship becomes stronger with experience by temporarily storing these information sources after a movement and updating the relation between them. For future movements, the person can then supply the recall schema with the desired outcome and the initial state of the muscles, and the recall schema will produce the corresponding motor specifications that will lead to the desired outcome.

The recognition schema is defined as the relationship between the sensory consequences of a movement, the initial state and the actual outcome. As with the recall schema, the recognition schema becomes stronger with practice. For a movement, the recognition schema is used to predict the expected sensory consequences of that movement. By comparing the actual sensory feedback to the expected sensory consequence, errors can be detected, and this information can be used as the actual outcome information for updating the recall schema. In producing a movement using a particular GMP, the desired outcome combined with the initial conditions activate the appropriate parameters through the recall schema. A novel movement can thus be produced as long as the desired outcome, the initial conditions, and the motor program are known.

Evidence for generalized motor programs comes from a number of different sources, including EMG studies showing that when an arm-swing movement is blocked, the EMG patterns of the muscles are produced unaltered nonetheless, for up to about 110 ms after blocking (Wadman, Denier van der Gon, Geuze, & Mol, 1979). This suggests that the pattern of muscle activity is specified in advance and is executed as a single ballistic movement despite the presence of sensory feedback



blocking the actual movement, and argues against a dynamical systems perspective in which movements are not preprogrammed but rather emerge from interactions with the environment and the dynamics of the system. If the onset of the antagonist muscle is determined by the position of the structure (e.g., elbow angle), then the antagonist muscle should not have been activated when this position could not be reached (as in the case of external blocking). Here, the environment alters the dynamics, preventing the system from reaching the position (or phase angle) required to activate the opposing muscle. In addition, Carter and Shapiro (1984) showed that when a sequential arm movement was sped up in time, the relative duration of various movement components as well as of EMG patterns remained constant, providing support for the notion that motor programs are generalized.

A second line of evidence is based on the logic that within-unit variability should be smaller than between-unit variability. One way this has been operationalized is by computing correlations between kinematic landmarks derived from position, velocity, and acceleration records of repeated productions of a movement (e.g., Clark, Robin, McCullagh, & Schmidt, 2001; Heuer, Schmidt, & Ghodsian, 1995; Schneider & Schmidt, 1995; Wulf, Schmidt, & Deubel, 1993; Young & Schmidt, 1990). For example, Schneider and Schmidt (1995) used a coincident timing task in which participants moved a lever with their right arm, first to their right and then to their left (backswing component), and intercept a moving target light during the movement to the right; the lights could move either fast or slow. The participant's movement was captured, and essentially represented a time-position record, from which velocity and

acceleration traces could be derived. Kinematic landmarks were defined as peaks, valleys, and zero-crossings of the trace (e.g., at position-peaks, the velocity trace is around zero since the movement changes direction which requires slowing and then reversal of movement). If a part of the movement is governed by one unit, the correlations between the various kinematic landmarks that comprise the movement should approach 1 since they would have been programmed together. If a correlation between two or more landmarks is close to zero, this suggests that these parts of the movement were not programmed together, but rather are part of different units. Furthermore, a sudden drop in correlations between two successive landmarks would suggest a unit boundary. Using this approach, Schneider and Schmidt found that the backswing component of the movement (which involved a movement reversal) consisted of a single unit, which was then followed by the forward swing as a second unit. This approach has been supported for oral-facial movements (Clark & Robin, 1998; Clark et al., 2001).

Finally, investigations of motor learning have also provided support for the notion of GMPs as distinct from parameters by showing that various conditions of practice such as order of presentation and feedback frequency affect relative timing and amplitude (GMP) and absolute timing or amplitude (parameters) differentially (e.g., Shea, Lai, Wright, Immink, & Black, 2001; Shea, Wulf, Park, & Gaunt, 2001; Wulf & Lee, 1993; Wulf & Schmidt, 1989, 1996; Wulf et al., 1993; see Maas et al., in preparation, and Shea & Wulf, 2005, for a review). However, it should be noted that Shea and Wulf (2005), in a recent re-evaluation of Schema Theory, argued for the

term *scalable response structure* (SRS) rather than GMP, to reflect the idea that the relative timing aspect of movements is not restricted to short-duration movements but extends to longer movement sequences as well, where a stable (though not necessarily invariant) relative timing pattern emerges with practice.

### **3.2.2. Schema Theory and Speech Production**

Although Schema Theory has been applied to understanding of motor speech disorders, in particular AOS (see next section), its impact on models of speech motor control and learning has been rather limited. One major difficulty lies in specifying (and empirically verifying) the size and content of motor programs for speech in relation to the nature of the input. It has been noted previously that it is unclear which aspects are to be considered GMPs and which aspects can be considered parameters (e.g., Ballard et al., 2000, p. 984; Maas et al., in preparation). Indeed, the minimal unit of speech production, a likely candidate for a GMP, is as yet unknown. In the Nijmegen model, motor programs would be syllable-sized (Levelt et al., 1999; Aichert & Ziegler, 2004), which is consistent with the original notion of a GMP as governing relatively short-duration movements (on the order of about 200 ms), but there are other possibilities, including the phoneme (Van der Merwe, 1997; Rogers & Spencer, 2001; and the indirect route of Levelt et al., 1999), the word (Klapp, 2003), the stress group (Sternberg et al., 1978), and the phrase (Varley et al., 2006).

In terms of the content, GMPs for speech likely specify the relative timing and force of muscle contractions. For example, voice onset time (VOT) of syllable-initial

plosives might be specified as part of the GMP, with voiced and voiceless plosives being governed by separate GMPs (Ballard, Maas, & Robin, in press). A recent study that examined VOT in persons who stutter under different speaking rate conditions found that persons who stutter show abnormal absolute VOT values relative to control speakers, whereas their relative timing of VOT did not differ from controls (Max & Gracco, 2005). Such findings suggest that the concept of GMPs and parameters can be usefully applied to speech motor control, and that control of GMPs and parameters may dissociate in disorders of speech production.

Based on the conceptualization of GMPs outlined in the previous section, GMPs for speech would not include specification of effector (articulator); rather, the particular articulator to be used for execution of the action would be specified by an effector parameter. In this case, the syllables ‘sue’ and ‘shoe’ would be governed by the same GMP, with one different parameter setting (alveolar vs. palato-alveolar place of articulation). In other words, a GMP is centrally represented and selects a given structure to be used for its execution. This assumption would predict transfer across place of articulation within a manner class, a prediction that is borne out by some speech treatment studies (e.g., Ballard et al., in press).

However, there is also some evidence to suggest that when a particular structure is repeatedly used to execute a particular action, the biomechanical properties of this structure become part of the memory representation, as indicated by greater amounts of transfer across effectors with small than with large amounts of practice (Park & Shea, 2003, 2005). Certainly, we use speech so often that it is possible that

the particular articulator used to produce a given phoneme becomes an integral part of a new GMP; in this case, motor programs for speech would be generalized in the sense that all productions of a particular syllable are governed by the same GMP.

While GMPs have been operationally defined in the motor literature in terms of coherence between various kinematic landmarks (e.g., Schneider & Schmidt, 1995; Young & Schmidt, 1990; Section 3.2.1), these measures have largely been applied to relatively simple movements involving a limited number of degrees of freedom (e.g., an arm-swing movement; Schneider & Schmidt, 1995). The difficulty with respect to speech is that it is by nature a motor skill that requires coordination across a wide range of structures, and the measurement of all potentially relevant structures poses daunting, if not insurmountable, challenges. While kinematic analyses have been applied to speech production (e.g., Abbs, Gracco, & Cole, 1984; Robin, Bean, & Folkins, 1989; Smith, Johnson, McGillem, & Goffman, 2000), these studies typically involve only a small number of structures (e.g., upper lip, lower lip, jaw). As such, important relations between structures and systems may be overlooked, especially with respect to the tongue, perhaps the most important articulator, but also with respect to respiration and phonation. In this dissertation, a reaction time approach is adopted, since this methodology has been shown to be sensitive to the number of units (as discussed in more detail in Sections 3.3.1 and 3.3.2).

In sum, Schema Theory offers a detailed account of how discrete movements are controlled and learned, and in particular it offers perhaps the best-defined unit of action, namely the generalized motor program (GMP). The notion of GMPs provides a

mechanism that in principle allows one to capture similarities between speech and nonspeech movements. For instance, in this view there is an abstract level of representation at which relative timing and force are specified independent of effector system, and thus it is possible that movements that share relative timing patterns are controlled by the same GMP, regardless of whether the GMP is implemented with a finger or with a jaw movement. However, the precise size and content of GMPs as applied to speech is as yet unclear, and as a result the impact of Schema Theory on understanding of speech motor control has been limited. Nonetheless, the general consensus that speech production involves units such as phonemes, syllables, and words suggests that a motor program approach to speech motor control is viable. In this dissertation, the syllable and the (pseudo) word are examined as possible units of speech motor control.

### **3.3. Processes of Motor Control**

In addition to units of action, a theory of motor control must also specify the processes involved in motor programming, including those processes that govern the serial ordering of movements. In contrast to the structuring of individual motor programs, the processes responsible for sequencing successive motor programs has received comparatively little attention in Schema Theory. One influential model to account for serial ordering in speech and finger movements was proposed by Sternberg and his colleagues (1978). A more recent model (the INT/SEQ model; Klapp, 1995, 2003) that incorporates many of the features of the Sternberg model will

form the main focus of this dissertation; however, given its significant historical role the Sternberg model will first briefly be discussed.

### **3.3.1. The Sternberg Model**

The Sternberg model was based on findings from a number of simple reaction time (RT) studies. In a simple RT paradigm, a participant is informed of the target response by a cue, which is then followed after a delay by an imperative signal that prompts the participant to produce the specified response as quickly and accurately as possible. In a series of experiments, Sternberg et al. (1978) found that the latency to begin articulation increased linearly with the number of items to be produced, as did the increase in mean item duration (producing a quadratic total sequence duration function). In the speech experiments, they observed these effects when the responses consisted of digits or weekdays, regardless of order or number of different words. Sternberg et al. also examined finger movements, i.e. typing letters, and observed the same latency and rate effects as a function of the number of keystrokes.

To account for these findings, Sternberg et al. (1978) proposed that motor programming involves loading a representation of the entire movement sequence (a motor program, in their terminology) into a motor buffer. This motor buffer is thought to be distinct from short-term memory (Sternberg et al., 1978: 133) based on the finding that an additional short-term memory load did not affect the latency and rate functions. A motor program in this model consists of a set of subprograms that make up the entire sequence, and the motor program is loaded during the interval between

the cue that specifies the response and the imperative signal; the subprograms are not ordered in the buffer.

When the imperative signal arrives, the first subprogram is located and retrieved from the buffer (retrieval stage), and its muscle commands are unpacked (unpacking stage), and then executed (command stage). The retrieval stage is thought to involve a search through a nonshrinking buffer, since simple RT varied with sequence length. If the first subprogram could be accessed directly, or if it held some special status (if it were “on top of the pile”, so to speak), then simple RT should be independent of sequence length. The assumption of a nonshrinking buffer was made to account for the fact that the slopes of the latency function and the rate function were linear and similar to one another. If production of a subprogram were to result in its removal from the buffer, then each subsequent subprogram would have fewer competitors and thus the slope of the rate function would be smaller than that of the latency function.

The unpacking stage is assumed to depend on the complexity of the subprogram, not on the number of subprograms in the buffer, and the retrieval and unpacking of subprograms are thought to be independent processes. This conclusion was based on an experiment in which the number of words was crossed with the number of syllables per word (e.g., limb-limit, cow-coward, rum-rumble). The results showed that mean simple RT (i.e. the intercept) was longer for disyllabic words than for monosyllabic words, but that the slope of the sequence length function (in terms of number of words) was identical for monosyllabic and disyllabic words. In other



words, the effects of number of words and number of syllables were additive. These results suggest, first, that retrieval and unpacking are independent processes, and second, that the subprogram appears to be the size of a word rather than a syllable (otherwise the number of words and the number of syllables should have interacted, with a greater sequence length effect for disyllabic words).

In a follow-up experiment it was found that padding a list of nouns by inserting unstressed function words such as *and*, *or*, and *of* also increased the mean latency but not the slope of the latency function. Based on these findings, Sternberg et al. tentatively assumed that a subprogram consisted of a stress group (one primary stress plus unstressed syllables); unpacking the first subprogram in the sequence takes longer for more complex subprograms (in this case, those containing more syllables). Thus, simple RT includes a component related to the number of words in the buffer (retrieval stage) and a component related to the complexity of the first subprogram (unpacking stage). Complexity presumably also affects unpacking of subsequent subprograms, but these effects are observable not on simple RT but during execution (i.e. on the rate function). The observation that the effect of subprogram complexity is greater on execution rate than it is on latency suggests that the command stages for successive subprograms are sequential and separated by retrieval and unpacking operations (which take longer in longer sequences).

While the Sternberg model has been very influential in the motor programming literature, one major limitation of the model is that it is based on findings from a simple RT paradigm only, and as such cannot provide a full picture of all processes involved

in programming the response. In particular, the loading or constructing of the motor program (the overall sequence representation) is presumed to occur prior to the imperative signal, i.e. it is preprogrammed, and therefore is not reflected in simple RT (e.g., Klapp, Abbott, Coffman, Greim, Snider, & Young, 1979). Klapp et al. replicated some of the Sternberg et al. (1978) results, namely the observation that simple RT varied with the number of identical digits to be pronounced (e.g., one-one-one), and further noted that the sequence length effect disappeared when speakers were required to count to a specific number starting at 1 (e.g., one-two-three). Klapp et al. reasoned that if the sequence length effect observed in simple RT were due to programming the articulation of the upcoming response, then the same effect should be observed when counting to a given number. The fact that this was not the case suggested to Klapp et al. that the sequence length effect was not due to programming but rather due to different demands on short-term memory or planning when to terminate the response.

Klapp et al. (1979) argued from these findings that choice RT is a more appropriate paradigm to study motor programming. In a choice RT paradigm, the imperative signal specifies the response, making advance preparation impossible, and therefore, choice RT is thought to reflect programming of the response. However, Klapp et al. failed to find a sequence length effect on choice RT, which they suggested was due to speakers programming the sequences in segments rather than as a whole, with programming of the second unit occurring during execution of the first unit, etc. (cf. cascaded processing; Rapp & Goldrick, 2000).

There has been considerable debate in the literature about which paradigm is most appropriate to study response programming, either simple RT (e.g., Marteniuk & MacKenzie, 1981; Sternberg et al., 1978) or choice RT (e.g., Klapp et al., 1979; Klapp & Greim, 1981; Sheridan, 1981). Many have noted (e.g., Kerr, 1978; Klapp et al., 1979; Marteniuk & MacKenzie, 1981; Sternberg et al., 1978) that there are several possible confounds in choice RT, such as stimulus processing and response selection variables. For Klapp and his colleagues, this underscores the importance of including appropriate experimental controls, whereas others argue that choice RT should not be used to study motor programming (e.g., Sternberg et al., 1978; Marteniuk & MacKenzie, 1981). As we shall see in the next section, more recently Klapp (1995, 2003) proposed a motor programming model that incorporates findings from both simple and choice RT paradigms.

In sum, based on findings from a simple RT paradigm, Sternberg et al. (1978) proposed a hierarchical, three-stage serial model of motor programming that included 1) a retrieval stage (affected by the number of subprograms in the buffer but not by subprogram complexity), 2) an unpacking stage (affected by the complexity of the subprogram but not by the number of subprograms), and 3) a command or execution stage (affected by subprogram complexity but not by number of subprograms). Applied to speech, they further proposed that the subprogram consists of a stress group, rather than a word or syllable.

There has been controversy over the appropriate RT paradigm to study motor programming. One major argument against the exclusive use of simple RT is that in a

simple RT paradigm the participant can preprogram the upcoming response, thereby placing much of the motor programming stage outside the RT interval. Finally, the model has not been applied to understanding of motor speech disorders, especially AOS.

In the next section, a recent model of motor programming will be discussed that forms the primary focus of this dissertation, and which integrates findings from both simple and choice RT paradigms (Klapp, 1995, 2003). Moreover, there has been some initial application of the model to AOS (Deger & Ziegler, 2002), providing further motivation to explore the model as applied to speech and AOS.

### **3.3.2. The INT/SEQ Model**

The INT/SEQ model originally proposed and developed by Klapp (1995, 2003) includes two processes, namely INT and SEQ. The INT process represents a preprogramming stage in which the internal structure of a program is activated prior to the onset of the movement and stored in a motor buffer until needed. A basic assumption is that the demands placed on the INT process increase with increasing complexity of the program. The next stage of processing is the SEQ process during which the sequence of multiple programs is assembled on-line to control serial ordering, by locating and retrieving the first item from the buffer. The SEQ process cannot be preprogrammed, and the load on the SEQ process increases with an increasing number of units in the buffer. The INT process is consistent with Schema Theory, in that activation of a GMP and assignment of appropriate parameters would

all occur as part of the INT process as these aspects of programming relate to the internal structure of a single program. The SEQ level of the model would then be used to assemble the individual GMPs into the proper sequence. Finally, the model assumes that with extended practice, a sequence of units can become integrated, or “chunked”, into a single unit, essentially transferring the control of serial order from SEQ to INT.

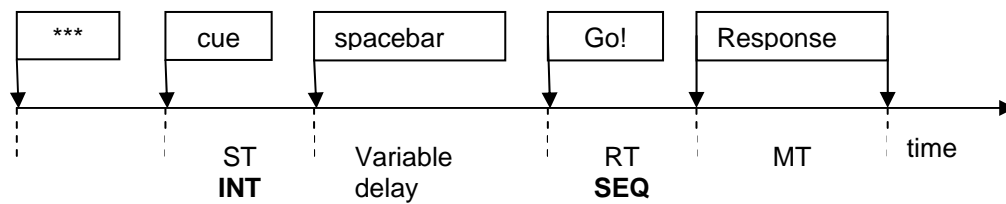
Support for the model has come from reaction time studies involving simple and choice RT, to index SEQ and INT, respectively (e.g., Klapp, 1995, 2003; Klapp & Wyatt, 1976; Khan, Lawrence, Buckolz, & Franks, 2006). Since SEQ cannot be preprogrammed, the number of units in a response should affect simple RT. Unit complexity effects, on the other hand, should be observed in choice RT only. Using button press responses, with single button presses of short (150 ms) or long duration (450 ms) and sequences of four button presses, Klapp (1995) found that a single long button press resulted in longer choice RT than a short button press did, whereas duration did not affect simple RT. This finding suggested that a single button press constituted a unit whose complexity could be defined in terms of its duration (longer movement is more complex), and moreover, that the internal features of a single unit (in this case, its temporal requirements) can be preprogrammed.

Conversely, comparing single button presses (short and long) to sequences consisting of short and long presses (short-long-long-short and long-short-short-long), Klapp (1995) found that the number of button presses affected simple RT (longer RT for sequences than for single button presses) but not choice RT. This pattern of results further supported the notion that button presses were units and that the SEQ process is

sensitive to the number of units. The absence of a sequence length effect on choice RT was explained by assuming that the INT and SEQ processes occur in parallel during choice RT and that the INT process takes longer to complete (Klapp, 1995), so that only INT effects are visible in choice RT. A recent study involving rapid aiming movements in simple and choice RT for one-target and two-target responses provided further support for the model (Khan et al., 2006). Finally, Klapp (1995) showed that after 8 days of practice on these sequences, the sequence length effect on simple RT disappeared, supporting the notion that the button presses were chunked into a single unit.

These findings have been replicated using the “self-selection paradigm” (see Figure 3-1), a modified reaction time paradigm that allows measurement of both INT and SEQ on each trial, and thus within subject (Immink & Wright, 2001; Wright, Black, Immink, Brueckner, & Magnuson, 2004). In the self-selection paradigm, subjects are given a cue that specifies the response required on the current trial (as in simple RT), and are asked to prepare the response as much as they can. They press a button to indicate that they are ready to respond, and this interval between cue onset and button press is termed Study Time (ST), which captures the INT demands. After subjects press the button, they withhold the response until the imperative signal, which appears after an unpredictable delay. They initiate movement as soon as possible upon presentation of the go-signal. The time from the imperative signal until the onset of the response is termed reaction time (RT) and indexes the SEQ process (as in simple RT). Immink and Wright (2001) and Wright et al. (2004) replicated the Klapp findings

using this self-selection paradigm: button press duration affected ST but not RT, and the number of presses affected RT early in practice but not after extensive random practice. In addition, interpress intervals became less variable, suggesting formation of a single larger unit from the smaller parts (Wright et al., 2004).



**Figure 3-1.** Self-selection paradigm (Immink & Wright, 2001). See text for details.

However, contrary to Klapp's (1995) findings using choice RT, Immink & Wright (2001) and Wright et al. (2004) also obtained a sequence length effect on their measure of INT (ST), in addition to a sequence length effect on RT. This suggests that sequences resulted in additional processing during preprogramming, at least in the self-selection paradigm. One hypothesis about the nature of this additional processing is that the sequences comprised two distinct units, and that each unique unit was loaded during INT (Immink & Wright, 2001: 436). An alternative hypothesis is that all units, rather than all *unique* units, were loaded during INT (Magnuson, Robin, & Wright, in press).

A recent study by Magnuson et al. (in press) addressed these two hypotheses by comparing sequences consisting of identical button presses (i.e. either all short or all long button presses). The rationale was that if only unique units are

preprogrammed, then there should be no difference between sequences consisting of one, two, or three button presses. If, on the other hand, all units are preprogrammed, a sequence length effect should emerge on ST. The findings revealed additive effects of sequence length and button press duration on ST, supporting the hypothesis that INT involves preprogramming all units in these sequences. An alternative account that claims that the increase of ST was caused by an additional, separate “multiplier” process that specifies the number of repetitions of a single preprogrammed units (e.g., Magnuson et al., in press) was ruled out by the fact that differences were observed between sequences of two presses and sequences of three presses, rather than just between single and multiple presses. Presumably, loading a multiplier would only occur for sequences, but the value of the multiplier should not affect processing time since this would negate the potential benefit of such a process.

Thus, it appears that INT involves preprogramming all units in a sequence. It is important to note here that there was also a sequence length effect on RT in these studies (Immink & Wright, 2001; Magnuson et al., in press; Wright et al., 2004), which suggests that even though all units were preprogrammed, these units were not integrated into a single unit. In other words, when examining the nature of motor programming and in particular the processes that occur during preprogramming, it is important to consider the entire pattern of results. If a sequence length effect is observed on ST and RT, this suggests preprogramming of all units without integration; if a sequence length effect is observed on ST only, then this suggests that all elements of a sequence are preprogrammed and integrated during the INT stage.



### 3.3.3. The INT/SEQ Model and Speech Production

Recently, Klapp (2003) has extended the INT/SEQ model to speech production and examined the nature of speech motor programs by comparing simple and choice RT paradigms. Klapp argued that if words constitute motor programs, then the number of syllables per word should define unit complexity, and thus, number of syllables should affect choice RT (INT) but not simple RT (SEQ). If, on the other hand, syllables are motor programs, then one would expect to see an effect of number of syllables on simple RT. In other words, simple RT provides a way to determine whether a sequence of movements represents a series of units or a single unit.

In Klapp's (2003) experiment 1, he demonstrated that the number of syllables in items presented as pseudowords affected choice RT but not simple RT, supporting the hypothesis that motor programs for speech are word-sized, not syllable-sized. These results are in agreement with those from earlier studies in which the number of syllables affected choice RT but not simple RT (e.g., Klapp, Anderson, & Berrian, 1973; Klapp, 1974). Choice RT effects of number of syllables have also been found in more recent studies involving picture naming (e.g., Meyer, Roelofs, & Levelt, 2003; Santiago, MacKay, Palma, & Rho, 2000) and nonword production (Deger & Ziegler, 2002). In addition, a study by Schönle, Hong, Benecke, and Conrad (1986) showed that choice RT was longer for sequences consisting of different syllables (e.g., /badaga/) than for sequences of repeating syllables (e.g., /bababa/), suggesting that the number of different syllables may also define complexity of a unit.

Note that Klapp's (2003) simple RT findings contrast with those obtained by Sternberg and colleagues (1978) who did observe an effect of complexity (number of syllables) on simple RT mean. However, the syllable effect observed by Sternberg et al. was rather small (appr. 5 msec), and it is possible that the absence of an effect reflects lack of power (e.g., Sternberg et al., 1978: 129). Another possibility is that when using small set sizes, speakers are more likely to integrate sequences into a single unit. In Klapp's experiment 1, each subject produced only three different items, whereas the Sternberg et al. experiment involved a set size of 72 nouns in various combinations.

Klapp (2003) further showed that when integration of syllables was discouraged by cueing each syllable separately during the foreperiod and by including a larger set size (all possible combinations of three syllables), the number of syllables affected simple RT, as predicted by the model. In another experiment (experiment 4), speakers were required to produce the same syllable, /bi/, one, two, or three times in a row in either a choice RT or a simple RT paradigm. The results revealed a sequence length effect on simple RT but not on choice RT, which suggests that repeated syllables are not integrated into a single word-sized unit. Similar findings were obtained by Deger and Ziegler (2002; see Chapter 4, section 4.2.3).

The suggestion that a sequence of syllables is processed by one mechanism in some cases (INT for different syllables) but by another mechanism in other cases (SEQ for repeating syllables) requires an explanation, as it calls into question the basic assumptions of the model, i.e. that INT is sensitive to complexity of a single unit and

SEQ is sensitive to the number of units. The use of number of syllables as a metric both of complexity and of sequence length exposes the threat of circular argumentation: INT is sensitive to unit complexity but not to number of units, and thus if INT is affected by the number of syllables then number of syllables must define complexity of a single unit; SEQ is sensitive to the number of units but not to unit complexity, and thus if SEQ is affected by number of syllables then number of syllables must define the number of units of the sequence.

This issue points to the interdependence of INT and SEQ: though INT and SEQ are separate, they are not independent because failure to integrate movement elements into a single unit during INT will increase the processing load for SEQ, i.e. INT and SEQ are in a trading relationship. The basic assumptions of the model can only be maintained by assuming different unit sizes for different sequences of the same number of syllables (e.g. /dadada/ is three units, /dabaga/ is one unit). However, flexible unit size is used as an adhoc solution and impedes the falsifiability of the model, unless an independent, principled account can be provided. If we accept the assumption that choice RT captures unit complexity and simple RT captures the number of units, then the basic question is why only sequences of different syllables become integrated into a single unit, while sequences of repeated syllables are programmed as separate units. Neither Klapp (2003), nor Deger and Ziegler (2002) provided a satisfactory explanation for this discrepancy. This issue will be addressed in Experiments 3 and 4 of this dissertation (see Chapter 7).

In order to further detail the SEQ process, specifically whether SEQ requires knowledge of the particular elements to be executed or not, Klapp (2003) conducted a set of partial precue experiments, in which the number of repetitions of a syllable (e.g., three) was specified in advance but the specific syllables to be produced were specified by the imperative stimulus (e.g., B, leading to Bibibi in this example). In other words, this paradigm was similar to the simple RT condition in that the number of syllables was specified in advance, and similar to the choice RT condition in that the particular syllable to be produced was cued by the imperative stimulus. The results showed that RT increased with the number of syllables to be produced, which is inconsistent with the notion of INT and SEQ occurring in parallel (as proposed in the original formulation of the model; Klapp, 1995), because in that case INT should take longer than SEQ thereby masking the SEQ process. Since complexity of the first element was identical across sequence lengths, there should have been no difference in RT between single syllables and sequences of syllables.

To account for these findings Klapp (2003) revised his original model by assuming that a sequence of units is controlled by an abstract temporal frame; this temporal frame specifies the initiation times of each unit independent of its content, and can be loaded into the buffer during preprogramming. The frame must be scanned to locate the first unit upon initiation of the response, and it is assumed that this scanning process (= SEQ) takes longer when there are more slots in the frame. The idea of scanning a temporal frame is very similar to the proposal made by Sternberg et al. (1978); the critical difference is that the temporal frame in the INT/SEQ model

does not contain effector-specific information. A further assumption is that in a choice RT task, no buffering is required and therefore no scanning (SEQ) is needed to locate the first unit, which explains why there is no effect on choice RT of the number of units (Klapp, 2003). In choice RT, the temporal frame is loaded in parallel with the INT process, but since the movement must be executed as soon as possible, the first unit can be inserted into its slot and executed immediately. Note that this interpretation implies that in addition to INT and SEQ, there is an additional process, namely loading of a temporal frame into the motor buffer that supports the serial ordering of the sequence elements (Magnuson et al., in press).

The notion of an abstract temporal frame is reminiscent of the metrical frame proposed by Levelt et al. (1999). According to Levelt et al., the metrical frame of a word is presumed to specify the number of syllables and, for words with irregular stress patterns, the location of primary stress (Levelt et al., 1999). Evidence for metrical frames comes from implicit priming studies (Roelofs & Meyer, 1998). In their implicit priming paradigm, Roelofs and Meyer grouped responses for a block of trials in sets of words with similar numbers of syllables or similar stress patterns (homogeneous sets), or in sets that differed in number of syllables or stress pattern (heterogeneous sets). In essence, this task is a partial pre-cue paradigm in which the partial precue is not specified before each trial but derives from the grouping of responses into homogeneous sets. Roelofs and Meyer found that reaction times were faster when the words appeared in the homogeneous sets than when these same words appeared in heterogeneous sets, which was interpreted to mean that metrical structure

(in terms of number of syllables and stress pattern) is indeed represented at the level of phonological encoding (Roelofs & Meyer, 1998; Levelt et al., 1999).

However, while the similarities between Klapp's (2003) abstract temporal frames and Levelt et al.'s (1999) metrical frames appear obvious, the studies supporting each of them differ in several ways, disallowing any firm conclusions at this point. For example, Klapp (2003) investigated and found a length effect (longer RTs for longer items) but did not investigate (or report on) the stress patterns of pseudowords, whereas Roelofs and Meyer (1998) investigated and found priming for words with similar metrical frames but did not investigate length effects per se. In addition, as also pointed out by Schiller, Fikkert, and Levelt (2004), the implicit priming effect found by Roelofs and Meyer depended on the homogeneous sets sharing both metrical structure as well as initial segments, weakening the case for completely separate representation of frames and fillers (cf. also Varley & Whiteside, 2001b: 80).

Schiller et al. (2004) addressed whether metrical structure can be primed independent of the segments, using a picture-word interference paradigm in which Dutch speakers named pictures while a (similar or dissimilar) word was presented aurally. In a series of experiments with different stimulus-onset-asynchronies (SOAs) between picture onset and auditory stimulus onset, Schiller et al. found no evidence for metrical priming, and thus no support for the idea that metrical frames are independent of the segments. Thus, it is possible that the metrical frames in the Nijmegen model and the temporal frames in the INT/SEQ model are different in that

metrical frames may be tied to the specific segments that fill the frame, whereas the temporal frame in the INT/SEQ model is independent of its fillers.

It should be noted that the speech findings reported by Klapp (2003) could also be explained in reference to phonological constructs such as syllables and phonemes, since the motor programs were in fact defined in phonological terms. While this is a convenient strategy to manipulate motor programming complexity (Maas et al., 2002), it means that effects are essentially ambiguous in terms of their phonologic or motoric nature. The generally accepted separation between phonological encoding and motor programming in most models of speech production suggests that it will ultimately be important to tease these factors apart. For example, one could find variables that are thought to affect each stage differently and directly compare their effects in a single experiment.

Psycholinguistic models postulate abstract, context-independent, discrete units such as phonemes or syllables (e.g. Levelt et al., 1999; Dell, 1986; Garrett, 1982) which may be seen as conceptually distinct from more contextualized, continuous motor patterns. These models include a level of processing at which phonemes are activated and sequenced, and point to speech errors of serial order such as anticipatory substitutions (e.g., *dig date* for *big date*) and transpositions (e.g., *dig bate* for *big date*) as evidence for the phoneme (e.g., McNeil et al., 2000). However, the notion that relatively abstract motor programs are sequenced is capable of explaining serial order errors equally well (e.g., Mowrey & MacKay, 1990; Wertz, LaPointe, & Rosenbek, 1984) and has been used to explain such errors in non-speech actions.

For the purpose of this dissertation, I shall assume the basic distinction between phonological planning and motor programming, acknowledging that the empirical evidence for their separation is not unequivocal (e.g., Rogers & Storkel, 1999: 796). Although teasing apart phonological planning and motor programming is not the primary focus of this dissertation, an attempt was made to examine a variable that can be argued to be strictly motoric without obvious phonological consequences. The rationale here is that the only way to study motor programming uncontaminated by phonological planning effects is to keep the phonological input to the motor system constant and vary only motor-level variables. This argument does not imply that phonologically-defined variables will not affect motor programming; it is assumed that they do (or can). Given that phonological constructs are conceived of as abstract and static in nature (Levelt et al., 1999: 31), the motoric variable studied here was absolute duration of a movement, specifically of a CV syllable in which vowel length is not phonemically contrastive in English. Absolute duration of a movement has been shown to affect the INT process in finger movements, and thus provided an a priori viable candidate for examination of the INT process in speech production free from phonological confounds.

In sum, the INT/SEQ model postulates two programming processes, namely a preprogramming process that organizes the internal structure of a unit (INT), and a sequencing process that places units in their correct serial order (SEQ). The model has been supported by a range of reaction time studies with unimpaired individuals, and has been applied to both speech and nonspeech motor programming. However, the



INT and SEQ processes require further specification, for example with respect to preprogramming of longer sequences, and the factors that underlie integration of successive gestures into a single unit. In speech, it appears that SEQ involves activating and scanning an abstract temporal frame, but the exact nature of this frame remains to be explicated further. At the very least, it appears that such a temporal frame specifies the number of syllables; it remains to be seen whether stress patterns (e.g., relative timing) are also specified in these frames.

### **3.4. Summary**

The topic of this dissertation is speech motor programming, with a particular focus on the units and processes involved, and on the nature of neurogenic breakdown of speech motor programming, i.e. AOS. In Chapter 2, several major issues in motor control, including speech motor control, were discussed, namely the degrees of freedom problem and the serial order problem. A common solution to these problems is to assume that patterns of muscle commands can be structured in advance as a single unit, whose order is controlled by a higher level process. Specification of the units and processes involved in generating actions such as speech is a major, and challenging task for models that adopt this assumption of a hierarchical system.

In Chapter 3, several models of motor programming were discussed in relation to speech production. It was noted that even relatively comprehensive models of spoken language production such as the Nijmegen model (Levelt et al., 1999) do not provide much detail about the units and processes of speech motor programming, and

that there is debate about the size of speech motor programs. To remedy this situation, I suggested that it is useful to look to the general motor behavior literature since this provides us with constructs and methods that facilitate the study and understanding of speech as a motor skill and allows for continuity between different motor skills.

The model that forms the primary focus of this dissertation is the recently proposed two-stage INT/SEQ model (Klapp, 1995, 2003), which was based on reaction time evidence from both simple and choice RT paradigms with unimpaired participants. The model distinguishes between a preprogramming process called INT that structures the internal features of a motor program (e.g., its temporal structure), and an on-line process called SEQ that sequences successive units into their correct serial order. The model has been applied both to nonspeech movements (e.g., button presses: Klapp, 1995; aiming movements: Khan et al., 2006) and to speech movements (Klapp, 2003).

Several key findings supporting the model have been replicated for the finger movements using the self-selection paradigm, a modified RT paradigm that allows assessment of both the INT and the SEQ process on each trial (e.g., Immink & Wright, 2001; Wright et al., 2004). Specifically, effects of unit complexity (press duration) were found on Study Time (ST), the paradigm's index of INT, whereas sequence length effects were observed on RT, the paradigm's index of SEQ. One interesting difference with previous work using choice and simple RT was that a sequence length also emerged on ST, suggesting that for sequences, more than one unit can be preprogrammed. Whether performers preprogram all units (Magnuson et al., in press),

or only all unique units (Immink & Wright, 2001) is an open question, although there is evidence to suggest that in finger movements, performers preprogram all units in a sequence (Magnuson et al., in press).

The application of the model to speech production has indicated that words (or pseudowords) are programmed as units (Klapp, 2003; Sternberg et al., 1978), whose internal complexity can be defined in terms of the number of different syllables (Klapp, 2003; Schönle et al., 1986; Sternberg et al., 1978). It was noted that these effects could have their origin at a higher, phonological level of processing, although the assumption in most models of spoken language production that the output of phonological planning constitutes the input to the motor system implies that these variables (e.g., number of syllables) will also impact motor programming. In order to validate the distinction between phonological planning and motor programming, it will be necessary to identify variables thought to affect each stage differently and directly compare them in a single experiment. This enterprise was not undertaken here; instead, this dissertation represents the first study to determine whether a variable that can be thought to be motoric but not phonological in nature (absolute duration) produces reliable effects on reaction time. This variable was chosen based on findings from the limb motor programming literature.

One potentially problematic finding for the INT/SEQ model as applied to speech is that the same variable (number of syllables) affects choice RT in some cases (different syllables) but simple RT in other cases (repeated syllables). In other words, based on the logic of the model this means that the serial order of successive syllables

is programmed by the INT process in the former case but by the SEQ process in the latter case. While the assumption of flexible unit size has been made by others as well (e.g., Levelt et al., 1999; Varley et al., 2006), the factors responsible for this flexibility must be independently established in order to avoid the assumption becoming an ad hoc solution that undermines the falsifiability of the model. In Experiments 3 and 4, two hypotheses regarding the factors responsible for processing a sequence as a single unit are directly contrasted. The first hypothesis claims that integration into a single unit occurs to control coarticulation and/or prevent serial order errors; the second hypothesis claims that integration occurs only when there is a rhythmic structure to the sequence.

While the INT/SEQ model has been replicated using the self-selection paradigm for finger movements, this has not been done for speech production. One goal of this dissertation was to extend the application of the self-selection paradigm to the study of speech production, in order to assess both INT and SEQ for each trial. Furthermore, in Experiments 3 and 4, patterns of results are compared between choice RT and Study Time (ST), in an effort to cross-validate paradigms.

#### **4.0. Apraxia of Speech (AOS)**

Apraxia of speech (AOS) in adults is an acquired motor speech disorder typically caused by a stroke affecting the left frontal cortices (Duffy, 2005), although primary progressive AOS as a result of a progressive illness also occurs (Duffy, Peach, & Strand, 2006). Darley and colleagues (Darley, Aronson, & Brown, 1975) first defined AOS as a disorder of speech motor programming. However, due to the relatively underdeveloped models of motor programming at the time, a debate ensued in the literature that has led to considerable confusion regarding the nature of AOS. An influential paper by Martin (1974) questioned the use of the term apraxia of speech to describe the condition, given that linguistic constructs such as phonemes were used to describe speech errors made by these patients. While Martin's argument has been (mis)interpreted as a denial of the existence of AOS or that apraxia is a form of aphasia (McNeil, Doyle, & Wambaugh, 2000), his argument was merely that a different set of descriptors was needed if one wanted to maintain a distinction between AOS and phonological paraphasias (cf. McNeil et al., 2000; Rosenbek, 2001).

#### **4.1. Characteristics of AOS**

Nowadays, there is a consensus among clinicians and researchers about the fact that AOS is indeed a disorder of speech motor programming, as originally conceived by Darley et al. (1975), as well as about the symptoms that are indicative of AOS. These defining characteristics of AOS are increased segment durations, increased intersegment durations (including the intrusive schwa), speech sound

distortions, errors that are consistent in type and position within the utterance (though the actual error may vary from trial to trial) and dysprosody (difficulty with stress placement with a tendency toward equal stress and segmentation) (McNeil et al., 1997, 2000; Wambaugh, Duffy, McNeil, Robin, & Rogers, 2006). Normal or fast rate are exclusionary criteria, as is normal prosody (Wambaugh et al., 2006). Other errors may be present but do not differentiate between AOS and phonemic paraphasia or dysarthria; these include initiation difficulties, articulatory groping, and sensitivity to automaticity and length of the utterance (Kent & Rosenbek, 1983; McNeil et al., 1997, 2000; Nickels, 1997). Finally, features that cannot be used to define AOS because they are more likely to reflect impairments at other levels in the system include serial order errors such as perseverative and anticipatory substitutions and metathetic errors (Wambaugh et al., 2006).

Evidence for these characteristics comes from perceptual, acoustic, kinematic, and physiologic studies. For example, McNeil, Odell, Miller, and Hunter (1995) conducted a perceptual analysis of three repeated productions of target words by speakers with AOS, conduction aphasia and ataxic dysarthria. These researchers found that location of error within an utterance and error type were consistent for speakers with AOS but not for speakers with conduction aphasia. Acoustic analyses have provided further evidence regarding the increased segment- and intersegment durations. For example, Kent and Rosenbek (1983) analyzed the speech of seven speakers with AOS and unimpaired control speakers, and observed reduced speech rate with longer segment durations and transitions and longer pauses between

syllables, decreased variation in intensity across stressed and unstressed syllables, incoordination between voicing and articulation, and slow and inaccurate movements of articulators for both consonants and vowels. In another acoustic study, Seddoh, Robin, Sim, Hageman, Moon, and Folkins (1996) analyzed four words in a carrier phrase (“That’s a pop/pea/Bob/bee a day”) repeated ten times by speakers with AOS, speakers with conduction aphasia, and unimpaired control speakers. Seddoh et al. focused on temporal measures, and found that relative to the control speakers, speakers with AOS exhibited longer and more variable vowel duration, consonant-vowel duration, and stop-gap duration (period of closure for a stop consonant such as ‘p’), suggesting that precise motor control of temporal aspects may be disrupted in AOS (Seddoh et al., 1996). Seddoh et al. also found that perceptual ratings of intelligibility were lower for speakers with AOS than for the other groups, who did not differ.

Kinematic studies have shown reduced peak articulatory velocity and inconsistent timing between the lip, tongue and velum (Itoh, Sasanuma, & Ushijima, 1979) and greater lip and jaw displacements (McNeil, Caliguiri, & Rosenbek, 1989). However, Robin, Bean, and Folkins (1989) investigated peak articulatory velocity for the upper and lower lip with and without a biteblock, and found that speakers with AOS did not differ from control speakers in terms of peak velocity or temporal coordination of the lips. Robin et al. suggested that interarticulator coordination for the lips may be less prone to disruption, and furthermore that the reduced speech rate often reported may not bear a direct relationship with articulatory velocity. Indeed,

Forrest, Adams, McNeil, and Southwood (1991) have found inter-articulatory coordination to be impaired in AOS compared to normal speakers. As well, Kelso and Tuller (1981) reported particular difficulty with temporal coordination of speech movements in AOS.

Taken together, perceptual, acoustic, and kinematic studies suggest a motor-level impairment in AOS, reflected in temporal and spatial abnormalities. Such temporal and spatial abnormalities cannot be ascribed to a phonological encoding system, since at this level of processing the units are assumed to be unspecified for kinematic variables. Thus, these symptoms address Martin's (1974) objections against the use of linguistic constructs to describe AOS, and clearly implicate the motor system in this disorder. In turn, a motor-level impairment would suggest that nonspeech motor control may also be impaired.

Indeed, several studies to date have provided evidence for nonspeech motor impairments in individuals with AOS. For example, McNeil, Weismer, Adams, and Mulligan (1990) investigated the control of isometric fine force and static position, tasks that require sustained force or position at a specified level (high or low). Participants were asked to maintain a given force level or position with the upper lip, lower lip, jaw, tongue, and finger, by matching their output force level or position (measured by transducers connected to the structures and displayed on an oscilloscope) to the target force or position, also indicated on the oscilloscope. The results showed that for fine force control as well as static position control, individuals with AOS and individuals with ataxic dysarthria were more variable than control



speakers for nearly every structure (including finger). Individuals with conduction aphasia did not differ from either the control speakers or the speakers with AOS and those with ataxic dysarthria. McNeil et al. concluded that at least some individuals with AOS exhibit more general motor control impairments, and furthermore that such impairments are also apparent in non-oral structures such as the finger.

Further evidence for nonspeech motor impairments comes from studies employing visuomotor tracking tasks (e.g., Ballard & Robin, in press; Hageman et al., 1994; Robin et al., submitted). As discussed in Section 1.4, patients with AOS were unable to track predictable patterns compared to control participants and to patients with conduction aphasia, while tracking of unpredictable patterns was normal (Hageman et al., 1994). This suggested that patients with AOS were unable to create a motor program for the predictable movement, and the implication is that the disorder in AOS is not confined to speech but also affects other motor control tasks that share certain properties with speech. For instance, the tracking task discussed above (e.g., Hageman et al., 1994) shared two important features with speech, i.e. its dynamic (as opposed to static) nature and the fact that peak velocity occurs in the center of the movement (Gracco & Abbs, 1986).

In a more recent visuomotor tracking study involving eight individuals with AOS, Ballard and Robin (in press) further tested the notion that AOS involves a difficulty in the development of feedforward control mode. Using predictable targets with and without continuous visual feedback for jaw movement, Ballard and Robin reasoned that if new programs can be developed but are poorly specified, then

continuous visual feedback should improve performance (i.e. removal of the target pattern should decrease error and variability), whereas if new programs can be developed but feedback cannot be adequately processed to modify the ongoing movement, then removal of the target pattern should lead to performance enhancement. The results showed that there was no difference between feedback and no-feedback conditions, but that performance on both was impaired relative to unimpaired controls, suggesting that both specification and implementation of a new motor program was impaired. Visual feedback did improve the spatial accuracy (in terms of amplitude) for both groups, but not the temporal accuracy in the AOS group, indicating difficulties in control of timing of movements.

Further support for nonspeech deficits in AOS comes from a study by Clark and Robin (1998), who employed a task in which participants were required to reproduce a regular target movement pattern with their lower lip and jaw after it had been removed from the screen. This task, developed by Wulf et al. (1993) for limb movements, used an analysis which was able to separate GMPs from parameters of movements. Clark and Robin found that their speakers with AOS were less accurate in reproducing either the GMP or the parameters, but not both, relative to control speakers, while speakers with conduction aphasia performed similarly to the control speakers.

In short, the current consensus among researchers and clinicians is that AOS reflects a disruption of motor programming (e.g., Ballard et al., 2000; Code, 1998; Deger & Ziegler, 2002; McNeil et al., 1997, 2000; Rosenbek, 2001; Varley &

Whiteside, 2001a,b; Wambaugh et al., 2006), although differences exist with respect to the exact nature of this motor programming deficit. An influential definition of AOS was provided by McNeil et al. (1997):

Apraxia of speech is a phonetic-motoric disorder of speech production caused by inefficiencies in the translation of a well-formed and filled phonological frame to previously learned kinematic parameters assembled for carrying out the intended movement, resulting in intra- and interarticulator temporal and spatial segmental and prosodic distortions (p. 329).

This definition captures the consensus that impairments of phonological processing are not part of the disorder in AOS (though such impairments may co-occur with AOS), and are in agreement with the findings from perceptual, acoustic, and kinematic studies of speech in AOS reviewed above, which suggest a disorder of speech motor programming that involves intra- and inter-articulator spatiotemporal coordination. In addition, there is evidence to suggest that the impairment in AOS is not confined to speech production. However, further specification of the exact nature of the motor programming deficit in AOS depends on the development and application of models of motor programming.

## **4.2. AOS in Relation to Models of Speech Motor Programming**

In order to further delineate the precise mechanisms disrupted in AOS, a model-driven approach is required to generate testable hypotheses. In this section, several accounts of AOS framed within the models discussed in Chapter 3 will be related to the characteristics of AOS outlined above.

### **4.2.1. The Nijmegen Model and AOS**

Within the Nijmegen model, one proposal to account for AOS places the locus of the deficit in the phonetic encoding stage of the model (Varley & Whiteside, 2001a,b; Whiteside & Varley, 1998). Specifically, Varley and Whiteside (2001a,b) have proposed that AOS represents a deficit to the direct route of phonetic encoding (i.e. impaired access to the mental syllabary), forcing speakers with AOS to rely on the indirect, phoneme-by-phoneme route of phonetic encoding. This indirect route is more resource-intensive and error-prone, resulting in a greater number of speech errors, loss of automaticity, prolongations or segments, reduced coarticulation, increased variability, and reductions in speech rate.

Several substantive arguments against this view have been noted (e.g., Ballard, Barlow, & Robin, 2001; McNeil et al., 2000; Rogers & Spencer, 2001). For example, this account predicts that high-frequency words or syllables should be produced equally well or poorly as low-frequency items, since both would rely on the indirect route. However, the often-reported finding that speakers with AOS have more difficulty with less frequent items than with highly frequent items suggests is at odds

with this prediction (Rogers & Spencer, 2001). In fact, Aichert and Ziegler (2004) showed that speakers with AOS do show sensitivity to syllable frequency, in that high-frequency syllables were produced with fewer errors than low-frequency syllables. Aichert and Ziegler interpreted these findings as reflecting intact access to the mental syllabary that contains damaged syllable motor programs.

Another argument against the dual route hypothesis of AOS is that if the indirect route is intact, then presumably speakers with AOS should be able to generate speech at the same level of efficiency as unimpaired speakers producing novel words, and this is clearly not the case (e.g., McNeil et al., 2000; Ziegler, 2001). In order to account for this discrepancy, Varley and Whiteside assume that the indirect route will in many cases also be impaired to some extent, perhaps due to deficits in resource allocation resulting from brain damage (Whiteside & Varley, 1998: 223) or due to neural proximity of damaged regions (Varley & Whiteside, 2001: 44), so that compensation for damage to the direct route will be incomplete (Whiteside & Varley, 1998). However, this reduces the elegance of the proposal, since many of the symptoms are now to be accounted for by mechanisms other than the presumed core deficit in AOS (Ziegler, 2001).

Another account of AOS within the Nijmegen model relates to the notion of temporary processing buffers (Rogers & Storkel, 1998, 1999). Specifically, Rogers and Storkel (1999) propose that AOS reflects a limitation of the buffer capacity to a single syllable. This proposal accounts for core symptoms such as syllable segregation, dysprosody, and slow speech rate. Rogers and Storkel (1999) examined

five patients with AOS and aphasia, five patients with aphasia but no AOS, and five unimpaired control speakers, on a task involving rapid recitation of two-word sequences. Previous studies had shown that when speakers produce single words presented one at a time in rapid succession, there is a cost (longer naming latency) when successive words share place and manner of articulation, relative to when the words do not share features (Rogers & Storkel, 1998). This increase in naming latency was interpreted as reflecting inhibition of articulatory gestures when the buffer had to be cleared and reprogrammed with a similar gesture. In their AOS study, Rogers and Storkel (1999) reasoned that presenting two words at a time would allow unimpaired speakers to preprogram both words, thereby eliminating the similarity effect. However, if speakers with AOS are limited to programming only a single syllable, then a similarity effect should still be present as production would involve reprogramming the buffer. Their findings supported the hypothesis, in that indeed a similarity effect was found for the speakers with AOS but not for speakers without AOS (with or without aphasia).

While both the dual route account of AOS and the reduced buffer capacity account of AOS can account for some of the characteristics of AOS, sound distortions are not easily captured. In addition, neither account provides an explanation or mechanism for the consistent presence of nonspeech motor deficits in patients with AOS, as both accounts are grounded in a psycholinguistic model of speech motor programming.

#### 4.2.2. Schema Theory and AOS

Schema theory has been applied to understanding AOS by various authors (e.g., Ballard et al., 2000; Clark & Robin, 1998; Kent & Rosenbek, 1983; McNeil et al., 2000; Robin, 1992). For example, Ballard et al. (2000) suggested that AOS may reflect a disruption of activating and/or selecting the relevant GMPs, a disruption of parameterization, or both. This idea was supported by the nonspeech oral movement study by Clark and Robin (1998) discussed in Section 4.1. These findings, together with the visuomotor tracking results discussed above (e.g., Ballard & Robin, in press; Hageman et al., 1994; Robin et al., submitted), support the notion that retrieving, selecting, and/or parameterizing GMPs is impaired in AOS. Thus, AOS can be viewed as the manifestation in speech of a more wide-spread problem with accessing, selecting or parameterizing GMPs for oral movements (e.g., Ballard et al., 2000).

An alternative hypothesis that has been proposed is that the (syllable) programs themselves are somehow degraded (Aichert & Ziegler, 2004). Evidence for damaged motor programs comes from an analysis of speech errors produced by apraxic speakers, who were found to show a syllable frequency effect, with fewer errors on syllables with a very high frequency (consistent with the notion that frequent syllables are retrieved as pre-stored units) (Aichert & Ziegler, 2004). In other words, the frequency effect suggested that pre-stored syllables were retrieved from the syllabary, but the errors on these items suggested that the information stored in the programs was “corrupted” (Aichert & Ziegler, 2004: 156). While it is not specified exactly how the programs are damaged, this proposal places the locus of the deficit in the structural

integrity of the motor programs, as opposed to the retrieval or selection of the programs. However, an analysis of speech errors alone cannot exclude the possibility that retrieval is also impaired, and conversely, an analysis of performance on a nonspeech task alone cannot determine whether the mapping between linguistic and motor levels is disrupted in AOS. Therefore, it will be necessary to investigate speech production itself with respect to both retrieval and execution processes.

A third possible alternative is that sensory information about the initial conditions or the consequences of the movement are unavailable in AOS (e.g., Kent & Rosenbek, 1983). Schema Theory suggests that in order to produce a movement, the person must select the relevant GMP based on the goal as well as the initial state of the muscles. Thus, if sensory information about the initial conditions is not available, the appropriate GMP may not be retrieved. Similarly, if sensory feedback from the produced movement is unavailable after the movement, no schema updating can occur. This hypothesis compatible with the finding that learning (as measured by transfer) typically observed in treatment studies for AOS is relatively limited (e.g., McNeil et al., 2000). The findings by Clark and Robin (1998) and Hageman et al. (1994) that speakers with AOS appeared unable to develop a GMP for the predictable target pattern suggests that motor learning may indeed be impaired as well. Presumably, developing a GMP for a movement constitutes an essential part of motor learning – without the movement specifications represented in the GMP no schema updating can occur.



At present however, this hypothesis is empirically indistinguishable from the hypothesis that GMP retrieval or parameterization is impaired, because both hypotheses can account for the limited generalization and the failure to track predictable patterns, as well as for the speech characteristics. Similarly, the findings that when the target tracking pattern is removed, speakers with AOS resort to a preferred tracking frequency (Ballard & Robin, in press) and that speakers with AOS do appear to have access to at least some (high-frequency) syllables (Aichert & Ziegler, 2004) can be explained under both views.

#### **4.2.3. The INT/SEQ Model and AOS**

Based on the evidence reviewed in Section 4.1, a primary hypothesis to be addressed in this dissertation is that AOS reflects an impairment of process INT but not the SEQ process. Specifically, features characteristic of AOS, such as the prominence of speech sound distortions, the temporal and spatial incoordination and variability (difficulty in organizing the internal structure of units), and dysprosody (difficulty with integrating units), all point to an impairment of processes involving (pre)programming. Further, the absence of serial order errors such as anticipatory and perseverative substitutions and phoneme transpositions (no difficulty with sequencing) suggest that the SEQ process is intact. This hypothesis is also consistent with the frequently observed articulatory groping, although this feature is not considered specific to AOS.

Further support for this hypothesis derives from findings that nonspeech motor control is also impaired in AOS when tested carefully (e.g., Clark & Robin, 1998; Hageman et al., 1994; McNeil et al., 1990). As argued above, these studies with carefully selected patients with AOS revealed that the control of static and dynamic aspects of nonspeech motor control are also impaired in AOS. In addition, the hypothesis of an INT impairment in AOS is consistent with several of the accounts of AOS discussed above, in particular those accounts framed within Schema Theory that suggest impaired retrieval, integrity or parameterization of GMPs (e.g., Aichert & Ziegler, 2004; Ballard et al., 2000; Clark & Robin, 1998; McNeil et al., 1997).

The hypothesis of a disruption of INT but intact SEQ in AOS makes a number of predictions. First, it is expected that individuals with AOS will take longer to complete the INT process than controls. Second, time to complete the SEQ process should not differ from controls. Third, individuals with AOS should show disproportionate effects of complexity on the INT process, relative to controls. Fourth, it is possible that if individuals with AOS fail to integrate multiple syllables into a larger word-sized unit, then there should be a greater load on the SEQ process for sequences.

Deger and Ziegler (2002) conducted the only study to date to apply the INT/SEQ model to speech motor programming in AOS. They investigated speech production of speakers with AOS and aphasia, speakers with aphasia but without AOS, and control speakers in a simple RT task using sequences of two and three syllables ('dada', 'data', 'daba', and 'dadada'). Responses were produced in blocks,

with each block containing 15 trials, and the experiment was repeated on three separate days, resulting in 45 trials for each response. Deger and Ziegler found an RT difference between /dada/ and /dadada/ for control speakers and speakers with aphasia but not for apraxic speakers; conversely, they found a difference in RT between /dada/ on the one hand and /data/ or /daba/ on the other hand for apraxic speakers but not for control speakers or aphasic speakers.

Notice first that the results for control speakers were consistent with those obtained by Klapp (2003), in that a sequence length effect was observed for repeating syllables. This confirms the idea that repeating syllables are not integrated into a single word-sized unit. However, notice further that the absence of a difference between repeating and alternating syllables of the same length is inconsistent with the INT/SEQ model: if repeating syllables such as /dada/ are not integrated, the SEQ load should consist of two units in this case, whereas sequences such as /daba/ are assumed to result in an integrated representation resulting in an SEQ load of one unit. Thus, the model would predict that the alternating sequences should be initiated faster than the repeating sequences. Though not significant, the numerical difference was in fact in the opposite direction. However, it could be argued that two-syllable sequences are integrated whereas sequences of three repeating syllables are not.

Turning to the predictions for AOS, only predictions 2 and 4 could be evaluated, since Deger and Ziegler (2002) used only simple RT. Regarding prediction 2 (no simple RT differences), the findings were mixed. For the analysis involving sequence length, there was indeed no group difference between individuals with AOS

and controls. However, for the syllable alternation analysis, the speakers with AOS did show longer RT than controls, although this difference was caused primarily by longer RTs to alternating sequences rather than repeating sequences, as indicated by a significant group by task interaction. This interaction speaks to prediction 4; specifically, it confirms the idea that these speakers with AOS were unable to integrate the alternating syllables into a single unit, thus resulting in a greater load on SEQ. Deger and Ziegler suggested that their apraxic speakers had difficulty with alternating syllables (switching between different movements), and that these speakers attempted to produce syllable sequences as single “entities”. However, they were unable to do so for the alternating sequences, and therefore decomposed these into separate units, resulting in longer simple RTs (as a result of sequencing). Thus, it was assumed that integration of syllable sequences is more demanding for alternating syllables than for repeated syllables. The difficulty with alternating syllables was further supported by a comparison of mean inter-syllable pause duration, which revealed that pauses between alternating syllables were longer than pauses between repeated syllables.

Although the Deger and Ziegler (2002) study provided initial evidence to suggest that INT is disrupted in AOS, it is important to point out that INT processing was not assessed directly and in isolation from SEQ processing (e.g., by examining single syllables). Furthermore, only group data were presented, and it is important to inspect individual patterns. In Experiments 1 and 2, the hypothesis of a deficit in INT in AOS is tested directly, by examining movement sequences in the context of the self-selection paradigm which allows measurement of both INT and SEQ on each trial.

### 4.3. Summary

With respect to the second aim of increasing our understanding of motor speech disorders, the discussion of AOS in this chapter indicated that this syndrome is best considered a disorder of speech motor programming. Although the speech-specific nature of this disorder remains a subject of debate, the growing body of evidence for nonspeech motor control deficits in AOS, the uncertainty about the proper definition of speech, and the philosophical position that in order to understand the nature and speech-specificity of the deficit in AOS one must examine nonspeech motor control, all suggest that the models, constructs, and methods from the motor behavior literature may be usefully extended to the conceptualization and investigation of this disorder. Conversely, the study of AOS as a motor programming disorder may provide further support for and extension of current models of motor programming.

Although the study of AOS has relied on a variety of measures, including perceptual, acoustic, and kinematic, these measures are all based on the final output of the system and potentially confound motor programming with motor execution difficulties. Given that one key aspect of motor programming is the fact that it can occur prior to the initiation of movement, it is logical to pursue measures of the processes preceding initiation in AOS. That is, a motor programming deficit should be reflected in processing time measures during the interval in which programming is thought to occur, rather than exclusively in the final product of that process. This dissertation represents one of the few studies of AOS to take such a time course approach. The novelty of the present work is that it extends the self-selection

paradigm to speech production in AOS, making it the first study to include measures both of the putative motor programming stages in the recent INT/SEQ model, as well as of the execution of the programmed movements.

## **5.0. Speech Motor Programming in AOS and Aphasia**

Experiment 1 was designed to address the nature of the speech motor programming deficit in AOS by investigating speech movements from a time course perspective, within the context of the recent INT/SEQ model of motor programming. The primary purpose of this study was to test the hypothesis that AOS involves a deficit in the INT process, whereas the SEQ process is thought to be intact. By extending the INT/SEQ model to disordered populations, we may be able to further delineate the nature of motor programming deficits in AOS, as well as obtain neuropsychological evidence about the separation between the INT process and the SEQ process. In addition, a secondary purpose of this experiment was to replicate previous findings using the self-selection paradigm, which allows measurement of both INT and SEQ within the same subject on each trial. Thus far, the model, especially as related to the speech findings, has been based on between-subject comparisons and comparisons between different paradigms. Thus, this experiment was intended to establish the robustness of the findings on which the INT/SEQ model is based.

To test the hypothesis of an INT deficit in AOS for speech movements, Experiment 1 examined speech movement patterns with identical temporal requirements as those used in the original finger movement studies on which the INT/SEQ model is based (e.g., Klapp, 1995; Immink & Wright, 2001). As discussed in Chapter 3, the INT/SEQ model has been shown to apply to speech as well as finger movements (e.g., Deger & Ziegler, 2002; Klapp, 2003). Typically, the word (Klapp,

2003) or stress group (Sternberg et al., 1978) has been assumed as the unit of speech, and complexity has been defined in terms of the number of syllables in a (pseudo)word (e.g., Deger & Ziegler, 2002; Klapp, 2003). One finding that has emerged from these studies is that alternating syllable sequences (e.g., ‘tegabi’) are treated as single units and thus load on the INT process, not the SEQ process (Klapp, 2003), while repeating syllable sequences (e.g., ‘bibibi’) are treated as separate units and thus load on the SEQ process, not the INT process (Deger & Ziegler, 2002; Klapp, 2003; but see Sternberg et al., 1978).

However, while the number of syllables presumably affects the complexity of motor programming, this manipulation also affects complexity at the level of phonological encoding, and thus it is not clear that observed effects should be attributed (solely) to the level of speech motor programming. In the present experiment, a response factor was manipulated whose effects can be isolated to the motor programming stage (i.e. timing). Specification of temporal parameters is thought to be the domain of motor programming rather than phonological encoding (e.g., Levelt et al., 1999; Van der Merwe, 1997). Timing of a simple CV syllable (/ba/) was chosen as the response factor, since vowel length is not phonemically contrastive in this context in English (unlike in certain CVC syllables, where vowel length differentiates between voicing of the final consonant; Ladefoged, 2001). Specifically, this experiment used the exact same temporal patterns as those used in the original finger movement studies that provided the basis for the INT/SEQ model. For optimal comparison to the finger movement studies, the syllable sequences consisted of



repeating syllables varying only in duration (comparable to pressing the same button with the same finger).

In order to exclude the possibility that concomitant aphasia, rather than AOS, would be responsible for any potential deficits, a small number of individuals with aphasia but without AOS was also included for comparison purposes. Note that subjects with pure or primarily AOS (and not aphasia) are extremely infrequent, thus the low number of subjects in the AOS group. For the individuals with aphasia, 12 were contacted but only 4 were willing to participate in this study. For these individuals, both INT and SEQ are hypothesized to be intact.

Thus, the main purpose of this experiment was to provide a direct test of the hypothesis of an INT deficit in AOS using specification of timing as the motor programming variable of interest. A secondary purpose was to establish the utility of the self-select paradigm to speech production research.

## **5.1. Methods**

### **5.1.1. Participants**

Four patients with AOS, four patients with aphasia, six age-matched healthy control speakers, and fifteen young adult control speakers were recruited for this study. All participants read and signed an informed consent form, and all procedures were approved by the local Institutional Review Board. One age-matched control speaker was excluded based on a self-reported history of stuttering, and three young adult control speakers were excluded (2 non-native English speakers, 1 with a

neurological history). Thus, the total sample for analysis consisted of four patients with AOS, four patients with aphasia, five age-matched control speakers, and twelve young adult control speakers. Background information for these participants is provided in Table 1. Prior to the experiment, all patients were tested for aphasia using the Boston Diagnostic Aphasia Examination (BDAE; Goodglass & Kaplan, 1983) or the Western Aphasia Battery (WAB; Kertesz, 1982), for AOS using the Apraxia Battery for Adults – 2 (ABA-2; Dabul, 2000), for dysarthria using an oral motor examination (e.g., Duffy, 2005), and for auditory perceptual discrimination using the auditory word discrimination subtest of the Psycholinguistic Assessment of Language Processing in Aphasia (PALPA; Kay, Lesser & Coltheart, 1992). Test results are provided in Appendix A.

Participants with AOS. All patients with AOS were diagnosed as such by two certified speech-language pathologists experienced in motor speech disorders, based on presence of the characteristics listed by McNeil et al., (1997, 2000) as indicative of AOS, i.e. slow and effortful speech, speech sound prolongations and distortions, sound and syllable segmentation, and dysprosody. In addition, all patients with AOS were classified by the ABA-2 as having apraxia of speech. There were three men and one woman in the AOS group, with a mean age of 57 years (range 27 to 72 years) and a mean time post onset of 4 years and 2 months (range 2;9 to 6;8).<sup>1</sup>

---

<sup>1</sup> Several of the patients included in this study have also been part of other studies in our lab. In the interest of full disclosure, we list here the cross-references: AOS1 is Subject 2 in Austermann Hula et al. (subm.); AOS3 is Subject 4 in Austermann Hula et al. (subm.); AOS4 is Participant 1 in Ballard et al. (in press).

**Table 5-1.** Participant information for Experiment 1. AMC = age-matched controls; YCON = young adult controls; TPO = time post onset (years; months).

	<b>Sex</b>	<b>Age</b>	<b>Hand</b>	<b>TPO</b>	<b>Aphasia</b>	<b>AOS</b>	<b>Oral/Limb apraxia</b>	<b>Dysarthria</b>
AOS1	M	69	L	3;7	Mild nonfluent	Mild-mod.	None/None	Mild unilateral weakness
AOS2	M	27	R	6;8	Mild-mod. nonfluent	Mild	None/None	None
AOS3	F	72	R	3;7	Mild anomia	Mild-mod.	None/None	None
AOS4	M	59 <i>57 (21)</i>	R	2;9 <i>4;2</i>	Mild anomia	Mild	None/None	Mild unilateral weakness
APH1	M	74	R	2;6	Mod. fluent	None	Mild/None	None
APH2	M	55	L	7;0	Mild-mod. fluent	None	Mild/None	Mild unilateral weakness
APH3	F	63	R	1;9	Very mild fluent	None	None/None	None
APH4	M	62 <i>64 (8)</i>	R	9;6 <i>5;3</i>	Mild anomia	None	None/None	Mild unilateral weakness
AMC1	M	65	R					
AMC2	M	72	R					
AMC3	F	72	R					
AMC4	F	54	L					
AMC5	M	57 <i>64 (8)</i>	L					
YCON (N=12)	11F 1M	24 (5) (20-35)	10R 1L, 1A					

AOS1 is a 69-year old left-handed man who suffered a single left-hemisphere stroke 3 years and 7 months prior to the onset of the experiment. He is an English-Spanish simultaneous bilingual, and up to his CVA was a college professor in Spanish literature. He considers English to be his primary language, and uses English almost exclusively since his stroke. In addition to a mild to moderate AOS, he also exhibited a mild nonfluent aphasia (BDAE language competency index = 81); his reading ability was intact. There was no evidence for oral or limb apraxia, but he did demonstrate a mild right-sided weakness of the lower face.

AOS2 is a 27-year-old right-handed man who suffered a left hemisphere stroke during surgery for a congenital heart problem, more than 6 years before the experiment. His native language is English, and prior to his CVA he was a college student. His AOS was rated as mild, and there was no evidence of oral or limb apraxia or dysarthria. However, he did have right-sided paresis in the extremities. In addition to AOS, he also had a mild to moderate nonfluent aphasia (WAB Aphasia Quotient = 69.3), characterized mainly by word finding problems and agrammatism. Reading abilities were within the range needed for this experiment.

AOS3 is a 72-year-old right-handed woman, who was a retired manager of data processing before her single left-hemisphere stroke, 3 years and 7 months prior to the onset of the experiment. Her native language is English. In addition to a mild to moderate AOS, she exhibited a mild anomia (WAB Aphasia Quotient = 95.5). Her reading ability was intact. There were no signs of oral or limb apraxia, nor was there evidence for dysarthria.

AOS4 is a 58-year-old right-handed man with a Ph.D. in physics who was a college professor and Vice President of a communications firm prior to his left-hemisphere stroke, which occurred 2 years and 9 months prior to the experiment. His native language is English. In addition to a mild AOS, he had a very mild anomia (WAB Aphasia Quotient = 97.2). Reading abilities were intact. There was no evidence for oral or limb apraxia, but he did demonstrate a mild right-sided weakness of the lower face and the extremities.

Participants with aphasia. Four individuals with aphasia were also tested, to address the possibility that brain damage, or aphasia would account for any potential differences between unimpaired controls and participants with AOS. The average age of the aphasia control group was 64 years (range: 55 to 74), and the average time post onset was 5 years and 3 months.

APH1 is a 74-year-old right-handed man who suffered a single left-hemisphere stroke 2 years and 6 months prior to the experiment. His native language is English and he continues to practice law as an attorney. He had a fluent aphasia of mild severity (AQ = 89.6) characterized mainly by semantic and phonological paraphasias, circumlocutions and mild auditory comprehension deficits. Reading abilities were adequate for this experiment. He exhibited no evidence of AOS, dysarthria, or limb apraxia, but he did have a mild oral apraxia.

APH2 is a 55-year-old left-handed man whose native language is English. He suffered a single left-hemisphere CVA in the frontal cortex and basal ganglia 7 years before the experiment. His mild-to-moderate fluent aphasia was mainly characterized

by word-finding problems, occasional semantic and phonological paraphasias and pronoun errors, and largely intact auditory comprehension. Reading abilities were within the normal range. There was no evidence for AOS or limb apraxia. However, there was a mild oral apraxia and mild right-sided weakness of the lower face.

APH3 is a 63-year-old right-handed woman, who experienced a left-hemisphere CVA 1 year and 9 months prior to this experiment. She had a fluent aphasia that was rated as very mild (AQ = 96.9) and that was characterized by mild word-finding problems during conversational speech and occasional semantic paraphasias. There was no evidence of AOS, oral or limb apraxia, or dysarthria. While she produced several graphemic errors in writing, her reading abilities were intact. Her history also included a traumatic brain injury approximately 15 years prior to the experiment.

APH4 is a 62-year-old right-handed man who has a degree in civil engineering and who is a pastor in a local church. He suffered a left-hemisphere CVA almost 10 years before the experiment. His aphasia was characterized mainly by mild word finding difficulties, occasional phonemic paraphasias, and mild comprehension problems for sequential commands. Reading abilities were intact, and there was no evidence for AOS or oral or limb apraxia. However, he did demonstrate evidence of mild right-sided weakness of oral structures.

Control participants. Two control groups were included in this experiment. An age-matched control group consisted of five individuals with no history of neurological, speech, or language problems. There were two women and three men

(mean age: 64; range: 54 to 72); three were right-handed and two were left-handed. In addition to the age-matched control group, twelve healthy younger adult controls participated as well (mean age = 24, SD = 5; range = 20-35). The young control group consisted of eleven women and one man; ten participants were right-handed, one was left-handed, and one was ambidextrous. They were all monolingual English speakers without neurological history, speech or language impairments, or uncorrected visual or auditory impairments. All were college students who participated for partial course credit.

### **5.1.2. Materials and Equipment**

All target responses in this experiment consisted of the syllable /ba/. Responses differed in terms of the number of syllables (sequence length: 1 or 4) and the duration of the syllables (short: 150 ms, long: 450 ms). Each response was paired with a visual symbol that functioned as the cue in the experiment. Cues were 1S (single short: 'ba'), 1L (single long: 'baaa'), 4S (SLLS sequence: 'babaaabaaaba'), and 4L (LSSL sequence: 'baaabababaaa'). Auditory models were created for each response based on recordings of by a male native speaker; both syllables were produced to approximate the target duration as closely as possible, to ensure naturalness (the drop in fundamental frequency was 7 Hz for the short syllable and 10 Hz for the long syllable). Recordings were digitized at 20 kHz with 10 kHz low-pass (antialias) filtering. The waveform was then imported into a sound editing software program (Adobe Audition v.1.5; Adobe Systems, Inc.) and the vowel was truncated to produce

the exact target syllable duration (from release of stop to vowel offset: 150 ms or 450 ms). Next, the 4S and 4L sequences were constructed from these single syllables, with 100 ms pauses in between, resulting in a total duration of 1500 ms for each of the two sequences. Finally, each response was recorded onto both channels (22.05 kHz, 16-bits) for presentation during the experiment.

The experiment was controlled using E-Prime software, version 1.1 (Psychology Software Tools, Inc.), and was run under Windows XP on a Dell desktop computer (Optiplex GX 280) for the young control participants and on a Dell notebook computer (Inspiron 600m) for the patients and the age-matched controls. Speech onset latencies were collected with the voicekey feature of the Serial Response Box (Psychology Software Tools, Inc.) using an omnidirectional lapel condenser microphone. All sessions were recorded at 44.1 kHz with a DAT recorder (model Sony PCM-M1) using a separate microphone of the same model. Throughout the experiment, both microphones were kept in place at a constant mouth-to-microphone distance of about 5 cm on the left side using a padded, adjustable headset.

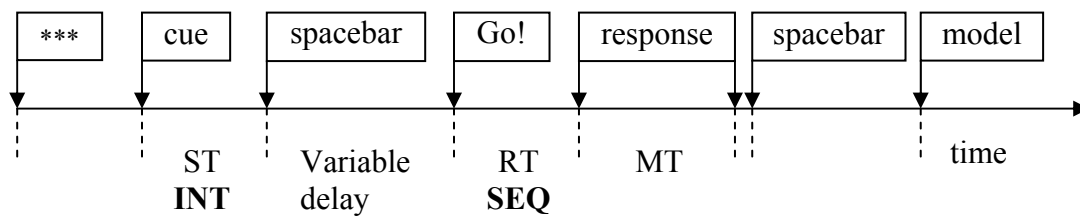
### **5.1.3. Task and Procedures**

The experimental task was the self-selection paradigm (see Figure 5-1). Specifically, the sequence of events was as follows. Each trial started with presentation of a warning signal (three asterisks, \*\*\*) in the center of the screen for 1000 msec. The warning signal was then replaced by the cue (1S, 1L, 4S, or 4L) in the center of the screen. At this time, the participant was to prepare the indicated response



as much as possible without producing it. The cue remained on the screen until the participant pressed the space bar with their left index finger to indicate readiness to respond. The time between presentation of the cue and the participant's press of the space bar was termed Study Time (ST), and constituted our measure of the INT process. The space bar press removed the cue from the screen and initiated a delay interval, which varied in duration (between 800 and 1200 msec, as in Wright et al., 2004) to minimize anticipatory responses. Next, a go-signal (the word "Go!") was presented in the center of the screen, accompanied by a 75 ms tone of 2000 Hz. Participants were instructed to initiate the prepared response as quickly as possible after the go-signal. The go-signal remained on the screen until a vocal response was detected by the voicekey. Time between presentation of the go-signal and onset of speech as detected by the voicekey was called Reaction Time (RT) and captured the SEQ process.

Upon completion of the response, the participant pressed the space bar again to call up the auditory model of the target response. The auditory model was presented after each response, and was preceded for 2 seconds by the message "The correct response is ..." in the center of the screen; this message remained on the screen during the auditory model. After the model, the experimenter judged the accuracy (correct/incorrect) of the response against the target by pressing one of two keys on the button box (see below for criteria). After a 2-second intertrial interval, the next trial started. All trials involved white stimuli against a black background, using bold Arial font (point size 24).



**Figure 5-1.** Self-selection paradigm adapted to represent trial events as used in Experiment 1. ST = Study Time; RT = Reaction Time; MT = Movement Time. See text for details.

The experiment involved an acquisition phase and a retention phase, each conducted on separate days (24 or 48 hours apart). Stimulus presentation in both phases was in random order. The acquisition phase was preceded by presentation of the auditory models of each response. Each model was presented four times in blocked order; the order was the same for all participants (1S, 1L, 4S, 4L). The acquisition phase consisted of 12 blocks of trials, in which each of the four responses was to be produced correctly four times (for a minimum of 16 trials per block). As in previous studies using the self-selection paradigm, incorrect trials were rerun at the end of each block until criterion was reached to ensure an equal number of observations per condition, and to ensure that all analyses of the processing time measures (ST and RT) were based only on correct trials. However, in order to avoid the potential of infinite block size and resulting frustration for the patients due to consistent errors, an upper limit was set so that the total number of trials in a block would never exceed 40 trials (10 opportunities for each response type).

Several error types were defined. First, premature errors were responses initiated during the delay interval, before the go-signal. Premature responses

immediately elicited an error tone of 3000 Hz for 500 ms, with the message “Too early! Wait for the go-signal.” presented in the center of the screen for 1000 ms, after which the model was presented and the next trial started. Second, slow errors were responses initiated more than 1100 ms after the go-signal, and immediately elicited the same 3000 Hz error tone, with the message “Too slow!” presented for 1000 ms, after which the model was presented and the next trial started. Third, perceptual errors were those judged on-line by the experimenter to differ from the target in terms of segmental content (distortions were accepted as correct), number of syllables, and temporal structure of sequences (i.e. a SLLS response for the 4L cue or vice versa). Note that no judgments were made of absolute duration of single syllables and sequences, nor of relative timing structure other than pattern reversals. Participants were not informed of the experimenter’s judgments of accuracy, to avoid drawing differential attention to sequence accuracy; no error messages were associated with these errors. It was expected that the auditory model would provide sufficient information for fine-tuning responses.

A rest interval was provided after each acquisition block. After the last trial in a block, a message appeared (white letters in bold Arial 24-pt font, on a teal background) that announced the end of the block and prompted the participant to press the space bar when ready to proceed with the next block. Participants were reminded in each rest interval to respond as accurately and as quickly as possible. The retention phase consisted of a single block of 20 trials (5 of each type), which was performed

without error messages and without the auditory models. Otherwise, trial events were identical to those during acquisition.

Each participant was tested individually; control participants were tested in a labroom, and patients were tested either in a room in the University's Communications Clinic or at their home. Before the experiment, the general nature and time commitment of the study was explained, participants read and signed the consent form, and filled out a questionnaire. Next, specific instructions were provided by the experimenter using a Powerpoint slideshow that presented the cues with verbal and orthographic descriptions of the associated responses. Similarly, the sequence of trial events was simulated using the slideshow, with written and verbal instructions about each event. When the participant indicated understanding of the task, the headset with the microphones was placed on the participant's head, and the experimenter assured that the participant was seated comfortably, with the index finger of the left hand resting on the space bar. A teal screen with white letters prompted the participant to press the space bar to hear the response models; after the models had been presented, a screen prompted another space bar press to begin the first acquisition block. Upon completion of the acquisition phase, participants were reminded about their retention session.

The acquisition session lasted approximately 60-75 minutes for control participants (age-matched and young controls), including paperwork, instructions, and the twelve acquisition blocks. Some patients took more time to complete acquisition

(range between 75 and 120 minutes for AOS patients, between 65 and 110 minutes for patients with aphasia). Retention testing lasted less than 5 minutes for all participants.

#### **5.1.4. Design and Analysis**

The two primary dependent variables of interest were Study Time (ST) as an index of the INT process, and Reaction Time (RT) as an index of the SEQ process. The independent variables were Duration (short versus long) of single presses, and Sequence Length (1 versus 4 syllables). First, data from younger controls were analyzed using 2 (Duration) x 2 (Sequence Length) x 12 (Block) repeated measures ANOVAs. The effect of Duration was also examined for the single syllables separately to assess the INT process without the confound of SEQ. Retention was analyzed in 2 (Duration) x 2 (Sequence Length) x 2 (Phase) ANOVAs with repeated measures on all factors, comparing the last acquisition block to the retention block. Follow-up testing was conducted using Tukey tests using an alpha level of .05.

Next, patient data were analyzed in two ways. First, separate group analyses were performed comparing each patient group against the age-matched control group, using 2 (Group) x 2 (Duration) x 2 (Sequence Length) x 12 (Block) ANOVAs with repeated measures on the last three factors for acquisition data, and 2 (Group) x 2 (Duration) x 2 (Sequence Length) x 2 (Phase) repeated measures ANOVAs for the retention data. It is recognized that with such a small sample size, these group analyses suffer from lack of power. However, in the apraxia literature parametric analyses are frequently performed on small groups and thus this analysis was conducted as it is the

standard in the apraxia literature. Because of the small sample size, a second analysis was conducted to determine whether each patient conformed to the group pattern using a relatively new individual subject analysis method developed by Crawford and Howell (1998). This analysis is used as the primary index of statistical differences between the patients and control participants. For this analysis, the ST or RT (collapsed across blocks) of each patient was compared against the ST and RT of the age-matched control group, using this modified t-test method. This method was developed specifically for comparing scores from a single individual against a the mean score from a small control group using a t-distribution rather than a normal distribution; the control mean and standard deviation are treated as sample statistics rather than population parameters (which would be inappropriate given the relatively small sample sizes). This test returns a t-value, and a one-tailed p-value that can be used as a point estimate of effect size (Crawford & Howell, 1998); the p-value essentially represents the percentage of controls who would obtain the individual's score.

In addition, to test for potential dissociations between different tasks (i.e. disproportionate complexity effects), the Revised Standardized Difference Test (RSDT; Crawford & Garthwaite, 2005) was applied. This test compares the standardized difference of an individual's performance on two tasks X and Y against the difference between tasks X and Y in the control sample. This test returns a t-value, a two-tailed p-value, and a one-tailed p-value that essentially represents the percentage of the controls who would obtain a difference in scores on X and Y as different as that

for the individual (Crawford & Garthwaite, 2005). This test was developed specifically to determine the presence of neuropsychological dissociations, especially for tasks whose scores and standard deviations are different. Crawford and his colleagues have demonstrated using Monte Carlo simulations that these tests maintain control of Type I error with small control sample sizes (Crawford & Garthwaite, 2005; Crawford & Howell, 1998), and are robust to departures from normality in most cases (Crawford, Garthwaite, Azzalini, Howell, & Laws, in press).<sup>2</sup> Although this method is increasingly being used in neuropsychology (e.g., Friedmann & Gvion, 2006), these analyses have not been used in the motor control/learning or motor speech disorders literature and provide unique ways to examine data of the sort reported in this experiment and that typical of the apraxia literature.

In addition to measures of the two motor programming stages preceding initiation (INT and SEQ), we also included measures of the execution stage, in particular with respect to the accuracy of timing. These measures were based on acoustic analyses, performed on a subset of the total data set. Specifically, all correct responses in even acquisition blocks plus the retention block were analyzed for eight randomly selected young controls and the four patients with AOS. Measures taken included the onset of each consonant, as defined by the release burst of the stop, and the onset and offset of each vowel, as defined by the first and last occurrence of vertical striations through the first and second formants (F1 and F2). Measures were based on wide-band (300 Hz) spectrograms with a visible frequency range of 0-10

---

<sup>2</sup> Free software programs that perform the necessary computations are made available on Dr. Crawford's website (<http://www.abdn.ac.uk/~psy086/dept/>).

kHz in a 1000 ms window, using the TF32 software (Milenkovic, 2000). Two independent raters performed the acoustic analyses, and analyzed the same data from four participants. Reliability assessment on the data for these four speakers (3 controls, 1 patient with AOS; 33% of the acoustic data) indicated excellent reliability for the consonant onset and vowel onset measures: the mean absolute difference for both these measures was less than 3.5 ms (for all positions in the sequence), with the percentages of measurement differences less than 10 ms between 95 and 100%. Reliability for the vowel offset measurement was lower, as expected: the mean absolute difference ranged from 3.5 to 22 ms, and the percentages of measurement differences less than 10 ms ranging from 35 to 85%. Reliability was similar for the control speakers and the patients with AOS.

Based on these acoustic measures, syllable durations and stop-gap durations were calculated for assessment of absolute and relative timing accuracy. Specifically, we used E (Total Error) as a measure of absolute timing error, and AE-prop (proportional absolute error) as a measure of relative timing error; both measures have been used in the literature on motor sequence learning (e.g., Shea, Wulf, Park, & Gaunt, 2001; Wright & Shea, 2001). E is an error measure that is sensitive to both response bias (overshoot and undershoot) and response variability, and captures control of absolute timing (controlled by a parameter in schema theory). E is computed as  $\sqrt{(CE^2 + VE^2)}$ , where CE is the average of the constant error (target total duration – actual total duration) and VE (variable error) is the standard deviation of the constant error (Shea et al., 2001; Wright & Shea, 2001).



To capture the ability to control the relative timing structure (generalized motor program) of the syllable sequences, we computed AE-prop (proportional absolute error), which represents as a measure of deviation from a target relative temporal structure, regardless of overall absolute duration (Shea et al., 2001; Shea & Park, 2003). AE-prop is computed as the sum of absolute differences in proportion of actual syllable movement time (MT) to the actual total sequence MT, and the proportion of target syllable MT to target total sequence MT. In formula form,

$$\text{AE-prop} = \Sigma ( |\text{Actual MT}_1 \text{ prop} - \text{Target MT}_1 \text{ prop}| + |\text{Actual Pause}_1 \text{ prop} - \text{Target Pause}_1 \text{ prop}| + |\text{Actual MT}_2 \text{ prop} - \text{Target MT}_2 \text{ prop}| + |\text{Actual Pause}_2 \text{ prop} - \text{Target Pause}_2 \text{ prop}| + |\text{Actual MT}_3 \text{ prop} - \text{Target MT}_3 \text{ prop}| + |\text{Actual Pause}_3 \text{ prop} - \text{Target Pause}_3 \text{ prop}| + |\text{Actual MT}_4 \text{ prop} - \text{Target MT}_4 \text{ prop}| ),$$

where  $\text{MT}_i \text{ prop} = \text{MT of syllable}_i \text{ divided by the total sequence MT}$ , and  $\text{Pause}_i \text{ prop} = \text{duration of stop-gap interval}_i \text{ divided by total sequence MT}$ . Thus, the 4S response had a target structure of [.10 - .07 - .30 - .07 - .30 - .07 - .10] and the 4L response had a target structure of [.30 - .07 - .10 - .07 - .10 - .07 - .30]. To illustrate, a 4S response produced as 200-75-550-125-400-100-150 would have actual proportions of .13 (200/1600), .05 (75/1600), .34 (550/1600), .08 (125/1600), .25 (400/1600), .06 (100/1600), and .09 (150/1600). Entering these numbers into the formula, this results in an AE-prop value of  $\Sigma (|.13 - .10| +|.05 - .07| +|.34 - .30| +|.08 - .07| +|.25 - .30| +|.06 - .07| +|.09 - .10|) = \Sigma (.03 + .02 + .04 + .01 + .05 + .01 + .01) = .17$ . Note that the last syllable duration in this example is the intended 150 ms, yet as a proportion of the

total sequence duration it is slightly too short (it would have to be 160 ms for 0 error, since  $160/1600 = .10$ ). The error measures were analyzed for each patient separately using the modified t-test (Crawford & Howell, 1998) and the RSDT (Crawford & Garthwaite, 2005). Thus, these measures were designed to further specify the nature of any deficits in control of timing for speech in AOS.

### **5.1.5. Predictions**

Based on previous findings of effects of duration and sequence length for motor programming of finger movements (e.g., Immink & Wright, 2001; Klapp, 1995), and the demonstrated applicability of the INT/SEQ model to speech production (e.g., Deger & Ziegler, 2002; Klapp, 2003), our predictions for the young control group were that if duration indeed captures motoric complexity in speech as it does for finger movements, duration of single syllables would affect ST (longer ST for longer syllables) indicating a greater INT load, but not RT since the internal features of a movement can be preprogrammed. Furthermore, considering that repeating syllable sequences were used, we expected to see an effect of sequence length on RT since such sequences have been shown to load on the SEQ process (e.g., Deger & Ziegler, 2002; Klapp, 2003) as they are presumably programmed as separate units.

The primary hypothesis of impaired INT but intact SEQ in AOS predicts that patients with AOS should demonstrate longer ST than controls, whereas RT should not be different. Furthermore, on the assumption that preparation of longer syllables is more complex than preparation of short syllables, it was expected that patients with

AOS would show disproportionate complexity effects on ST (i.e. larger short-long differences for patients with AOS than for controls). Patients with aphasia were not expected to show any differences for either ST or RT. Finally, for the timing error measures, it was expected that the patients with AOS would show greater error than the controls; no specific predictions were made with respect to whether absolute timing, relative timing, or both would be impaired.

## **5.2. Results**

For the ANOVA results, only significant effects ( $p < .05$ ) and trends ( $.05 < p < .10$ ) are reported, unless specific predictions were made for the effect.

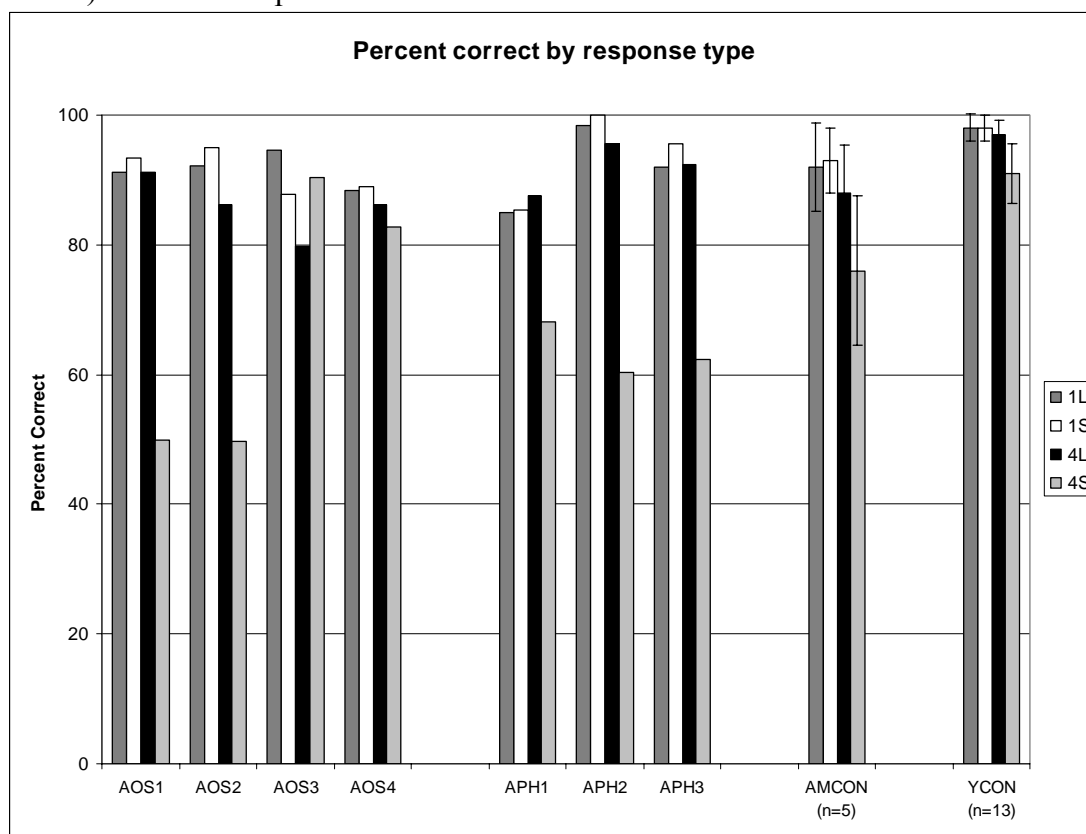
### **5.2.1. Accuracy: Number of Error Trials**

As a first step, the number of errors produced during acquisition (i.e. the number of rerun trials) was used to calculate a percent accuracy score for each participant for each response type separately (collapsed across blocks) (see Graph 5-1). The total accuracy percentage (collapsed across response type) for each group by block, and the distribution of error types as a percentage of the total number of errors are provided for each group separately in Appendix B.

For the younger control group, a 2 (Sequence Length) x 2 (Duration) x 12 (Block) ANOVA on accuracy revealed a main effect of Sequence Length ( $1 > 4$ ;  $F[1,11] = 5.50, p < .05$ ) and of Block ( $F[11,121] = 5.17, p < .0001$ ), indicating that accuracy in Block 1 was lower than in all subsequent blocks. There was also a

marginal Sequence Length by Duration interaction ( $F[1,11] = 4.83, p = .0502$ ), suggesting that accuracy for the 4S response was lower than accuracy for all other responses, with no differences between the other response types. Finally, the three-way interaction was also significant ( $F[11,121] = 2.43, p < .05$ ), which indicated that accuracy for the 4S sequence was lower than accuracy for 1S and 4L in block 1, whereas there was no difference with 1L accuracy.

**Graph 5-1.** Percent correct for each patient and both control groups (collapsed across blocks). Error bars represent standard error.



For the comparison between the AOS group and the age-matched controls, a 2 (Group) x 2 (Sequence Length) x 2 (Duration) x 12 (Block) repeated measures

ANOVA revealed main effects of Sequence Length ( $1 > 4$ ;  $F[1,7] = 16.83$ ,  $p < .005$ ), Duration ( $L > S$ ;  $F[1,7] = 6.92$ ,  $p < .05$ ), and Block ( $F[11,77] = 6.68$ ,  $p < .0001$ ), indicating that accuracy in Block 1 was lower than in all subsequent blocks except Block 3, which differed from Blocks 10 and 12. There was also a marginal Sequence Length by Duration interaction ( $F[1,7] = 3.71$ ,  $p = .095$ ) suggesting that only the 4S sequence differed from the single syllables, whereas the 4L sequence did not differ from any other response type in terms of accuracy. In addition, there was a significant Sequence Length by Block interaction ( $F[11,77] = 2.37$ ,  $p < .05$ ) which indicated that sequences were less accurate than single syllables during blocks 1-4 but not thereafter. While the Group effect was not significant, there was a marginal Group by Block interaction ( $F[11,77] = 1.89$ ,  $p = .0541$ ), which indicated that the age-matched controls had lower accuracy in Block 1 than in all other blocks, whereas the AOS group showed no differences in accuracy across blocks.

For the APH group, the ANOVA revealed significant main effects of Sequence Length ( $1 > 4$ ;  $F[1,6] = 23.00$ ,  $p < .005$ ), Duration ( $L > S$ ;  $F[1,6] = 26.37$ ,  $p < .005$ ), Block (Block 1 < subsequent blocks;  $F[11,66] = 8.86$ ,  $p < .0001$ ), and significant interactions between Sequence Length and Duration ( $F[1,6] = 15.01$ ,  $p < .01$ ), Duration and Block ( $F[11,66] = 2.03$ ,  $p < .05$ ), Sequence Length and Block ( $F[11,66] = 2.31$ ,  $p < .01$ ), and Sequence Length by Duration by Block ( $F[11,66] = 3.52$ ,  $p < .001$ ). Accuracy for the 4S sequence was lower than all other responses in blocks 1-4 but not in subsequent blocks; the other responses did not differ from each other in accuracy. There was also a marginal Group x Duration interaction ( $F[1,6] = 4.57$ ,  $p =$

.076), suggesting that accuracy was lower for S than for L responses in the APH group but not in the age-matched control group.

To examine accuracy for each patient separately, the accuracy scores of the age-matched control participants were averaged to provide a basis for comparison of each individual patient's score using the modified t-test (Crawford & Howell, 1998). A table with all comparisons is provided in Appendix B (Table B-1). None of the patients differed from the age-matched control group with respect to accuracy for any response type (all p-values > .20, one-tailed,  $df = 4$ ). However, APH4 produced a high number of incorrect responses for the sequences (reversing the intended temporal pattern, incorrect number of syllables). Due to time constraints, he only finished 6 of the 12 acquisition blocks with incomplete data (due to the upper limit on number of trials per block); as a result, his data are not included in the analyses reported below.

### **5.2.2. INT: Study Time (ST)**

For the ST analyses, STs less than 100 ms and greater than 10 seconds were removed as invalid data points. This resulted in a total data loss of 1.7% (0.4% for young controls, 1.1% for age-matched controls, 2% for patients with AOS, and 7.9% for patients with aphasia). Analyses (ANOVAs, modified t-tests, and RSdT) were performed on log-transformed ST means to meet the normality assumption (only untransformed means are reported to enhance interpretability).

Young Controls. The overall 2 (Duration) x 2 (Sequence Length) x 12 (Block) repeated measures ANOVA for the acquisition data revealed significant main effects

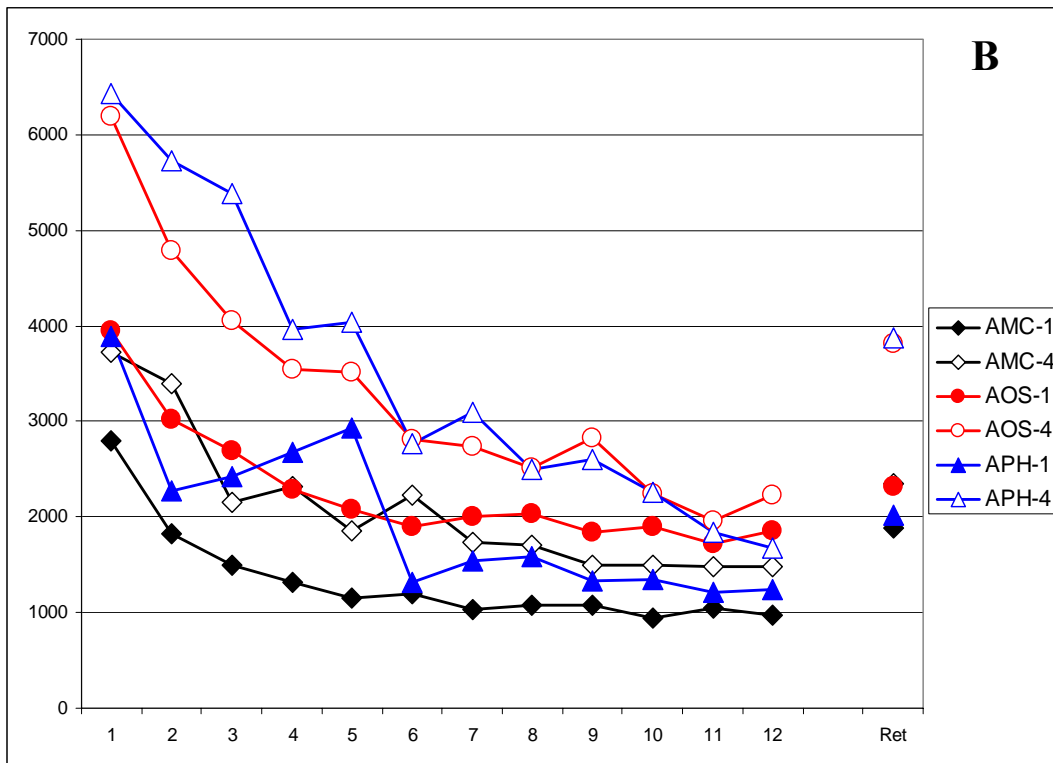
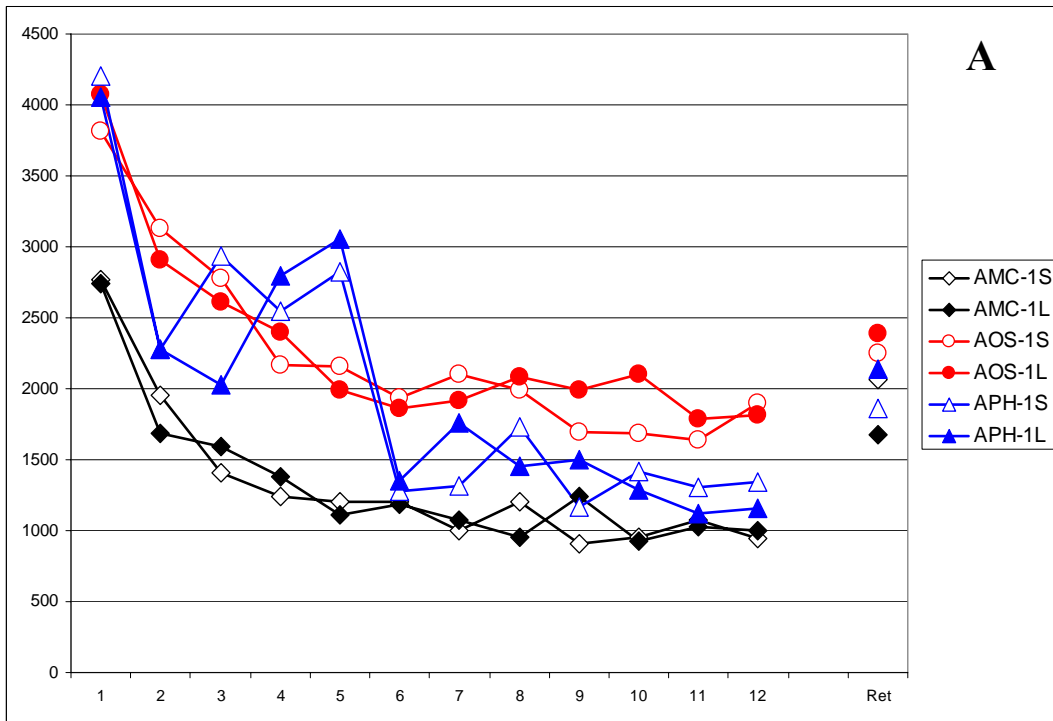
of Sequence Length ( $F[1,11] = 26.09, p < .001$ ), Duration ( $F[1,11] = 5.95, p < .05$ ), and Block ( $F[11,121] = 48.05, p < .0001$ ). In addition, there was a significant Sequence Length x Duration interaction ( $F[1,11] = 6.43, p < .05$ ), which indicated that the 4S response (SLLS) resulted in longer ST (mean = 1540, SD = 931) than the 4L (LSSL) response (mean = 1373, SD = 810), while there was no difference between single short and single long syllables (1S mean = 945, SD = 446; 1L mean = 947, SD = 488)<sup>3</sup>, both of which differed from the sequences. The Sequence Length by Block interaction was also significant ( $F[11,121] = 2.94, p < .01$ ), indicating that for single syllables, there was only a significant decrease in ST from Block 1 to Block 2, whereas for the sequences a further decrease was present from Block 2 to Block 3; in each acquisition block, single syllables and sequences differed significantly from each other. The retention analysis revealed only a main effect of Sequence Length ( $F[1,11] = 9.94, p < .01$ ) indicating longer ST for sequences (mean = 1293, SD = 681) than for single syllables (mean = 899, SD = 373), and a main effect of Phase ( $F[1,11] = 22.63, p < .0001$ ) indicating a significant increase in ST from the last acquisition block to the retention block.

Patients: AOS. ST results for the age-matched control group and the two patient groups are presented in Graph 5-2. Group comparisons with 2 (Group) x 2 (Sequence Length) x 2 (Duration) x 12 (Block) repeated measures ANOVAs for the

---

<sup>3</sup> Separate analyses performed on single syllable data only confirmed this pattern: There was no effect of duration in either acquisition or retention analyses ( $F_s < 1$ ).

**Graph 5-2.** Study Time (msec) by block, for single syllable duration, 1S vs. 1L (A) and sequence length, 1 vs. 4 syllables (B). AMC = age-matched control group (N=5), AOS = Apraxia of Speech group (N=4), APH = aphasia group (N=3).





AOS group against the age-matched controls revealed a significant main effect of Group ( $F[1,7] = 7.78, p < .05$ ), indicating that the AOS group showed longer ST (mean = 2739; SD = 1226) than the age-matched control group (mean = 1702; SD = 1136). In addition, there was a main effect of Sequence Length ( $F[1,7] = 14.42, p < .01$ ), indicating longer ST for sequences than for single syllables, and a main effect of Block ( $F[11,77] = 33.98, p < .0001$ ). There were no interactions with Group; the Sequence Length by Block interaction failed to reach significance ( $F[11,77] = 1.67, p = .097$ ). As with the younger controls, there was no effect of Duration ( $F < 1$ ).

Retention data from AOS2 were not analyzed because of a large number of premature responses and incomplete responses; thus, the retention analyses are based on the remaining three patients with AOS. Retention analysis revealed the same pattern of results as the acquisition analysis: There were main effects of Group ( $F[1,6] = 7.67, p < .05$ ) with longer ST for the AOS group (mean = 2481; SD = 971) than for the age-matched control group (mean = 1665; SD = 783), Sequence Length ( $F[1,6] = 10.29, p < .05$ ), and Phase ( $F[1,6] = 13.32, p < .05$ ).<sup>4</sup>

In order to determine whether the observed group effects on ST were present for all patients with AOS, each patient's mean ST (collapsed across blocks for each response type) was compared against the mean of the age-matched control group.

---

<sup>4</sup> A potential concern might be that the group and sequence length effects were driven largely by the 4S responses, which resulted in more error trials. Although no Sequence Length x Duration interaction was obtained, separate analyses (1S vs. 4S, and 1L vs. 4L) were performed for both acquisition and retention. The same pattern of results as in the overall analysis was obtained for all separate analyses, with the exception of the Group effect for the 1S-4S retention analysis where a nonsignificant trend was observed ( $F[1,6] = 3.91, p = .095$ ). In sum, the sequence length effect and the group effect appears to be robust and not an artefact of a single response. All analyses revealed the same pattern, i.e. absence of a group effect, significant effects of Sequence Length and Block, no interactions. Similarly, the individual analyses for each patient revealed the same pattern of results when analyzed separately.

Separate sets of analyses were performed to assess the Duration effect (including only single syllable data) and the Sequence Length effect (collapsed across S and L responses). The results for the Duration analyses are presented in Table 5-1, results for the Sequence Length analyses are presented in Table 5-2. As can be seen in Table 5-1, three of the four patients with AOS (AOS2, AOS3, and AOS4) demonstrated significantly longer ST than the age-matched control group for the single syllables during acquisition. AOS3 and AOS4 also had longer ST for sequences, while the difference for AOS2 failed to reach significance ( $t = 1.99$ ,  $p = .059$ , one-tailed,  $df = 4$ ). While STs for AOS1 during acquisition were also elevated, these differences were not significant for either 1S or 1L. For the retention comparisons, again all patients showed numerically longer STs than the age-matched controls, but only AOS4 showed a statistically significant difference (for 1L only). No evidence was found for disproportionate effects of Duration; however, since no Duration effect was found in any of the ANOVAs with the control groups, this may merely mean that single syllable duration is not a relevant complexity manipulation for speech.

The sequence length comparisons (Table 5-2) revealed that ST was longer than that of the age-matched control group for AOS2, AOS3, and AOS4, for the single syllables (confirming the findings from the Duration comparisons), but that ST for syllable sequences did not differ reliably from the age-matched control group for any patient, despite numerical differences in the predicted direction. There was no evidence for disproportionate sequence length effects.

**Table 5-2.** Study Time (ST) comparisons for each patient against age-matched controls for single syllables (ST means averaged across blocks). Untransformed means are presented for expository reasons; however, the statistical results reported in this table are based on log-transformed means. N/A = not analyzed (see text).

Subject	Duration	Acquisition	t-value	p-value	Retention	t-value	p-value
AOS1	1L	1925 (776)	1.441	0.111	2120	1.278	0.135
	1S	1784 (740)	1.410	0.116	2152	0.240	0.411
	Δ	141	0.052	0.961	-32	1.002	0.373
AOS2	1L	2248 (878)^	1.988	0.059			
	1S	2186 (842)*	2.347	0.039		N/A	
	Δ	62	0.585	0.590			
AOS3	1L	2365 (755)*	2.260	0.043	2213	1.499	0.104
	1S	2084 (768)*	2.188	0.047	2280	0.393	0.357
	Δ	281	0.118	0.912	-67	1.064	0.347
AOS4	1L	2643 (562)*	2.741	0.026	2838*	2.774	0.025
	1S	2943 (660)*	3.835	0.009	2319	0.438	0.342
	Δ	-300	1.709	0.163	519	2.151	0.098
APH1	1L	1702 (623)	1.041	0.178	1536	-0.375	0.363
	1S	1505 (435)	0.835	0.225	1397	-0.909	0.207
	Δ	197	0.336	0.754	139	0.520	0.630
APH2	1L	1509 (1225)	0.129	0.452	1815	0.483	0.327
	1S	1512 (1102)	0.228	0.415	1297	-1.107	0.165
	Δ	-3	0.162	0.879	518	1.507	0.206
APH3	1L	2747 (1582)*	2.486	0.034	3077*	3.188	0.017
	1S	3067 (1830)*	3.429	0.013	2886	1.109	0.183
	Δ	-320	1.489	0.211	189	2.011	0.115
AMC (N=5)	1L	1328 (337)			1673	(306)	
	1S	1323 (253)			2063	(730)	
	Δ	5			-390		
YCON (N=12)	1L	947 (173)			1141	(363)	
	1S	944 (186)			1062	(454)	
	Δ	3			79		

\*  $p < .05$ ,  $df=4$ , one-tailed (Crawford & Howell, 1998).

**Table 5-3.** Study Time (ST) comparisons for each patient against age-matched controls for sequence length (ST means averaged across blocks). Untransformed means are presented for expository reasons; however, the statistical results are based on log-transformed means.

Subject	Number	Acquisition	t-value	p-value	Retention	t-value	p-value
AOS1	4	2396 (1404)	0.356	0.370	4841 <sup>^</sup>	1.838	0.070
	1	1855 (703)	1.463	0.109	2138	0.561	0.302
	$\Delta$	541	1.361	0.245	2703	1.466	0.217
AOS2	4	3431 (1502)	1.082	0.170			
	1	2212 (814)*	2.159	0.048		N/A	
	$\Delta$	1221	1.327	0.255			
AOS3	4	3635 (1346)	1.265	0.137	3114	0.795	0.236
	1	2225 (738)*	2.247	0.044	2247	0.739	0.250
	$\Delta$	1410	1.215	0.291	867	0.066	0.951
AOS4	4	3682 (986)	1.335	0.126	3474	1.054	0.176
	1	2795 (578)*	3.262	0.016	2579	1.235	0.142
	$\Delta$	787	2.263	0.086	895	0.214	0.841
APH1	4	2577 (1552)	0.485	0.327	1922	-0.345	0.374
	1	1612 (495)	0.968	0.194	1489	-0.740	0.250
	$\Delta$	965	0.608	0.576	433	0.467	0.665
APH2	4	3154 (2310)	0.642	0.278	2579	0.349	0.372
	1	1510 (1133)	0.170	0.437	1556	-0.583	0.296
	$\Delta$	1644	0.0595	0.584	1023	1.086	0.339
APH3	4	4842 (1610) <sup>^</sup>	1.801	0.073		2.749	0.026
	1	2817 (1449)*	2.922	0.022	7119*	1.786	0.074
	$\Delta$	2025	1.379	0.240	3006 <sup>^</sup> 4113	1.120	0.325
AMC (N=5)	4	2091 (1015)			2349 (797)		
	1	1328 (280)			1880 (519)		
	$\Delta$	763			469		
YCON (N=12)	4	1457 (461)			1632 (729)		
	1	946 (171)			1106 (377)		
	$\Delta$	511			526		

\*  $p < .05$ ,  $df=4$ , one-tailed (Crawford & Howell, 1998).

<sup>^</sup>  $p < .10$ ,  $df=4$ , one-tailed (Crawford & Howell, 1998).

Patients: APH. Group analysis based on the 2 (Group) x 2 (Sequence Length) x 2 (Duration) x 12 (Block) repeated measures ANOVA for acquisition did not reveal a main effect of Group ( $F[1,6] = 2.80, p = .145$ ), despite a numerical difference in favor of the age-matched controls (controls mean: 1702, SD = 1136; APH mean = 2748, SD = 1946)<sup>5</sup>. The Sequence Length was significant however ( $F[1,6] = 18.48, p < .01$ ), as was the effect of Block ( $F[11,66] = 22.74, p < .0001$ ). No other effects approached significance. For retention, a similar pattern of results was obtained: There was no group effect ( $F < 1$ ; control mean = 1665, SD = 783; APH mean = 2168, SD = 1589), but there was a main effect of Sequence Length ( $F[1,6] = 8.19, p < .05$ ) and of Phase ( $F[1,6] = 21.50, p < .005$ ). No other effects approached significance, except for the Group x Duration x Phase interaction ( $F[1,6] = 4.24, p = .085$ ).<sup>6</sup>

Analysis of the acquisition data from individual patients using the modified t-test for Duration indicated that patient APH3 had longer ST than the age-matched controls (see Table 5-1), whereas the other two patients did not differ from the age-matched controls. At retention testing, APH3 also demonstrated longer ST than age-matched controls for the long syllable only. A similar pattern was found for the sequence length comparisons, in that APH3 showed longer ST than the age-matched controls during acquisition and retention, though the difference reached significance

---

<sup>5</sup> While the numerical difference appears to be of the same magnitude as that between patients with AOS and age-matched controls, recall that statistical analysis was performed on log-transformed means. However, analyses on untransformed means yielded the same pattern. The group mean in the APH group was elevated largely due to the means of a single patient (APH3, see individual analyses below).

<sup>6</sup> As was done in the AOS analysis, separate analyses were performed on S responses (1S vs. 4S) and L responses (1L vs. 4L), for acquisition and retention. All analyses revealed the same pattern, i.e. absence of a group effect, significant effects of Sequence Length and Block, with no interactions. Thus, the pattern from the overall analysis appeared to be stable.

only for the single syllables during acquisition (see Table 5-1) and only for sequences during retention (Table 5-2).

### 5.2.3. SEQ: Reaction Time (RT)

RTs under 100 ms were removed as invalid data points (anticipatory responses). This resulted in a total data loss of 4 data points (1 for young controls, 1 for age-matched controls, and 2 for patients with AOS). All analyses (ANOVAs, modified t-tests, and RSDT) were performed on untransformed means.

Young Controls. The 2 (Sequence Length) x 2 (Duration) x 12 (Block) repeated measures ANOVA for acquisition revealed only a main effect of Block ( $F[11,121] = 11.52, p < .0001$ ). Notably, the Sequence Length effect was not significant ( $F[1,11] = 2.15, p = .171$ ), despite a numerical difference in the predicted direction (single syllable mean = 534, SD = 112; sequence mean = 563, SD = 126). No other effects approached significance. Separate analysis for single syllables only revealed only a significant effect of Block ( $F[11,121] = 9.76, p < .0001$ ). As predicted, there was no effect of Duration ( $F[1,11] = 1.29, p = .280$ ). For retention, the Sequence Length effect was also not significant ( $F[1,11] = 3.75, p = .079$ ; single syllable mean = 522, SD = 92; sequence mean = 555, SD = 124), but the effect of Phase was significant ( $F[1,11] = 6.49, p < .05$ ).<sup>7</sup> Separate analysis on the single syllables only revealed only a main effect of Phase ( $F[1,11] = 6.97, p < .05$ ); there was no effect of syllable duration ( $F[1,11] = 1.73, p > .20$ ).

---

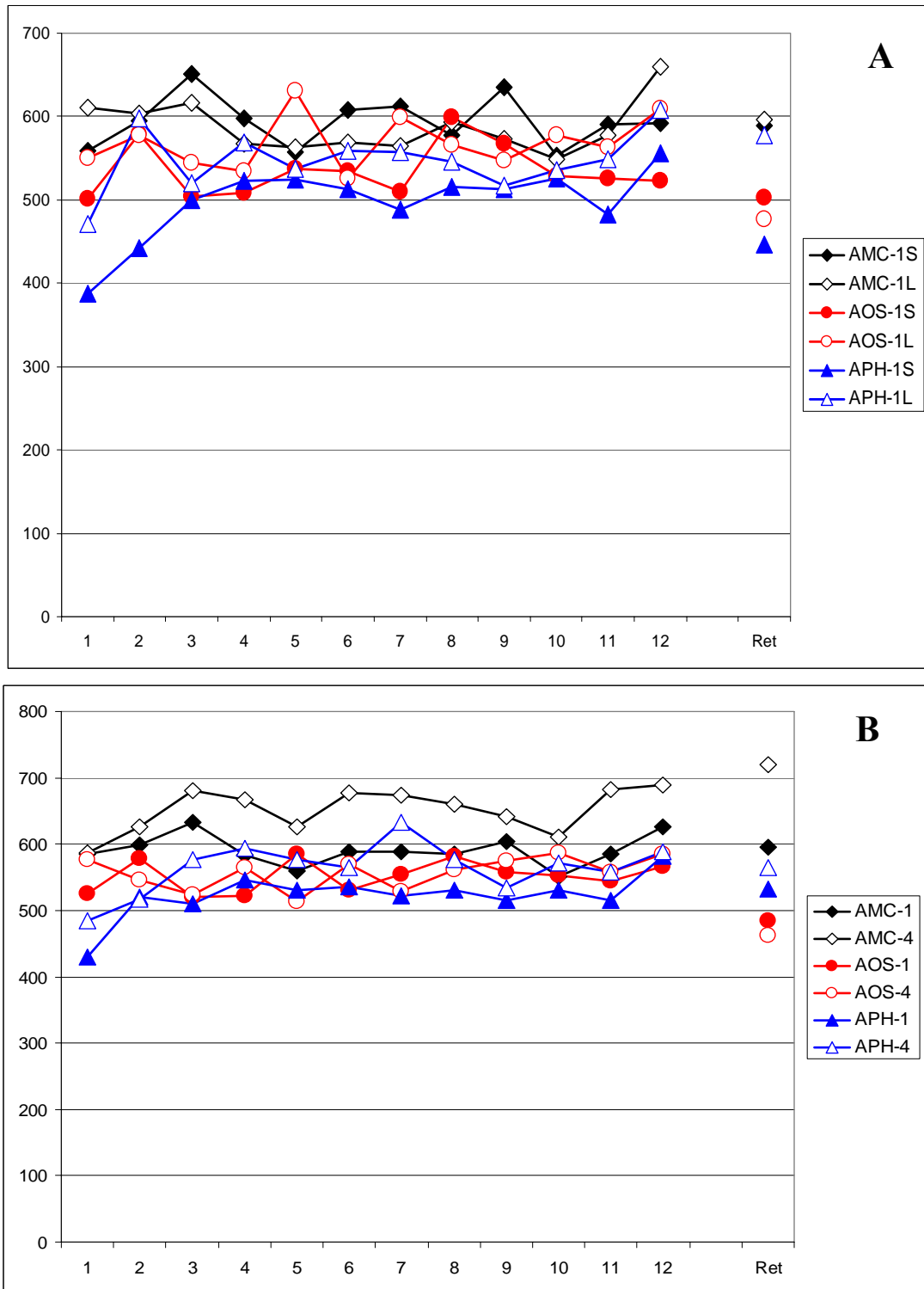
<sup>7</sup> As with the ST analyses, separate analyses were performed for S and L responses, for acquisition and retention. In all cases, the pattern was the same: significant effects of block/phase but no significant sequence length effects.

Patients: AOS. RT for the age-matched control group and the two patient groups are presented in Figure 3. The 2 (Group) x 2 (Sequence Length) x 2 (Duration) x 12 (Block) ANOVAs comparing age-matched controls and patients with AOS did not detect any significant effects. Most relevant to the hypothesis under investigation, there was no difference in RT between the two groups ( $F[1,7] = 1.62, p > .20$ ), with the patients in fact showing numerically faster RTs (mean = 555, SD = 101) than the age-matched controls (mean = 620, SD = 126). As with the younger controls, the Sequence Length effect was not significant ( $F[1,7] = 1.87, p > .20$ ).<sup>8</sup> Retention ANOVAs were not performed due to the fact that data for two patients (AOS1 and AOS2) included many premature responses.

---

<sup>8</sup> Again, the same pattern of results was obtained when analyzing S and L responses separately; i.e. no sequence length effects, no group effects.

**Graph 5-3.** Reaction Time across blocks for single syllable duration (A) and sequence length (B). AMC = age-matched control group (N=5), AOS = apraxia of speech group (N=4), APH = aphasia group (N=3).





Results from the modified t-test and RSDT comparisons are presented in Table 5-3 (single syllable duration) and Table 5-4 (sequence length). As can be seen in Table 5-3, none of the patients with AOS demonstrated longer RTs than the age-matched controls for the single syllables, with most of the comparisons revealing numerically faster RTs for the AOS patients. There was no evidence for disproportionate duration effects for any of the patients. During retention, AOS3 and AOS4 showed significantly faster RTs for long syllables than the age-matched controls. The same pattern held for the sequence length comparisons (Table 5-4): None of the patients with AOS differed from the age-matched control group with respect to RT, and numerical differences were in favor of the patients. At retention testing, AOS3 and AOS4 evidenced significantly faster RTs than the age-matched control group.

Patients: APH. Group analysis using 2 (Group) x 2 (Sequence Length) x 2 (Duration) x 12 (Block) ANOVAs failed to detect significant effects. Most importantly, there was no difference between groups ( $F < 1$ ), and the numerical difference was in favor of the APH group (APH mean = 542, SD = 158; control mean = 620, SD = 126). Again, the sequence length effect did not reach significance ( $F[1,6] = 3.07, p = .130$ ). Retention analyses were not performed on group data because data for APH1 included many premature responses.

**Table 5-4.** Reaction Time (RT) comparisons for each patient against age-matched controls for single syllables (RT means averaged across blocks). N/A = not analyzed (see text).

Subject	Duration	Acquisition	t-value	p-value	Retention	t-value	p-value
AOS1	1L	564 (65)	- 0.300	0.390	N/A		
	1S	542 (80)	- 0.688	0.265			
	Δ	22	0.726	0.508			
AOS2	1L	622 (113)	0.456	0.336	N/A		
	1S	602 (53)	0.106	0.460			
	Δ	20	0.657	0.547			
AOS3	1L	603 (84)	0.209	0.422	307*	- 4.644	0.005
	1S	508 (67)	- 1.138	0.159	583	- 0.048	0.482
	Δ	95^^	2.332	0.080	-276¶	3.034	0.039
AOS4	1L	487 (68)	- 1.304	0.131	436*	- 2.578	0.031
	1S	487 (61)	- 1.416	0.115	405	- 1.486	0.106
	Δ	0	0.210	0.844	31	0.777	0.481
APH1	1L	514 (69)	- 0.952	0.198	N/A		
	1S	383 (48)*	- 2.792	0.025			
	Δ	131¶	3.027	0.039			
APH2	1L	430 (67)^	- 2.047	0.055	537	- 0.961	0.196
	1S	421 (58)*	- 2.289	0.042	460	- 1.042	0.178
	Δ	9	0.454	0.673	77	0.058	0.956
APH3	1L	697 (45)	1.435	0.112	670	1.169	0.154
	1S	688 (91)	1.244	0.141	716	1.026	0.182
	Δ	9	0.360	0.737	-46	0.102	0.923
AMC (N=5)	1L	587 (70)			597 (57)		
	1S	594 (69)			589 (113)		
	Δ	-7			12		
YCON (N=12)	1L	541 (75)			561 (96)		
	1S	527 (81)			555 (94)		
	Δ	14			6		

\*  $p < .05$ ,  $df=4$ , one-tailed (Crawford & Howell, 1998).

^  $p < .10$ ,  $df=4$ , one-tailed (Crawford & Howell, 1998).

¶  $p < .05$ ,  $df=4$ , two-tailed (Crawford & Garthwaite, 2005).

^^  $p < .10$ ,  $df=4$ , two-tailed (Crawford & Garthwaite, 2005).

**Table 5-5.** Reaction Time (RT) comparisons for each patient against age-matched controls for sequence length (RT means averaged across blocks).

Subject	Number	Acquisition	t-value	p-value	Retention	t-value	p-value
AOS1	4	565 (48)	- 0.729	0.253	N/A		
	1	553 (66)	- 0.510	0.318			
	$\Delta$	13	0.207	0.846			
AOS2	4	680 (71)	0.234	0.413	N/A		
	1	612 (59)	0.282	0.396			
	$\Delta$	68	0.045	0.966			
AOS3	4	530 (62)	- 1.022	0.182	405*	- 5.045	0.004
	1	554 (40)	- 0.497	0.323	445^	- 1.863	0.068
	$\Delta$	-24	0.496	0.646	-40	2.042	0.111
AOS4	4	454 (54)^	- 1.658	0.086	372*	- 5.573	0.003
	1	487 (59)	- 1.396	0.118	421*	- 2.159	0.049
	$\Delta$	-33	0.248	0.816	-49	2.182	0.095
APH1	4	538 (86)	- 0.955	0.197	N/A		
	1	449 (49)^	- 1.906	0.065			
	$\Delta$	89	0.892	0.423			
APH2	4	394 (49)*	- 2.161	0.048	398*	- 5.157	0.003
	1	425 (46)*	- 2.228	0.045	498	- 1.209	0.147
	$\Delta$	-31	0.064	0.952	-100	2.496	0.067
APH3	4	763 (47)	0.930	0.203	782	0.993	0.188
	1	693 (64)	1.369	0.121	693	1.197	0.149
	$\Delta$	70	0.416	0.699	89	0.135	0.899
AMC (N=5)	4	652 (109)			720 (57)		
	1	591 (68)			596 (74)		
	$\Delta$	61			124		
YCON (N=12)	4	564 (95)			600 (116)		
	1	534 (75)			559 (89)		
	$\Delta$	30			41		

\*  $p < .05$ ,  $df=4$ , one-tailed (Crawford & Howell, 1998).

^  $p < .10$ ,  $df=4$ , one-tailed (Crawford & Howell, 1998).

Modified t-test and RSDT results are presented in Tables 5-3 (single syllables) and 5-4 (sequence length). It can be seen in Table 5-3 that APH1 and APH2 had significantly faster RTs than age-matched control for short syllables during

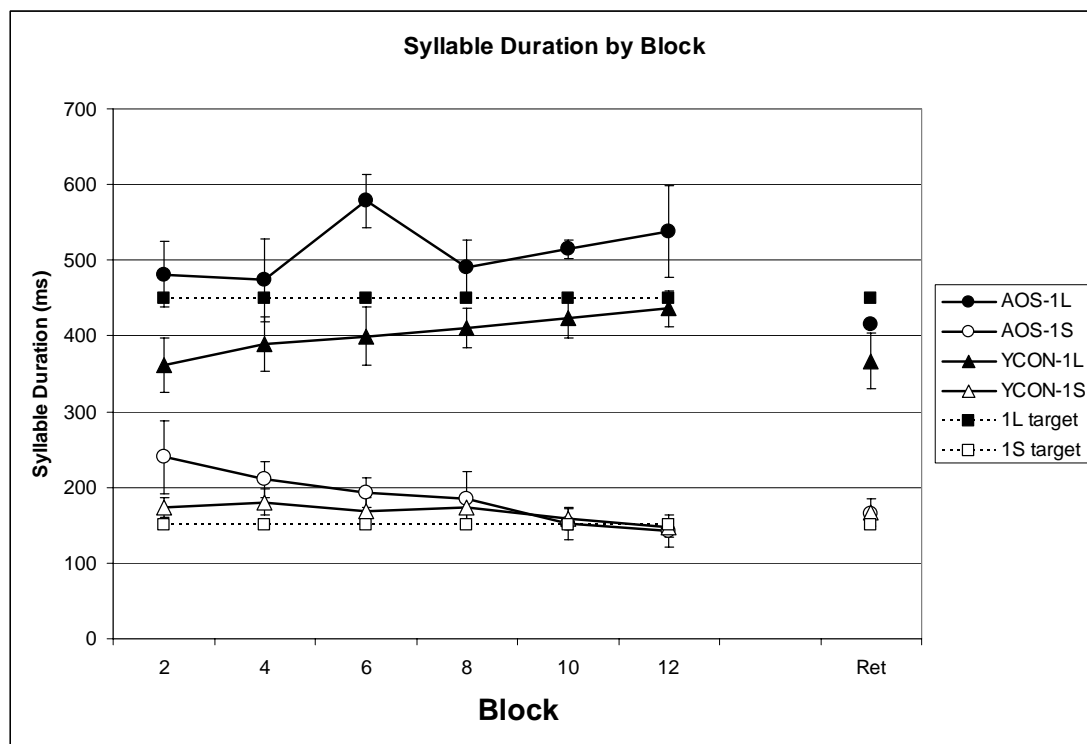
acquisition; this resulted in a disproportionate effect of syllable duration in the case of APH1. No other differences were significant in either acquisition or retention. APH3 was the only patient who consistently showed longer RTs than age-matched controls, but these differences did not reach significance. For the sequence length comparisons (Table 5-4), only APH2 demonstrated significantly faster RTs than age-matched controls, for both single syllables (acquisition) and sequences (acquisition and retention). APH3 was the only patient to consistently show numerically longer RTs than controls (though these differences did not reach significance).

#### **5.2.4. Execution: Duration, and Absolute and Relative Timing Error**

Single Syllable Duration. As a first step, syllable duration of single syllables was examined to determine whether participants differentiated these syllables despite the absence of augmented feedback on duration. Data are presented in Graph 5-4. For the younger control speakers, a 2 (Duration) x 6 (Block) repeated measures ANOVA revealed significant effects of Duration ( $F[1,7] = 176.97, p < .0001$ ) and Duration x Block ( $F[5,35] = 7.55, p < .0001$ ), which indicated that speakers reliably differentiated short and long syllables and that syllable durations diverged across acquisition blocks. The retention analysis revealed a main effect of Duration ( $F[1,7] = 156.29, p < .0001$ ) and a significant Duration x Phase interaction ( $F[1,7] = 20.17, p < .005$ ), which indicated that the duration of the long syllable was shorter in retention than in acquisition, whereas there was no difference for the short syllable.

Individual analyses comparing each patient against the control mean indicated that none of the patients differed significantly from the controls, with the exception of AOS2 who showed longer syllable duration for the long syllables during acquisition ( $t=1.921$ ,  $p = .048$ , one-tailed,  $df=7$ ; Crawford & Howell, 1998). See Appendix B (Table B-2) for individual comparisons.

**Graph 5-4.** Single syllable durations by block for speakers with AOS (N=4) and young control speakers (N=8), based on acoustic analysis of correct responses in even blocks.



Absolute Timing Accuracy. For absolute timing error (E) of single syllables, analysis of the acquisition data from young controls revealed only a marginal effect of Duration ( $F[1,7] = 5.04$ ,  $p = .0597$ ); the effect of Block was also not significant

( $F[5,35] = 1.89, p = .1213$ ). Thus, although these speakers significantly increased their differentiation between short and long syllables, this did not result in a significant reduction in absolute timing error relative to the target durations. However, the retention analysis revealed a significant Duration effect ( $F[1,7] = 8.16, p < .05$ ), with larger error for long syllables than for short syllables, and a significant effect of Phase ( $F[1,7] = 12.48, p < .01$ ) indicating greater error during retention than during the last acquisition block. Analysis of absolute timing error for sequences for control speakers revealed only a main effect of Block ( $F[5,35] = 2.68, p < .05$ ), indicating greater error in block 4 than in block 12. There were no differences between 4S and 4L sequences. There were no significant differences in the retention analysis.

Individual analyses of single syllable absolute timing error are presented in Table 5-5, and results for absolute and relative timing error of syllable sequences are presented in Table 5-6. For single syllables, only patient AOS2 showed greater absolute timing error than controls, for short syllables ( $t=5.161, p = .001, df=7$  one-tailed); the difference for long syllables failed to reach significance ( $t=1.423, p = .099, df=7$  one-tailed). For absolute timing of sequences, AOS2 and AOS3 had significantly greater absolute timing error than controls (both  $t_s > 4.50, p_s < .001$ ).

Relative Timing Accuracy. For relative timing error (AE-prop), the ANOVA on the acquisition data of the control speakers revealed a significant main effect of Block ( $F[5,35] = 2.91, p < .05$ ) and a significant Sequence Type by Block interaction ( $F[5,35] = 2.76, p < .05$ ). This pattern of results indicated that relative timing error of the 4S sequence was greater in Block 2 than in all subsequent blocks, whereas there

**Table 5-6.** Absolute timing error (E) for single syllables for each patient (collapsed across blocks), compared to young controls (N=8).

Subject	Duration	Acquisition	t-value	p-value	Retention	t-value	p-value
AOS1	1L	80 (29)	- 0.306	0.384		N/A	
	1S	66 (45)	1.056 <i>0.745</i>	0.163 <i>0.480</i>			
AOS2	1L	170 (71) <sup>^</sup>	1.423	0.099		N/A	
	1S	129 (68)*	5.161 <i>2.043<sup>^</sup></i>	0.001 <i>0.080</i>			
AOS3	1L	150 (73)	1.035	0.168	118	- 0.145	0.444
	1S	59 (23)	0.612 <i>0.231</i>	0.280 <i>0.824</i>	100	1.011 <i>0.626</i>	0.173 <i>0.551</i>
AOS4	1L	99 (29)	0.056	0.479	57	- 1.025	0.170
	1S	68 (44)	1.208 <i>0.631</i>	0.133 <i>0.548</i>	61	0.088 <i>0.603</i>	0.466 <i>0.566</i>
CON (N=8)	1L	96 (49)			128 (65)		
	1S	50 (15)			57 (40)		

\* p < .05, df=7, one-tailed (Crawford & Howell, 1998).

<sup>^</sup> p < .10, df=7, one-tailed (Crawford & Howell, 1998).

**Table 5-7.** Relative (AE-prop) and absolute timing error (E) for sequences (collapsed across 4S and 4L, and across blocks), compared to young controls (N=8).

Subject	Measure	Acquisition	t-value	p-value	Retention	t-value	p-value
AOS1	AE-prop	.367 (.053)	0.698	0.254		N/A	
	E	163 (51)	0.247 <i>0.281</i>	0.406 <i>0.787</i>			
AOS2	AE-prop	.358 (.082)	0.628	0.275		N/A	
	E	676 (223)*	7.547 <i>4.185<sup>¶</sup></i>	0.000 <i>0.004</i>			
AOS3	AE-prop	.369 (.041)	0.712	0.250	.528	0.915	0.195
	E	469 (72)*	4.600 <i>2.397<sup>¶</sup></i>	0.001 <i>0.048</i>	263	1.012 <i>0.070</i>	0.173 <i>0.946</i>
AOS4	AE-prop	.284 (.025)	0.070	0.473	.314	- 0.132	0.449
	E	115 (17)	- 0.440 <i>0.317</i>	0.337 <i>0.760</i>	82	- 1.031 <i>0.643</i>	0.168 <i>0.541</i>
CON (N=8)	AE-prop	.275 (.125)			.341 (.193)		
	E	146 (66)			173 (84)		

\* p < .05, df=7, one-tailed (Crawford & Howell, 1998).

<sup>¶</sup> p < .05, df=7, two-tailed (Crawford & Garthwaite, 2005).

were no differences between blocks for the 4L sequence. The effect of Sequence Type (4S vs 4L) was not significant ( $F < 1$ ), including for Block 2. Retention analysis revealed marginal effects of Block ( $F[1,7] = 3.62, p = .0987$ ) and Sequence Type x Block ( $F[1,7] = 5.57, p = .0503$ ), suggesting that only the 4S sequence showed a significant increase in error from the last acquisition block to retention.

For the AOS group, individual comparisons (see Table 5-6) showed that none of the patients demonstrated significantly greater relative timing error than the controls, despite numerically elevated error values for all patients. However, AOS2 and AOS3 (who both had greater absolute timing error than controls) demonstrated a significant dissociation between absolute and relative timing error ( $t_s > 2.00, p_s < .05$ ).

### **5.3. Discussion**

Experiment 1 was designed to test the hypothesis of a deficit in INT and not SEQ in AOS, focusing specifically on the control of timing of syllables and syllable sequences. Since this was the first experiment to apply the self-select paradigm to the study of speech production, discussion of the findings from the younger control speakers in relation to the INT/SEQ model is warranted.

#### **5.3.1. Speech Motor Programming in Younger Control Speakers**

Starting with the sequence length effect, the younger control speakers revealed longer programming time for syllable sequences than for single syllables, replicating previous findings using the self-selection paradigm (e.g., Immink & Wright, 2001;



Magnuson et al., in press; Wright et al., 2004). However, this effect was confined to ST and was not evident for RT, suggesting that programming of the sequences was completed during the INT process and resulted in a single integrated unit, rather than merely loading multiple syllables as separate units as was the case for button press sequences (Magnuson et al., in press). In other words, this pattern of ST and RT results suggests that the sequence, not the syllable, functioned as the speech motor program. At first glance, these findings are in contrast with the results from Klapp (2003) and Deger and Ziegler (2002), who found that repeated syllables are not integrated during the INT process but instead are processed as a series of separate units, thus increasing demands on the SEQ process.

This discrepancy may relate to a number of methodological differences with previous studies. For example, Klapp (2003) applied a screening procedure for his participants, in which they were included in the experiment only if their mean choice RT was less than 550 ms and their mean simple RT was less than 350 ms. No such screening was performed in the present experiment. And although Deger and Ziegler (2002) did not apply such screening procedures (and yet observed a sequence length effect), it should be noted that the average RTs in both Klapp (2003) and Deger and Ziegler (2002) did not exceed 350 ms, whereas the average RTs in the present experiment were in the 500 to 600 ms range. While these differences may merely reflect differences in sensitivity of measurement equipment, it is also possible that differences in participant characteristics and/or strategies are responsible for the discrepancy with these studies. Consistent with this view is the fact that the present

study was presented as a learning experiment, with a focus on both accuracy and speed, whereas the primary focus in Klapp (2003) and Deger and Ziegler (2002) was speed.

Another, perhaps more critical difference with previous studies is that in the present experiment a particular target prosodic (temporal) pattern was imposed on the sequences, whereas neither Klapp (2003) nor Deger and Ziegler (2002) specified the prosodic characteristics of their target responses, and it is possible that their speakers imposed primary stress on each syllable. The presence of a rhythmic pattern on the sequences in the present experiment may have encouraged speakers to integrate the sequence into a single unit (a stress group, or phonological word; Levelt et al., 1999; Sternberg et al., 1978). Thus, it may not be the alternation between different syllables per se that facilitates unit integration, but the presence of a target temporal pattern that binds the syllables together. This hypothesis is examined directly in Experiments 3 and 4.

Turning now to single syllable complexity (duration), we did not replicate the finger tap studies (e.g., Klapp, 1995; Immink & Wright, 2001), in that the younger control speakers did not show any evidence for ST differences related to syllable duration: Short and long single syllables took equally long to program. Although in a previous study using finger movements, Wright et al. (2004) also failed to find a main effect of duration on ST, these authors did observe the expected effect during the first 30 trials of the acquisition phase. Reanalysis of the duration effect in the early practice blocks of our Experiment 1 did not reveal any effects of duration, however. The

INT/SEQ model predicted such a difference on the assumption that syllable duration is an appropriate metric of speech motor program complexity. One interpretation of this null effect is that speech is such a highly practiced motor skill that relatively small variations in syllable duration simply do not place substantial (measurable) demands on speech motor programming.

However, there are other possible explanations relating to methodological differences with previous studies. For example, in contrast to previous studies which used a reaction time window<sup>9</sup> of 400 ms (e.g., Immink & Wright, 2001), the RT window in the present study was increased to 1000 ms to minimize frustration due to potentially longer RTs in the experimental subjects (i.e. those who had suffered strokes). It is possible that this increase in RT window might have encouraged speakers to postpone some aspects of motor programming until later, perhaps due to the cost associated with maintaining a programmed response in the buffer (e.g., Canic & Franks, 1989; Sternberg et al., 1978). Programming did not appear to be postponed until after the go-signal, else we should have seen an effect of syllable duration on RT. It remains possible however that programming occurred during the delay interval. Another possible methodological explanation is that detailed augmented feedback is necessary for the duration effect to emerge, perhaps due to greater effort on the participants' part to maintain a high degree of accuracy. Recall that in this study, no feedback was provided on the duration of single syllables, as the perceptual judgment of duration was not deemed reliable. However, analysis of the single syllable duration

---

<sup>9</sup> Reaction time window refers to the period after the go-signal during which a participant must initiate the response in order to avoid a "Too slow" error.

data for eight younger control speakers revealed that these speakers did successfully differentiate between short and long syllables, and furthermore that they increased the duration difference across practice blocks, suggesting that at least these eight speakers increased their accuracy even in the absence of feedback about absolute timing. Data from Experiment 1 do not allow us to distinguish between these explanations.

However, Experiment 2 (on finger movements) will allow us to differentiate between the two methodological explanations, in that detailed feedback is provided but the RT window is identical to the RT window in Experiment 1. Thus, if an effect of duration is found in Experiment 2, this would argue against the RT window hypothesis.

### **5.3.2. Speech Motor Programming in Apraxia of Speech and Aphasia.**

Turning now to the main purpose of Experiment 1, the primary predictions arising from the INT hypothesis of AOS were that ST would be longer for speakers with AOS than for age-matched control speakers, but that RT would not differ. Critically, the results of Experiment 1 confirmed both these predictions: STs were found to be longer than in the age-matched control group in the group analyses as well as in the individual analyses for three of the four patients with AOS. The one patient with AOS who did not show the predicted effect differed from the other patients in several respects, including handedness (he was left-handed) and language background (he was a simultaneous bilingual), which may have facilitated the task for him. Importantly, none of the patients differed from the controls with respect to RT, suggesting a localized impairment. Thus, this pattern of results is consistent with the

hypothesis that AOS involves a deficit in preprogramming (INT) of speech responses, whereas SEQ and response initiation are unimpaired. These findings are in agreement with Deger and Ziegler's (2002) interpretation of impaired INT in AOS. However, the present study is the first to assess the preprogramming stage directly in AOS, and allowed us to distinguish between INT and SEQ stages of processing, unlike previous work.

A third prediction from the INT-deficit hypothesis was that speakers with AOS should show disproportionate complexity effects on INT. However, since no effects of unit complexity (duration) were observed for either patients or controls, this prediction could not be assessed in this experiment. It remains to be seen in future work whether disproportionate complexity effects will be obtained when using different metrics of motor programming complexity (e.g., pitch or loudness manipulations, speaking with a biteblock).

Given the interpretation for control speakers that the sequences in this experiment were programmed as single units, a fourth prediction was that speakers with AOS might fail to integrate the successive syllables into a single unit, thus producing a sequence length effect on RT. Recall that Deger and Ziegler (2002) observed longer simple RT for /daba/ than for /dada/ for their speakers with AOS (but not for their control speakers), which they suggested resulted from an inability to fully preprogram two syllables, with INT occurring during the simple RT interval. If INT occurs during the simple RT interval in AOS, then effects of program complexity should emerge on simple RT, and programming different syllables has been shown to

increase programming time relative to programming identical syllables (Schönle et al., 1986). Deger and Ziegler phrased this effect in terms of programming syllable transitions, though it is also possible that their effect arose due to loading multiple different elements, even without integrating them.

Contrary to this fourth prediction, there was no effect of sequence length on RT for the speakers with AOS, essentially resulting in the same pattern of results as the controls, except with elevated ST. Thus, there is no reason to posit a different mechanism or programming strategy for the absence of a sequence length effect on RT; instead, parsimony suggests that the individuals with AOS in this experiment were able to integrate the syllables of a sequence into a single unit, but that it took them longer than controls to complete this process. The difference with Deger and Ziegler's (2002) findings may relate to differences in severity. The patients in our experiment were classified as having relatively mild AOS, whereas those in the Deger and Ziegler study were more severely impaired. Thus, it is possible that patients with more severe AOS would have shown a sequence length effect in this experiment. However, the interpretation that these individuals with AOS were able to integrate the syllables into a single unit would predict a disproportionate sequence length effect on ST, since in this case sequence length defines unit complexity. In fact, the ST group effect was stronger for the single syllables than for the sequences. One possible explanation is that the speakers with AOS did not program the sequence units as completely as the controls did. In other words, perhaps the control speakers spent more time programming the sequence relative to single syllables, whereas the speakers

with AOS may have programmed the single syllables fully yet may have left the sequences underspecified to some extent. This view is compatible with the finding of reduced absolute timing accuracy for sequences in two of the three patients with AOS who showed longer ST than controls. Perhaps the control speakers loaded the GMP and took time to adequately parameterize the program, whereas the speakers with AOS may have loaded the GMP but did not spend much time parameterizing the program.

With respect to timing accuracy, only two patients with AOS demonstrated greater error, and only for absolute timing, resulting in a significant dissociation between relative and absolute timing. This provides neuropsychological support for the distinction between these aspects of movement sequences, which is captured in Schema theory (Schmidt, 1975, 2003; Schmidt & Lee, 2005) by the constructs of GMPs and parameters, each of which can be affected independently. These findings are in agreement with those obtained by Clark and Robin (1998) using a nonspeech oral motor control task. Clark and Robin demonstrated that either GMP accuracy (as indexed by RMS error) or parameterization accuracy (as indexed by time and amplitude scaling factors) were compromised in AOS, but not both. Furthermore, these findings suggest that these constructs, and their measures as they have been used in the motor learning literature (e.g., Shea et al., 2001; Wright & Shea, 2001), hold relevance to understanding speech production (disorders).

Given that the interpretation of an INT deficit here is based primarily on the longer ST and equal RT, relative to controls, and given that ST in itself does not

specify the exact nature of the processes occurring in this interval, it is important to consider other possible explanations for the pattern of results. Several possibilities will be considered here, including the reduced buffer capacity hypothesis (Rogers & Storkel, 1999), a verbal working memory deficit hypothesis, the Dual Route hypothesis (Varley & Whiteside, 2001a,b), a phonological encoding hypothesis, a stimulus-response mapping hypothesis, and a self-judgment or monitoring hypothesis.

Reduced Buffer Capacity Hypothesis. Recall that according to the reduced buffer capacity hypothesis (Rogers & Storkel, 1999), the problem in AOS is that the articulatory buffer can only contain a single syllable. This hypothesis predicts that there should be no sequence length effect on RT, just as was observed. However, the hypothesis would have to be augmented in some way to account for the longer ST for single syllables, as well as for the observed sequence length effect on ST. Longer ST for single syllables could arise if it is assumed, for example, that the buffer capacity limitation is a problem of rapid decay (as soon as a syllable is entered into the buffer, its activation level decays), which would lead the speaker with AOS to reprogram the buffer again and again, either until a sufficient level of activation is reached due to summation of residual activation from multiple attempts or until the speaker decides to continue the trial without a full and complete syllable program in the buffer.

With respect to the sequence length effect on ST, it is unclear how the reduced buffer capacity hypothesis accounts for this, since this effect suggests that these patients did at least attempt to program more than one syllable. Perhaps the short and the long syllable (that make up the sequence) compete with one another for buffer



space, thereby increasing the time to load either one of them. To test this hypothesis, future studies could systematically manipulate similarity of the syllables in a sequence, and determine whether greater similarity leads to longer ST, as would be expected if increasing similarity increases competition or interference. Alternatively, the sequence length effect could arise at a level other than the articulatory buffer, such as phonological encoding or motor programming. Perhaps multiple syllables are planned at these preceding levels, even though ultimately only one of these syllables can enter the buffer.

Finally, a failure to program all syllables of a sequence and load them into the buffer as a single multisyllabic motor program should result in syllable-by-syllable programming, with programming of the second syllable occurring during or after execution of the first syllable (i.e. once the buffer is cleared). Such a programming strategy is likely to impede the integrity of a relative timing pattern defined over the entire sequence, in which syllable durations are defined in relation to each other. However, there were no differences in relative timing accuracy between groups. Thus, there was no evidence to support the idea of a limitation of the articulatory buffer to single syllable in these patients with relatively mild to moderate AOS, although it is possible that individuals with more severe AOS would reveal a reduction in buffer capacity.

Verbal Working Memory (vWM) Deficit Hypothesis. This hypothesis is similar to the reduced buffer capacity hypothesis in that it seeks an explanation in terms of memory rather than in terms of the programming process that creates the

input to the memory system. For example, in Baddeley's (1986) model of working memory, there is a visuospatial memory system, a verbal working memory system, and a central executive that allocates resources between these two systems. The verbal working memory system (vWM) includes a phonological storage component or buffer and an articulatory rehearsal loop that can be used to maintain activation of the items in storage. Whether or not vWM as measured by traditional tasks such as digit or word span is the same as the articulatory buffer assumed in models of speech production remains an issue of debate (e.g., Den Ouden, 2002; Klapp et al., 1979; Sternberg et al., 1978). There is evidence that speakers with AOS are impaired on such vWM tasks and essentially perform like controls under conditions of articulatory suppression (e.g., Waters, Rochon, & Caplan, 1992), suggesting a relation between vWM and speech motor programming.

The fact that simple RT paradigms (including the self-selection paradigm) involve a demand on memory due to the delay interval raises the question of whether the pattern of results for the patients with AOS can be accounted for by deficits in vWM. For example, an inability to maintain a programmed response in a buffer may lead to additional attempts to reprogram the buffer (as was suggested for the reduced buffer capacity hypothesis) or to early clearing of the buffer, predicting increased ST or a disproportionate number of premature responses, respectively<sup>10</sup>. Or, a sequence of syllables may decay rapidly after it has been programmed and loaded into the buffer,

---

<sup>10</sup> Inspection of error type distributions for individual patients suggested that premature responses formed a majority of errors only for AOS4 and APH1; for the other patients, a majority of errors were incorrect responses. In addition, premature responses formed a substantial proportion of total errors in the younger control group as well, suggesting that premature responses need not reflect working memory deficits.

reducing the sequence to a single syllable and thus explaining the absence of a sequence length effect on RT (note that this is essentially the same as the reduced buffer capacity hypothesis).

Although we did not assess vWM in our participants, there are several reasons why this account seems unlikely. First, the delays used in this experiment were relatively short (less than 1200 ms), and were not filled with distracting tasks, which should have minimized memory demands. However, since we have no data on vWM abilities in these participants it remains a possibility that their vWM was so severely impaired that even these delays were challenging. Second, if reprogramming were necessary due to failure to maintain a programmed response over a delay interval, one would expect to see longer RT in patients with AOS than in controls (who presumably have the response programmed and ready for action in the buffer); this is contrary to our findings. Third, if reprogramming occurred, then the sequence length should have (re)emerged on RT for these patients, which was not the case. Fourth, if reprogramming of responses occurred during the ST interval, this would predict a disproportionate sequence length effect on ST, since presumably reprogramming a longer, more complex response would take longer than reprogramming a shorter, simpler response (if both a single syllable and a sequence are reprogrammed once, the sequence length effect should double). Finally, even if it can be demonstrated that these patients also have vWM deficits, it would seem more logical to assume that the motor programming deficit is responsible for the vWM deficit than vice versa. That is, errors of spatiotemporal coordination (e.g. speech sound distortions) do not follow

naturally from impairment of the phonological store (unless one assumes that this store involves a motor-based code rather than a phonologically-based code); instead, a deficit in motor programming is likely to impede articulatory rehearsal, thus preventing adequate updating of activation of the items in the phonological store with the resultant decay and impaired vWM performance.

Dual Route Hypothesis. According to the Dual Route hypothesis (Varley & Whiteside, 2001a,b), speakers with AOS have lost access to the direct route of speech motor programming and instead are dependent on the indirect, assembly-based route. In the Nijmegen model, syllable programs are the maximum unit size and can be retrieved as single units via the direct route (mental syllabary), whereas the indirect route involves phoneme-by-phoneme assembly to create syllable motor programs. However, in Varley & Whiteside's (2001a,b; Varley et al., 2006) conceptualization of speech motor programming, there is no restriction (other than frequency of occurrence) on unit size, creating the possibility that syllable sequences can be programmed as single units (direct route) or as a series of separate syllable-sized units (indirect route).

Notice that this idea corresponds closely to the distinction between INT (direct route) and SEQ (indirect route): If a sequence is produced often enough, it will become a single chunk and can be programmed by the direct route (INT), whereas sequences that are not produced very often will remain separate syllable-sized programs and are programmed by the indirect route (SEQ). One critical difference between these views is that the SEQ process cannot be preprogrammed and must

occur after the decision to initiate speech, whereas the concept of an indirect route does not in itself imply that this process must await initiation. If the indirect route can be completed prior to initiation, then this route essentially falls in the INT domain and ceases to be similar to the SEQ process. Another difference is that the Dual Route hypothesis is restricted to speech motor programming, whereas the INT/SEQ model has applicability to other, nonspeech, motor behavior as well.

The Dual Route account for the performance of the AOS group relative to controls would be that the sequences cannot be programmed as single units but must be computed anew each time via the indirect route, resulting in longer ST, and a sequence length effect on ST. The assumption that indirect route encoding can be completed prior to the go-signal accounts for the absence of a sequence length effect on RT. Thus, the present experiment does not distinguish this version of the Dual Route hypothesis from the INT/SEQ model. Future studies aimed at dissociating these accounts might include frequency manipulations in addition to complexity or sequence length manipulations. Furthermore, investigation of nonspeech impairments in patients thought to have a direct route deficit might be able to decide between the two accounts, in that only the INT/SEQ model would predict a nonspeech deficit. Experiment 2 examined motor programming of finger movements in these patients, and thus may provide information pertinent to the distinction between these two hypotheses.

Phonological Encoding Deficit. It was argued that a disproportionate effect of single syllable duration in AOS would be a clear indication of a motor level deficit, as

opposed to a phonological level deficit. Since no effects of single syllable duration or interactions between group and duration were observed, we failed to find the unambiguous evidence for a motor programming deficit in AOS. We did obtain longer ST for the AOS group than for the controls, as well as a sequence length effect in both groups. However, the fact that the number of syllables is also a relevant aspect of processing at the phonological encoding stage (e.g., Levelt et al., 1999) suggests the possibility that the longer ST for the patients with AOS actually reflects a phonological encoding deficit, for example a problem in retrieving metrical frames or segments, or in prosodification (associating the segments to their syllable nodes).

While this possibility cannot be ruled out entirely, there are several reasons to doubt it. First, the primary diagnosis for these patients was AOS. Although it is possible that these patients also had a phonological disorder, there was no clear evidence to suggest this was the case. Notably, the main speech characteristic indicative of phonological impairment, i.e. phoneme order errors (McNeil et al., 1997), was not observed during assessment and informal conversation. Second, the patients with aphasia who did show positive evidence for phonological impairment without AOS (APH1 and APH2) did not show increased ST relative to controls. This might be considered somewhat surprising, since a deficit in preparation for speech at preceding levels could also affect ST. The absence of significant effects on ST might reflect the relatively mild nature of these patients' deficit (i.e. these patients evinced little or no phonological encoding difficulties in this experiment as indicated by virtual absence of phonemic errors), or it might reflect the possibility that phonological

processing difficulties do not increase processing time, or it could be that these patients did not fully preprogram the responses. This latter possibility is consistent with the fact that RT was generally faster in these two patients than in age-matched controls. This possibility further predicts that these patients should make more errors than controls. While no acoustic analyses were available for these patients and thus, timing accuracy could not be assessed, there were no differences in accuracy as determined based on the number of rerun trials.

Stimulus-Response Mapping. Since the ST interval includes processes relating to visual stimulus processing and response selection in addition to response programming, it is possible that the longer ST in the patients with AOS reflects deficits at these preceding levels. However, stimulus processing seems an unlikely candidate since in that case we would have also seen longer RTs (in response to the go-signal). Thus, the remaining option is that perhaps there is a difficulty in AOS in mapping the stimulus to a particular response. One might suppose that this would likely lead to producing the incorrect pattern. While these errors did occur, they appeared to be more frequent for the patients with aphasia, who did not show longer ST. Furthermore, this hypothesis does not explain the presence of timing errors as were observed in two of the patients. However, this account cannot be ruled out on the basis of these data.

Self-monitoring / Self-judgment. Another possible explanation for some of the results relates to the ability to monitor one's own speech and/or judge one's own state of readiness for action. This is a relevant factor to consider (even though it is difficult

to quantify), given the nature of the self-selection paradigm in which participants are supposed to indicate when they are ready to respond. According to this hypothesis, a major component of ST is the time it takes to monitor one's response and self-determine when one is ready to respond. As such, it is possible that the increase in ST for patients with AOS or aphasia is not caused by motor programming difficulties but by more conservative estimates of one's readiness to respond. The longer time needed to indicate readiness might relate to difficulties in monitoring one's internal speech for potential errors, or to a history of speech production difficulties (as a result of AOS or aphasia) that makes speakers less confident about their speech.

For this hypothesis to capture the data it must further be assumed that the self-judgment interacts with response characteristics such as sequence length, otherwise no sequence length would be observed on ST. Thus, at some level of processing there must be a representation of the task that encodes its complexity or anticipated production difficulty. The precise level(s) at which these task representations are monitored remains to be determined; possibilities include the phonological encoding stage and the motor programming stage.

This hypothesis can potentially explain the longer ST for APH3. Although there was no group effect on ST for the patients with aphasia (contrary to the findings for the AOS group), one of the patients (APH3) did show longer ST than controls. Her aphasia was very mild, but she complained that speaking was much more effortful since her stroke and expressed concern and frustration with this change in her language abilities. An interesting observation is that during the experiment, APH3



showed a substantial decrease in ST for single syllables from block 5 to block 6, after she was informed that the session was almost half-way completed. APH3 reacted surprised and commented that she would speed up, apparently shifting strategy to emphasize speed over accuracy. While these observations are anecdotal and disallow any firm conclusions, they do hint that strategic factors may also affect ST and may potentially explain differences between participants. An alternative explanation for why APH3 showed longer ST than controls is that she exhibited a more generalized reduction of processing speed (recall that APH3 had a history of a TBI prior to her stroke). While APH3 was the only patient with aphasia to show consistently longer RT than age-matched controls, these RT differences were not significant.

#### **5.4. Conclusions**

First, Experiment 1 provided evidence supporting the hypothesis that AOS involves a deficit at the preprogramming (INT) stage of processing, whereas the SEQ stage appears to be intact. By separating out preprogramming from response initiation, it was suggested that the often-reported initiation difficulties may in fact reflect preprogramming difficulties, and that once a response has been programmed, initiation itself is no longer problematic. The notion of a deficit in INT was based primarily on the finding of longer ST for individuals with AOS than for age-matched controls; two patients also demonstrated impaired absolute timing control of the sequences. Intact SEQ was inferred from the normal pattern of RT findings, i.e. no difference in mean RT and absence of sequence length effect on RT as in controls. While this pattern of

results was predicted from the hypothesis made based on the INT/SEQ model, alternative accounts were also discussed and found to be compatible with some aspects of the data.

Second, the results from the control speakers suggested that the repeated syllable sequences in this experiment were programmed as single units rather than as a series of separate syllable-sized motor programs. This inference was based on the presence of a sequence length effect on ST but not on RT. It is possible that the presence of a specific temporal pattern encouraged speakers to integrate the syllables into a sequence, unlike previous studies, which have generally not specified the temporal patterns of their target responses.

Finally, this study demonstrates the utility of the self-selection paradigm in the study of speech production. Specifically, by comparing ST and RT in the same trial, we were able to show that the syllable sequences used in this experiment were programmed as a single unit. In addition, the paradigm was able to capture a dissociation between processing stages by virtue of differential impairment in AOS, thus providing converging evidence for the INT/SEQ model from a divergent, neuropsychological source.

### **Acknowledgements**

Chapter 5 is being prepared for publication in a manuscript entitled “Motor Programming in Apraxia of Speech”, by Edwin Maas, Donald A. Robin, David L. Wright, and Kirrie J. Ballard. The dissertation author is the primary investigator and author of the manuscript.

## **6.0. Nonspeech Motor Programming in AOS**

Experiment 1 provided evidence in support of the hypothesis that AOS involves a deficit in speech preprogramming or INT and that SEQ and speech initiation are intact. The primary purpose of Experiment 2 was to further detail the nature of the motor programming deficit in apraxia of speech (AOS). Specifically, the following questions were addressed: 1) Do speakers with AOS show evidence for impairments in motor programming of nonspeech movements that require fine-grained temporal specification, and 2) If so, is this deficit localized to the INT motor programming process as it was shown to be in speech production?

Given that the INT/SEQ model that drove the hypotheses underlying the present experiments has been used to explain both speech and nonspeech motor programming (Klapp, 1995, 2003), it is possible that the INT stage includes a central (effector-independent) component that underlies both speech and nonspeech motor programming. For example, if INT involves activating generalized motor programs (GMPs) and specifying their parameters (Schmidt, 1975, 2003; Schmidt & Lee, 2005), and GMPs are indeed effector-independent, then a deficit in the INT process in AOS should also affect nonspeech (finger) movements that are based on the same GMP (unless the deficit is in the effector-parameter). Thus, if those patients with AOS who show evidence of an INT deficit in speech also show an INT deficit in nonspeech tasks, this would not only further narrow down the level of the deficit in AOS, but would also suggest that the INT process (or at least, parts thereof) is shared among

different motor control tasks rather than each task being programmed by its own separate, modality-specific INT process.

Thus, to test the hypothesis that AOS involves a central deficit in the INT process, this second experiment also used the self-selection paradigm, and involved non-speech (finger) movements only, i.e. the same finger tap responses as used by Klapp (1995) and Wright and colleagues (Immink & Wright, 2001; Wright et al., 2004). Examining the responses that provided the original basis for the distinction between INT and SEQ will provide the clearest test of the hypothesis of a central INT deficit in AOS, and will allow us to further narrow the range of possible alternative accounts for the findings of Experiment 1. Most of the alternative accounts discussed in Chapter 2 are suited to explain deficits in speech production, but provide no mechanism that would predict deficits in nonspeech tasks. For instance, the Dual Route hypothesis (Varley & Whiteside, 2001a,b) in its current form is restricted to explaining speech deficits. Similarly, an explanation of the deficit in AOS in terms of phonological encoding deficits could account for the speech findings in Experiment 1 but would not predict, or account for, any deficits in nonspeech motor control.

The hypothesis that speakers with AOS have a motor programming deficit that extends beyond speech (e.g., Ballard et al., 2000, 2003) predicts that these speakers will produce more errors and show increased programming time in this nonspeech task, relative to control subjects. Specifically, the hypothesis of a central INT deficit, not an SEQ deficit, in AOS makes several predictions. First, INT programming time (reflected in Study Time) should be longer than for control subjects. Second, there

should be a disproportionate effect of unit complexity on INT for speakers with AOS relative to control speakers. Third, since SEQ is hypothesized to be intact, there should be no differences in Reaction Time between speakers with AOS and control speakers. In contrast, if the deficit in AOS is specific to speech only (e.g., Ziegler, 2003a), then there should be no differences between speakers with AOS and control speakers in terms of errors and programming time.

A secondary purpose of this experiment was to further examine the absence of a syllable duration effect on ST for the younger controls in Experiment 1. This finding was unexpected given the fact that finger movement studies have shown effects of button press duration on ST (e.g., Immink & Wright, 2001; Magnuson et al., in press; Wright et al., 2004). In addition to the possibility that syllable duration is not an appropriate metric of motor program complexity in speech, two methodological explanations were discussed, namely the feedback hypothesis and the RT window hypothesis. According to the feedback hypothesis, participants in Experiment 1 did not emphasize accuracy of these responses due to the fact that no detailed feedback was provided on their timing accuracy (unlike in the finger movement studies). According to the RT window hypothesis, participants postponed some of their programming due to the fact that they had sufficient time after the go-signal to respond without eliciting an error. In Experiment 2, the RT window was the same as in Experiment 1, but the use of button press responses allowed for the provision of detailed feedback. Thus, if a duration effect emerges in Experiment 2, this would support the idea that the absence of detailed feedback was responsible for the lack of a duration effect in Experiment 1.

## **6.1. Methods**

### **6.1.1. Participants**

The same four patients with AOS as in Experiment 1, as well as aphasic patients APH1 and APH2 participated in Experiment 2. Again it is recognized that an N of 2 is extremely limiting; however only these two subjects agreed to participate in this phase of the experiment. Participant information is provided in Table 6-1. This experiment was conducted approximately one year prior to Experiment 1, and thus the age and time post onset differ slightly from those reported in Table 5-1. Language and speech profiles had remained stable for all patients, except APH1, who at the time of Experiment 2 had more severe comprehension problems, as well as a homonymous hemianopia and unilateral neglect that had resolved at the time of Experiment 1. Although the experiment was attempted, it was discontinued after several blocks due to frustration. Thus, data were available for the four patients with AOS and for APH2.

In addition to these patients, 17 unimpaired monolingual control participants were recruited through the subject pool of the School for Speech, Language, and Hearing Sciences of San Diego State University. Two control participants were unable to perform or learn the task, and were therefore not included in any analysis. In addition, one control participant failed to appear for retention testing, and one participant reported a history of dyslexia; as a result, these participants were also excluded from all analyses. Thus, the final control group consisted of 13 individuals (10 women, 3 men), with a mean age of 23 years (range = 19-41). All were right-

handed, and participated for partial course credit. All participants provided informed consent in accordance with local Institutional Review Board procedures.

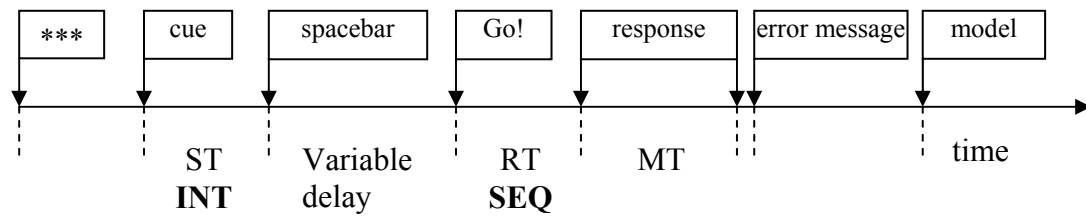
### **6.1.2. Task and procedures**

Experiment 2 used the self-selection paradigm, with four different key-press responses as targets. The four target responses were identical to those used by Klapp (1995) and Wright and colleagues (Immink & Wright, 2001; Wright et al., 2004), and involved pressing the F-key of a computer keyboard. Depression of the F-key produced a tone that indicated press duration. Each response was associated with the same cues as in Experiment 1 (1S = single short key press, 150 msec; 1L = single long key press, 450 msec; 4S = sequence of four key presses, 150-450-450-150 with 100 msec interpress intervals; 4L = sequence of four key presses, 450-150-150-450, with 100 msec interpress intervals). The experiment (trial events and data collection) was controlled by Micro Experiment Lab software (Psychology Software Tools, Inc.) run in DOS on a Sharp laptop computer either at the participant's home, the university's Communications Clinic, or available labrooms.

**Table 6-1.** Participant information for Experiment 2.

	<b>Sex</b>	<b>Age</b>	<b>Hand</b>	<b>TPO</b>	<b>Aphasia</b>	<b>AOS</b>	<b>Oral/Limb apraxia</b>	<b>Dysarthria</b>
AOS1	M	68	L	2;6	Mild nonfluent	Mild-mod.	None/None	Mild unilateral weakness
AOS2	M	27	R	6;1	Mild-mod. nonfluent	Mild	None/None	None
AOS3	F	72	R	2;7	Mild anomia	Mild-mod.	None/None	None
AOS4	M	58	R	1;8	Mild anomia	Mild	None/None	Mild unilateral weakness
		<i>56 (20)</i>		<i>3;3 (2;0)</i>				
APH1	M	73	R	1;1	Mod. fluent	None	Mild/None	None
APH2	M	54	L	6;0	Mild-mod. fluent	None	Mild/None	Mild unilateral weakness
YCON (N=13)	10F 3M	<i>23 (6)</i> <i>(19-41)</i>	13R					





**Figure 6-1.** Self-selection paradigm adapted to represent trial events as used in Experiment 2. ST = Study Time; RT = Reaction Time; MT = Movement Time. See text for details.

The sequence of events on each trial was similar to those in Experiment 1 (see Figure 6-1). Each trial started by presenting the word “READY” in the center of the screen for 500 msec, which was immediately followed by the cue indicating the required response (1S, 1L, 4S, or 4L). Participants prepared the required response and pressed the space bar with their left thumb as soon as they were ready to respond<sup>11</sup>. After a variable delay (ranging between 800 and 1200 msec), the go-signal (“Go!”) was presented for 300 ms in the center of the screen, which prompted participants to execute the response as quickly and accurately as possible using their left index finger on the F-key.

Following incorrect responses, an error message accompanied by three 500-msec tones of 3000 Hz was presented. Error messages could be any one of the following: 1) “Slow start error”, elicited by responses initiated more than 1000 msec after the go-signal, 2) “Wait for go-signal”, elicited when response initiation preceded the go-signal (the go-signal was provided after this message), 3) “Too short” or “Too

<sup>11</sup> The choice to use only the left hand for all button presses (in contrast to Immink & Wright, 2001, and Wright et al., 2004, who had participants use the right hand to press the “end” key to end the Study Time interval) was motivated by the potential right hemiparesis in the clinical groups and the desire to keep conditions as similar as possible between groups.

long”, indicating that the overall response duration exceeded a prespecified range of acceptability (100 msec above or below target duration for the single presses, 500 msec for the sequences), 4) “Pause too long”, which indicated that the interpress interval exceeded 200 msec (thus, this error message only applied to the sequences), and 5) “Incorrect response”, elicited when a key other than the F-key was pressed. After the feedback message, an auditory model consisting of 1000 Hz tones with the message “The correct response =” was presented to indicate what the correct response should have been. The auditory model and visual message were also presented after correct trials, and were provided to help participants learn the targets.

Experiment 2 also consisted of two phases: an acquisition phase and a retention phase, each conducted on separate days, with retention testing at either 24 hours or 48 hours after acquisition. Presentation of target responses was random during both acquisition and retention. Acquisition practice involved 12 blocks; each block terminated when four correct productions of each target had been produced. Incorrect responses were rerun at the end of each block, to ensure equal number of data points for all participants. Thus, given that there were four target responses, a block consisted of a minimum of 16 trials, plus however many errors were produced. There was a 20-second rest interval between blocks. Retention testing involved one block, without feedback and without rerunning of incorrect trials, for a total of 16 trials (four trials for each target response).

After providing informed consent (and after assessment testing for the patients), the experimenter explained the general nature of the experiment, and

provided detailed instructions about the specific events of the experimental task. During these instructions, associations between cue and target response were explained, as were the possible error messages that may be elicited during the experiment. Participants were also explained that incorrect trials would be rerun at the end of each block, and thus that the length of the session depended on their performance. Throughout the instructions, questions were encouraged and answered by the experimenter. When the participant indicated understanding of the instructions, the experiment began with a presentation of the auditory response models; each target response model was presented four times in a row with its visual cue, in the order 1S, 1L, 4S, 4L. Immediately following the models, the first acquisition block began. The first block involved reiteration of instructions and a few demonstrations of the correct responses by the experimenter, in order to ensure adequate understanding of the task. For this reason, the first block was considered a warm-up block, and responses for this block were excluded from the analyses.

### **6.1.3. Design, Analysis, and Predictions**

The design, dependent measures, and analyses were the same as for Experiment 1, except that no group ANOVAs were performed due to unequal group sizes. If the deficit in AOS is specific to the speech motor control system (e.g., Ziegler, 2003a,b), then there should be no difference in accuracy (number of rerun trials and timing accuracy) between patients with AOS and control participants, nor should there be any differences in ST or RT. However, if AOS reflects a central deficit

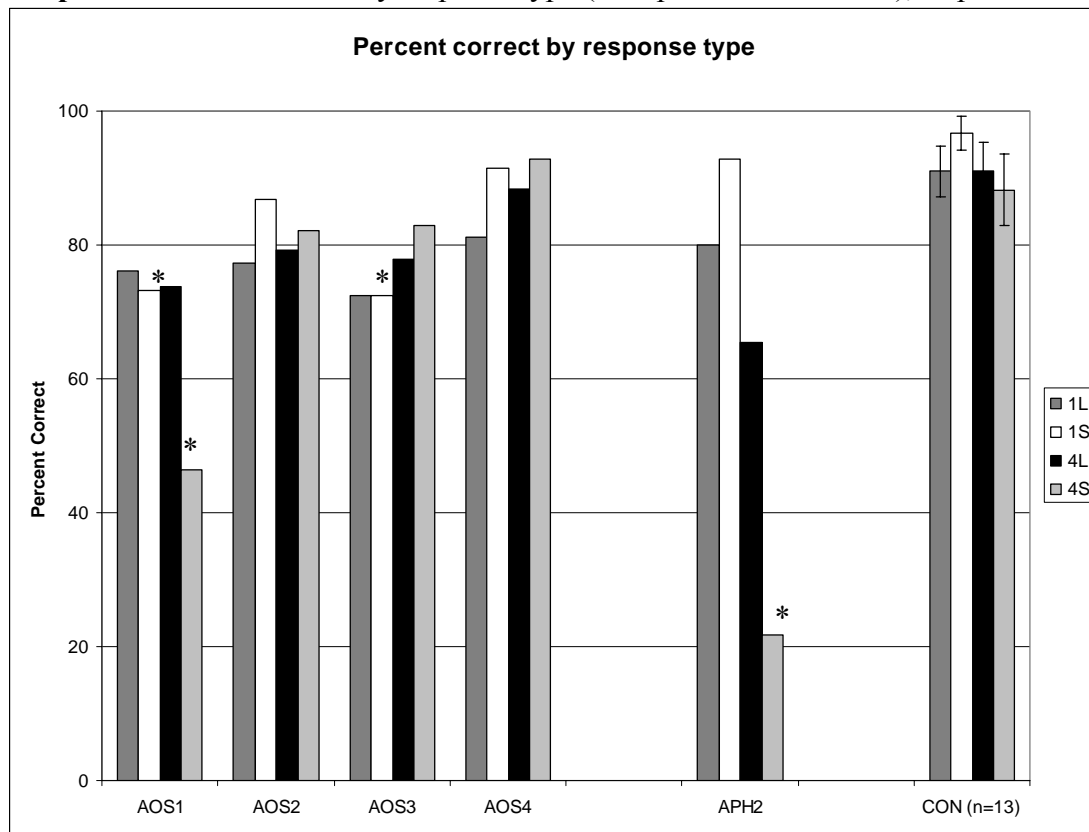
in the programming and control of timing (e.g., Ballard et al., 2003), then a greater number of errors (rerun trials) and larger timing error values are expected for patients with AOS relative to control participants, and, most importantly, ST should be longer for patients with AOS than for unimpaired controls. In addition, there should be a disproportionate effect of complexity (single press duration). No group differences were expected for RT.

## **6.2. Results**

Time taken to complete the acquisition component (Day 1) of the experiment varied between subjects. For the controls, acquisition was typically completed between 50 and 75 minutes. AOS1, AOS3, and AOS4 each completed acquisition within 100 minutes, whereas AOS2 took approximately two hours for completion. APH2 took approximately two hours and fifteen minutes to complete acquisition, and several 4S responses were completed by the experimenter due to great difficulty for the patient. Responses made by the experimenter were excluded from the analyses.

### **6.2.1. Accuracy: Number of Error Trials**

As a first step, the number of errors produced during acquisition (i.e. the number of rerun trials) was used to calculate a percent accuracy score for each participant for each response type separately (collapsed across blocks) (see Graph 6-1). The total accuracy percentage (collapsed across response type) for each group by block is provided in Appendix C (Graph C-1; a breakdown by error type was not

**Graph 6-1.** Percent correct by response type (collapsed across blocks), Experiment 2.

available). Data for the control participants were submitted to a 2 (Sequence Length) x 2 (Duration) x 11 (Block) repeated measures ANOVA. This analysis revealed significant effects of Sequence Length ( $F[1,12] = 7.72, p < .05$ ), Block ( $F[10,120] = 5.67, p < .0001$ ), and Sequence Length x Duration ( $F[1,12] = 6.78, p < .05$ ). This pattern indicated greater accuracy for the single short press (1S) than for the 4S sequence, whereas there was no difference between the single long press and the 4L sequence.

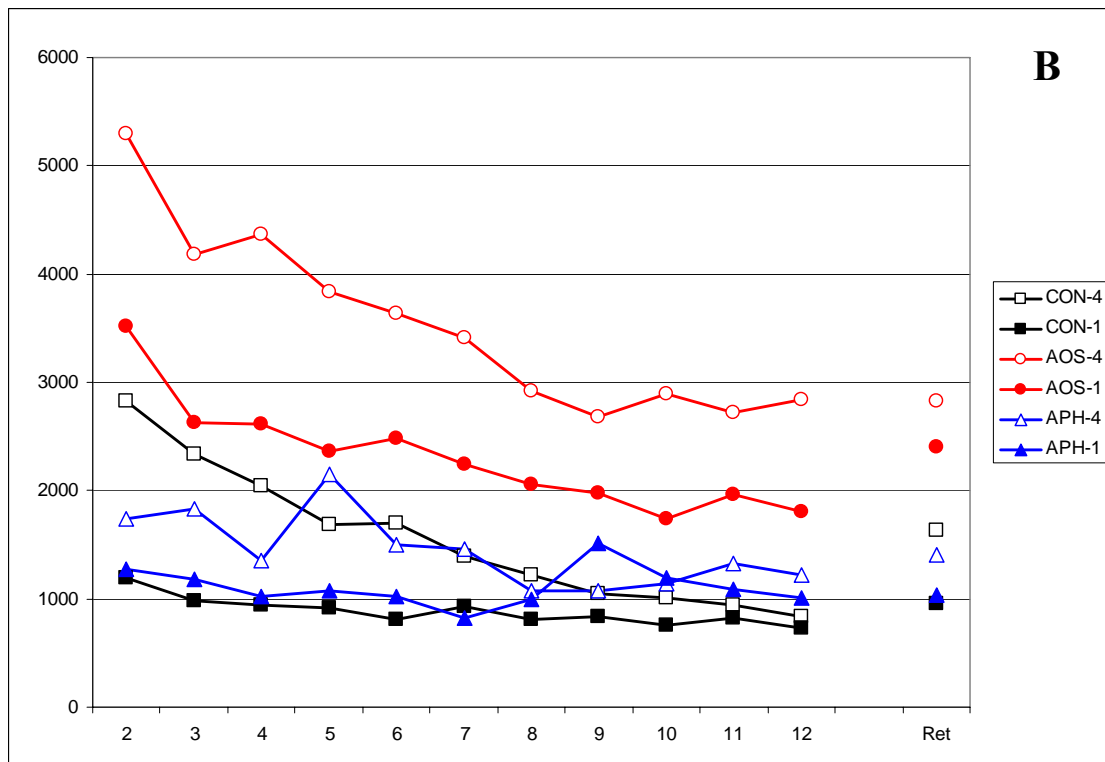
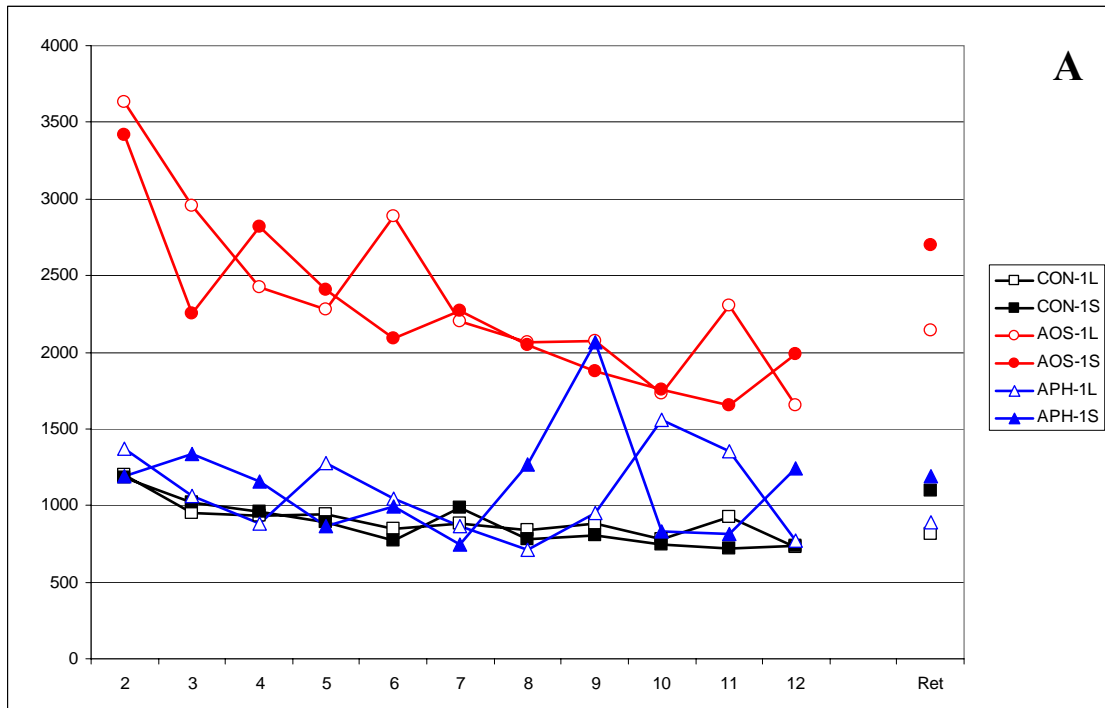
For the patients, statistical analysis using Crawford and Howell's (1998) modified t-test indicated that while the patients produced numerically more error trials

than controls (with the exception of the 4S responses for AOS4), only a few of these differences reached statistical significance (AOS1: 1S,  $t = -2.427$ ,  $p = .016$ ; 4S,  $t = -2.090$ ,  $p = .029$ ; AOS3: 1S,  $t = -2.524$ ,  $p = .13$ ;  $df = 12$ , one-tailed). The patient with aphasia also produced numerically more error trials than controls, significantly so for 4S responses ( $t = -3.333$ ,  $p = .003$ ;  $df = 12$ , one-tailed). See Appendix C (Table C-1) for comparisons for all patients.

### **6.2.2. INT: Study Time (ST)**

Prior to analysis, STs longer than 10 seconds and STs shorter than 100 msec. were removed as invalid data points. This resulted in a total data loss of 28 observations or 0.8% of all data (controls: 0.1%; AOS: 3.0%; APH: 1.1%). Analyses were performed on log-transformed means to meet the assumption of normality; however, untransformed means are presented for interpretability. ST results are presented in Graph 6-2.

**Graph 6-2.** Finger movement Study Time for duration, 1S vs. 1L (A), and for sequence length, 1 vs. 4 (B). CON = control group (N=13), AOS = apraxia of speech group (N=4), APH = patient with aphasia (N=1).



Controls. The overall 2 (Duration) x 2 (Sequence Length) x 11 (Block) repeated measures ANOVA for the acquisition data revealed significant main effects of Sequence Length ( $1 < 4$ ;  $F[1,12] = 20.34$ ,  $p < .001$ ) and Block ( $F[10,120] = 33.51$ ,  $p < .0001$ ), and significant interactions between Duration and Block ( $F[10,120] = 2.49$ ,  $p < .01$ ) and Sequence Length and Block ( $F[10,120] = 11.17$ ,  $p < .0001$ ). This pattern of results indicated that the sequence length effect was significant in blocks 2-7 but not in blocks 8-12.<sup>12</sup> The 2 (Duration) x 2 (Sequence Length) x 2 (Phase) retention analysis revealed significant main effects of Duration ( $F[1,12] = 9.80$ ,  $p < .01$ ), Sequence Length ( $F[1,12] = 24.48$ ,  $p < .0005$ ), and Phase ( $F[1,12] = 16.68$ ,  $p < .005$ ), as well as significant interactions between Duration and Phase ( $F[1,12] = 13.47$ ,  $p < .005$ ) and Sequence Length and Phase ( $F[1,12] = 5.85$ ,  $p < .05$ ). This pattern of results indicated that the sequence length re-emerged during retention, due to an increase in ST for the sequences only.

Separate analyses were performed on the single press data to assess the presence of a complexity effect (1S vs. 1L). For acquisition data, ANOVAs comparing short vs. long single presses for controls revealed only a main effect of Block ( $F[10,120] = 12.63$ ,  $p < .0001$ ), indicating longer ST in Block 2 than in all subsequent blocks, longer ST in Block 3 than in Blocks 6 and 8-12, longer ST in Block 4 than in Blocks 10 and 12, and longer ST in Blocks 5 and 7 than in Block 12. The effect of Duration was not significant ( $F[1,12] = 2.22$ ,  $p = .162$ ), despite a numerical difference in the predicted direction (1L mean = 901, SD = 369; 1S mean = 874, SD = 309). The

---

<sup>12</sup> Separate analyses comparing 1S to 4S and 1L to 4L both revealed the same pattern: main effects of sequence length and block and a significant interaction.



Duration x Block interaction was not significant ( $F[10,120] = 1.70, p > .05$ ). Retention analyses on duration of single presses revealed a significant main effect of Duration ( $F[1,12] = 5.06, p < .05$ ), indicating longer ST for 1S (mean = 915, SD = 449) than for 1L (mean = 770, SD = 260). In addition, there was a main effect of Phase ( $F[1,12] = 9.32, p < .05$ ), indicating longer ST at retention testing (mean = 951, SD = 439) than during the last acquisition block (mean = 734, SD = 248). While most of this increase was carried by the 1S response, the Duration x Phase interaction failed to reach significance ( $F[1,12] = 3.54, p = .084$ ).

Patients. Separate analyses were carried out comparing each patient to the control group, using Crawford & Howell's (1998) modified t-test (see Tables 6-2 and 6-3). ST comparisons for Duration (1S vs. 1L), collapsed across blocks, showed that during both acquisition and retention, three of the four patients with AOS differed significantly from the control group (all  $t$ -values  $> 2, ps < .05; df=12$ ), while the difference for the fourth patient (AOS1) failed to reach significance, but was in the predicted direction (1S:  $t = 1.688, p = .051$ ; 1L:  $t = 1.773, p = .059; df=12$ ) (see Table 6-2). At retention, all four patients with AOS demonstrated significantly longer ST than controls for both short and long presses (all  $t$ -values  $> 1.7, ps < .05; df=12$ ). In contrast, the patient with aphasia did not differ from controls on ST at either acquisition or retention. There was no evidence for disproportionate complexity effects on ST using the Revised Standardized Difference Test (Crawford & Garthwaite, 2005)

**Table 6-2.** ST comparisons for each patient against the control sample for duration of single presses. Analyses are based on log-transformed means; however, untransformed means are presented to enhance interpretability.

Subject	Duration	Acquisition	t-value	p-value	Retention	t-value	p-value
AOS1	1L	1622 (515) <sup>^</sup>	1.773	0.051	1526*	1.786	0.050
	1S	1541 (675) <sup>^</sup>	1.688	0.059	2381*	1.958	0.037
	Δ	81	0.517	0.615	-855	0.151	0.883
AOS2	1L	2009 (1170)*	2.169	0.025	2766*	3.321	0.003
	1S	2035 (848)*	2.474	0.015	3181*	2.609	0.011
	Δ	-26	1.837	0.091	-415	0.626	0.543
AOS3	1L	3589 (550)*	4.107	0.001	2215*	2.747	0.009
	1S	3084 (668)*	3.852	0.001	2991*	2.471	0.015
	Δ	505	1.537	0.150	-776	0.243	0.812
AOS4	1L	2304 (585)*	2.805	0.008	2073*	2.577	0.012
	1S	2272 (542)*	2.947	0.006	2234*	1.815	0.047
	Δ	32	0.862	0.406	-161	0.670	0.516
APH2	1L	1078 (275)	0.647	0.265	890	0.394	0.350
	1S	1138 (371)	0.858	0.204	1193	0.404	0.347
	Δ	-60	1.280	0.225	-303	0.009	0.993
CON (N=13)	1L	901 (278)			809 (263)		
	1S	874 (244)			1093 (538)		
	Δ	27			-284		

\*  $p < .05$ ,  $df=12$ , one-tailed (Crawford & Howell, 1998).

<sup>^</sup>  $p < .10$ ,  $df=12$ , one-tailed (Crawford & Howell, 1998).

For the Sequence Length comparisons, two of the four patients with AOS (AOS3 and AOS4) showed significantly longer ST for both single presses and sequences during acquisition ( $ts > 2$ ,  $ps < .05$ ;  $df=12$ ) (see Table 6-3); for AOS2, only single-press ST was significantly longer than in controls ( $t= 2.409$ ,  $p < .05$ ) while the numerically longer ST for sequences failed to reach significance ( $t= 1.453$ ,  $p= .086$ ). Patient AOS1 did not differ significantly from controls in ST for either sequences ( $t= 0.446$ ,  $p > .10$ ) nor for single presses ( $t= 1.758$ ,  $p = .052$ ). Again, the patient with

aphasia did not have longer ST than controls either during acquisition or during retention.

**Table 6-3.** ST comparisons for Sequence Length (collapsed across blocks) for each patient. Analyses were performed on log-transformed means; however, untransformed means are presented to enhance interpretability.

Subject	Number	Acquisition	t-value	p-value	Retention	t-value	p-value
AOS1	4	1692 (666)	0.446	0.332	1669	0.289	0.389
	1	1582 (493) <sup>^</sup>	1.758	0.052	1954*	2.145	0.027
	Δ	110 <sup>^</sup>	1.959	0.074	-285 <sup>^</sup>	1.982	0.071
AOS2	4	3051 (1555) <sup>^</sup>	1.453	0.086	2893	1.090	0.148
	1	2037 (866)*	2.409	0.017	2973*	3.304	0.003
	Δ	1014	1.434	0.177	-80 <sup>¶</sup>	2.358	0.036
AOS3	4	5638 (903)*	2.757	0.009	3821 <sup>^</sup>	1.496	0.080
	1	3336 (435)*	4.010	0.001	2547*	2.877	0.007
	Δ	2302 <sup>^</sup>	1.873	0.086	1274	1.482	0.164
AOS4	4	3727 (445)*	2.003	0.034	2937	1.113	0.144
	1	2286 (535)*	2.865	0.007	2142*	2.399	0.017
	Δ	1441	1.295	0.220	795	1.380	0.193
APH2	4	1443 (343)	0.220	0.415	1410	0.045	0.482
	1	1107 (181)	0.796	0.221	1041	0.407	0.345
	Δ	336	0.867	0.403	369	0.390	0.703
CON (N=13)	4	1548 (662)			1635 (950)		
	1	888 (259)			951 (342)		
	Δ	660			684		

\*  $p < .05$ ,  $df=12$ , one-tailed (Crawford & Howell, 1998).

<sup>^</sup>  $p < .10$ ,  $df=12$ , one-tailed (Crawford & Howell, 1998).

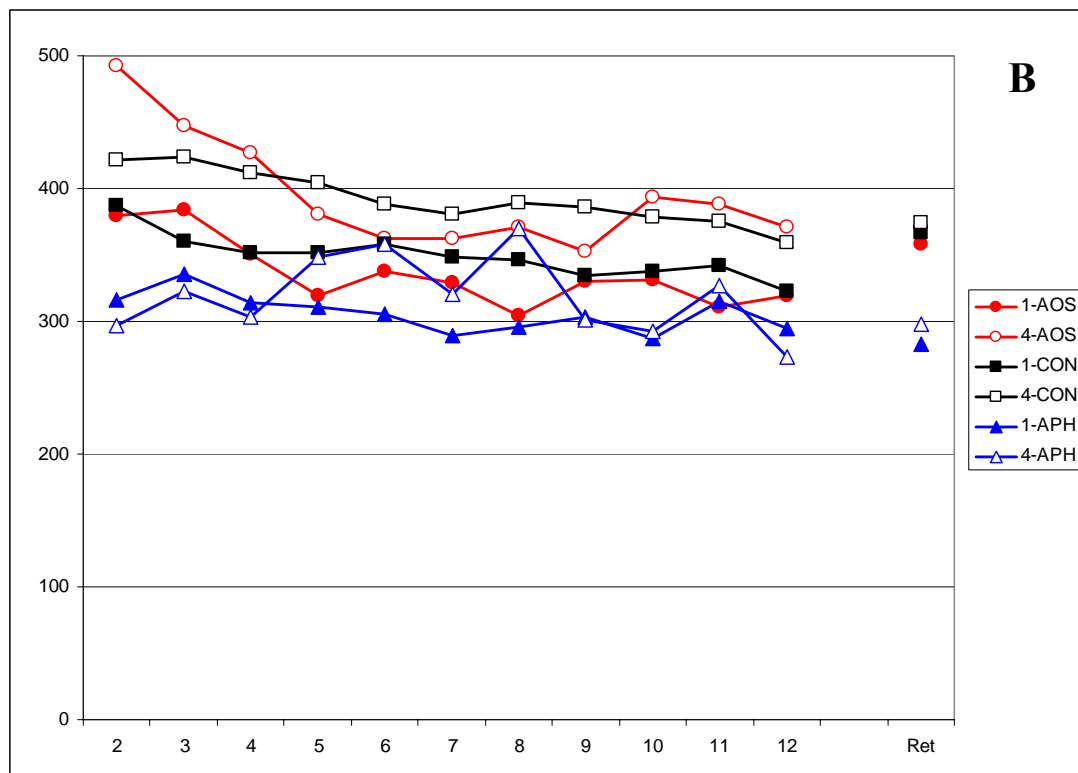
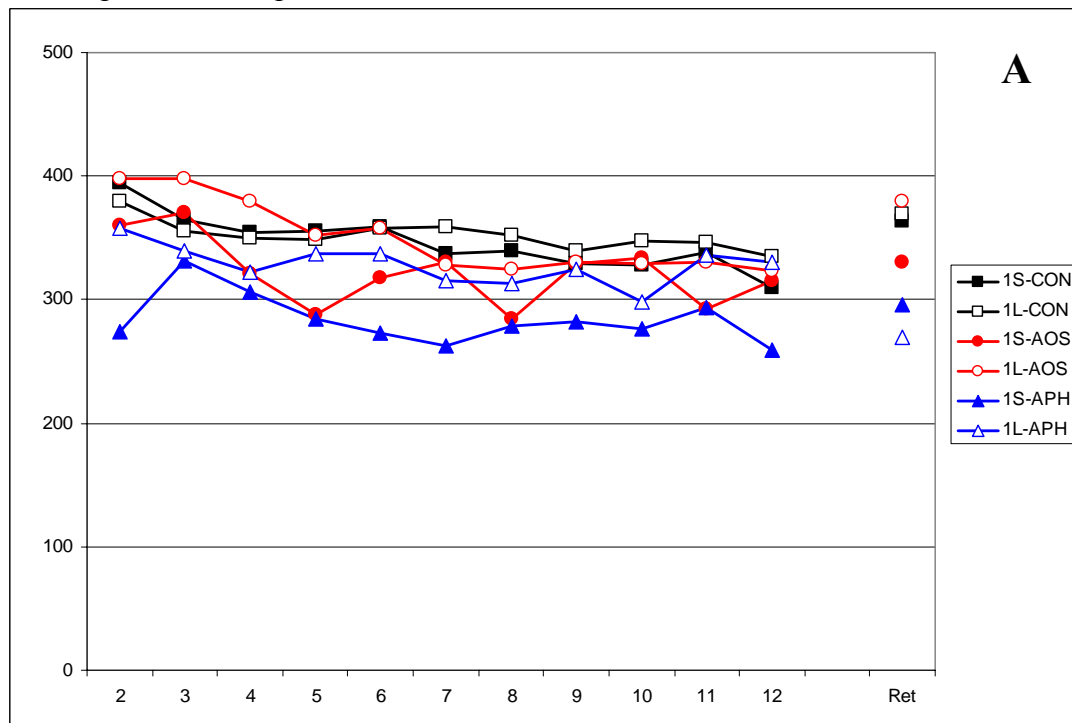
<sup>¶</sup>  $p < .05$ ,  $df=12$ , two-tailed (Crawford & Garthwaite, 2005).

### 6.2.3. SEQ: Reaction Time (RT)

Prior to analysis, RTs shorter than 100 ms were removed as invalid data points (anticipations), which resulted in a data loss of 4 observations (0.1%, all for controls).

Analyses were performed on untransformed means. Data are presented in Graph 6-3.

**Graph 6-3.** Finger movement RT for duration, 1S vs. 1L (A) and for sequence length, 1 vs. 4 (B). CON = control group (N=13), AOS = apraxia of speech group (N=4), APH = patient with aphasia.



Controls. The 2 (Duration) x 2 (Sequence Length) x 11 (Block) ANOVA on acquisition data revealed significant main effects of Sequence Length ( $1 < 4$ ;  $F[1,12] = 14.77$ ,  $p < .005$ ) and Block ( $F[10,120] = 5.45$ ,  $p < .0001$ ) and a marginal effect of Duration ( $L < S$ ;  $F[1,12] = 4.04$ ,  $p = .0674$ ). In addition, there was a significant interaction of Sequence Length and Duration ( $F[1,12] = 14.49$ ,  $p < .005$ ), and a marginal three-way interaction ( $F[10,120] = 1.85$ ,  $p = .0587$ ). This pattern of results indicated that the sequence length effect was greater for the 1S-4S comparison than the sequence length effect for the 1L-4L comparison, although the sequence length effect was significant in both cases. Retention analysis detected significant effects of Sequence Length ( $1 < 4$ ;  $F[1,12] = 4.97$ ,  $p < .05$ ) and Phase (Acquisition  $<$  Retention;  $F[1,12] = 5.63$ ,  $p < .05$ ), as well as a significant Sequence Length by Duration interaction ( $F[1,12] = 18.99$ ,  $p < .001$ ). This pattern of results indicated that ST for the 4S response was longer than for the other three responses, which did not differ from each other.

Separate analysis for single press duration did not reveal an effect of Duration ( $F[1,12] = 1.56$ ,  $p = .236$ ), as expected. There was a significant effect of Block ( $F[10,120] = 3.83$ ,  $p < .0005$ ), indicating that RT in Block 2 was longer than RT in Blocks 8-12. Retention analyses revealed only a main effect of Phase ( $F[1,12] = 5.74$ ,  $p < .05$ ), indicating longer RTs at retention (mean = 367, SD = 103) than during the last acquisition block (mean = 323, SD = 71).

**Table 6-4.** RT comparisons for Duration (collapsed across blocks) for each patient.

Subject	Duration	Acquisition	t-value	p-value	Retention	t-value	p-value
AOS1	1L	349 (64)	- 0.044	0.483	299 (92)	- 0.625	0.272
	1S	306 (94)	- 0.622	0.273	298 (28)	- 0.630	0.270
	Δ	43 <sup>^^</sup>	2.103	0.057	1	0.010	0.992
AOS2	1L	362 (88)	0.148	0.442	448 (142)	0.705	0.247
	1S	339 (81)	- 0.109	0.458	408 (113)	0.420	0.341
	Δ	23	0.946	0.363	40	0.580	0.573
AOS3	1L	364 (101)	0.178	0.431	468 (72)	0.883	0.197
	1S	339 (74)	- 0.109	0.458	333 (32)	- 0.296	0.386
	Δ	25	1.054	0.313	135 <sup>¶¶</sup>	2.366	0.036
AOS4	1L	326 (41)	- 0.385	0.353	302 (13)	- 0.598	0.281
	1S	305 (26)	- 0.637	0.268	284 (23)	- 0.763	0.230
	Δ	21	0.927	0.372	18	0.337	0.742
APH2	1L	328 (39)	- 0.356	0.364	269 (40)	- 0.892	0.195
	1S	284 (28)	- 0.964	0.177	296 (51)	- 0.649	0.264
	Δ	44 <sup>¶¶</sup>	2.212	0.047	-27	0.495	0.629
CON (N=13)	1L	352 (65)			369 (108)		
	1S	346 (62)			364 (101)		
	Δ	6			5		

<sup>¶¶</sup> p < .05, df=12, two-tailed (Crawford & Garthwaite, 2005).

<sup>^^</sup> p < .10, df=12, two-tailed (Crawford & Garthwaite, 2005).

Patients. Comparisons of individual patients against the control group are presented in Table 6-4 (duration) and Table 6-5 (sequence length). Analysis using the modified t-test (Crawford & Howell, 1998) on Duration during acquisition (collapsed across blocks) indicated that none of the patients with AOS exhibited longer RTs or greater RT differences than the control group (all  $t_s < 1$ ,  $p_s > .25$ , one-tailed,  $df=12$ ) (see Table 3-4). While all patients had larger numerical differences between 1L and 1S than controls, this difference approached significance only for AOS1 ( $t = 2.103$ ,  $p = .057$ ,  $df=12$ ). The patient with aphasia also had RTs within the normal range ( $t_s < 1$ ,  $p_s$

> .15); however, this patient demonstrated a greater short-long difference during acquisition than the control group ( $ts > 2$ ,  $ps < .05$ , two-tailed,  $df=12$ ). A similar pattern was found for the Sequence Length analysis during acquisition (see Table 3-6): none of the patients with AOS or aphasia showed any evidence for longer RTs relative to the control group (all  $ts < 1.25$ ,  $ps > .10$ ). One patient (AOS2) did have a greater sequence length effect on RT than controls during acquisition ( $t = 2.447$ ,  $p < .05$ , two-tailed,  $df=12$ ). The patient with aphasia did not show a greater sequence length effect than controls.

**Table 6-5.** RT comparisons for Sequence Length (collapsed across blocks) for each patient.

Subject	Number	Acquisition	t-value	p-value	Retention	t-value	p-value
AOS1	4	349 (64)	- 0.499	0.314	307 (58)	- 0.897	0.194
	1	327 (83)	- 0.337	0.371	299 (63)	- 0.642	0.266
	$\Delta$	22	0.319	0.756	8	0.483	0.638
AOS2	4	506 (169)	1.218	0.112	436 (101)	0.830	0.211
	1	350 (85)	0.015	0.494	428 (120)	0.576	0.288
	$\Delta$	156¶	2.447	0.031	8	0.482	0.639
AOS3	4	383 (113)	- 0.113	0.456	425 (123)	0.683	0.254
	1	352 (89)	0.046	0.482	410 (90)	0.406	0.346
	$\Delta$	31	0.313	0.760	15	0.525	0.609
AOS4	4	342 (39)	- 0.578	0.287	316 (29)	- 0.776	0.226
	1	315 (36)	- 0.520	0.306	294 (19)	- 0.690	0.252
	$\Delta$	27	0.114	0.911	22	0.165	0.872
APH2	4	323 (76)	- 0.794	0.221	297 (63)	- 1.031	0.162
	1	306 (41)	- 0.658	0.262	282 (45)	- 0.803	0.219
	$\Delta$	17	0.267	0.794	15	0.432	0.673
CON (N=13)	4	393 (85)			374 (72)		
	1	349 (63)			367 (102)		
	$\Delta$	44			7		

¶  $p < .05$ ,  $df=12$ , two-tailed (Crawford & Garthwaite, 2005).

#### 6.2.4. Execution: Duration, and Absolute and Relative Timing Error

Timing error measures were based on the press and interpress durations that were recorded by the MEL (Psychology Software Tools, Inc.).

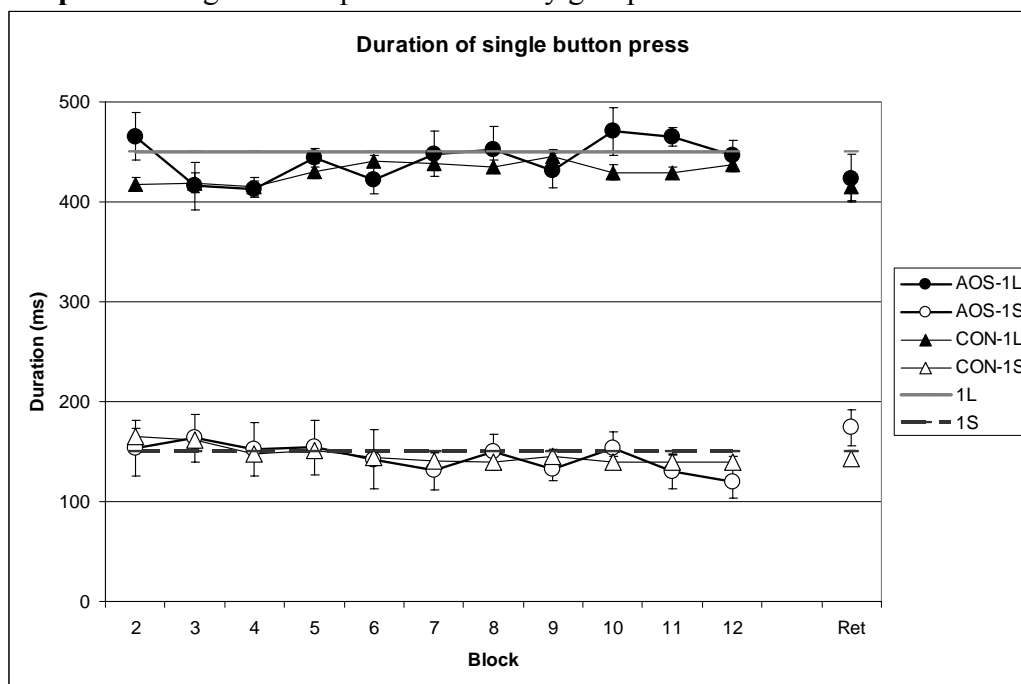
Single Press Duration. As a first step, duration of single button presses was examined to determine whether participants differentiated these presses, despite the absence of an effect on ST (or RT). An ANOVA on the acquisition data from the control group produced a significant effect of Duration ( $F[1,12] = 2119.56, p < .0001$ ), with averages for both long presses (431,  $SD = 27$ ) and short presses (147,  $SD = 27$ ) very close to their target durations of 450 and 150 ms, respectively. In addition, the Duration x Block interaction was significant ( $F[10,120] = 4.50, p < .0001$ ), indicating that the difference between the short and long presses became larger over the course of the acquisition session. Retention analysis revealed a significant effect of Duration ( $F[1,12] = 860.74, p < .0001$ ) and a marginal Duration x Phase interaction ( $F[1,12] = 3.80, p < .0749$ ), suggesting a decrease in duration from the last acquisition block to the retention block for the long press.

Duration data for the single presses for the controls and the AOS group are shown in Graph 6-5; comparisons of individual patients with the control group are presented in Table 6-6. As can be seen in Table 6-6, these analyses indicated that all four patients with AOS differed in single press duration during acquisition for at least one of the responses ( $ps < .05, df=12, one-tailed$ ), and that two of the patients showed a greater difference between short and long presses than controls ( $ps < .05, df=12, two-tailed$ ). At retention testing, only AOS3 demonstrated a difference relative to



controls. In contrast, the patient with aphasia did not differ significantly from controls during acquisition, although during retention his long presses were shorter than those of the control group ( $p < .05$ ,  $df=12$ , one-tailed). In short, these analyses suggest that the patients also differentiated between the short and long presses in terms of their execution, in some cases more so than the controls.

**Graph 6-5.** Single button press duration by group and block.



Absolute Timing. Results for absolute timing error (E) of single syllables are presented in Table 6-7; results for absolute timing error of sequences are presented in Table 6-8. For single presses, analysis of the acquisition data from the controls revealed a significant effect of Duration ( $F[1,12] = 16.78$ ,  $p < .005$ ), indicating greater absolute timing error for long presses than for short presses, and a significant effect of

Block ( $F[10,120] = 1.96, p < .05$ ), indicating a decrease in absolute timing error across blocks. Retention analysis produced only a significant effect of Duration ( $F[1,12] = 9.78, p < .01$ ), with greater absolute timing error for long presses than for short presses.

**Table 6-6.** Single press duration for each patient (collapsed across blocks), compared to controls (N=13).

Subject	Duration	Acquisition	t-value	p-value	Retention	t-value	p-value
AOS1	1L	408 (40)*	-1.847	0.04477	364 (100)	-1.003	0.16784
	1S	163 (34)	0.811	0.21645	176 (76)	1.178	0.13086
		245^^	1.855	0.08835	188	1.661	0.12264
AOS2	1L	465 (44)*	2.730	0.00913	459 (26)	0.865	0.20192
	1S	101 (21)*	-2.333	0.01893	135 (60)	-0.286	0.39006
		364¶	3.505	0.00434	324	0.878	0.39693
AOS3	1L	439 (50)	0.642	0.26634	410 (93)	-0.098	0.46165
	1S	190 (29)*	2.181	0.02491	222 (16)*	2.819	0.00774
		249	1.076	0.30326	188¶	2.217	0.04673
AOS4	1L	461 (56)*	2.409	0.01648	463 (104)	0.944	0.18190
	1S	121 (38)	-1.319	0.10595	165 (59)	0.785	0.22379
		340¶	2.593	0.02352	298	0.121	0.90545
APH2	1L	419 (48)	-0.964	0.17713	306 (11)*	-2.144	0.02663
	1S	116 (30)^	-1.572	0.07094	141 (42)	-0.071	0.47214
		303	0.426	0.66673	165	1.579	0.14040
CON (N=13)	1L	431 (12)			415 (49)		
	1S	147 (19)			143 (27)		
		284			272		

\*  $p < .05$ ,  $df=12$ , one-tailed (Crawford & Howell, 1998).

^  $p < .10$ ,  $df=12$ , one-tailed (Crawford & Howell, 1998).

¶  $p < .05$ ,  $df=12$ , two-tailed (Crawford & Garthwaite, 2005).

^^  $p < .10$ ,  $df=12$ , two-tailed (Crawford & Garthwaite, 2005).

Acquisition analysis of absolute timing error for sequences for the controls showed only a significant effect of Block ( $F[10,120] = 2.61, p < .01$ ), indicating greater error in block 2 than in blocks 6 through 11. There was no difference in

absolute timing error between 4S and 4L sequences, nor did sequence type interact with Block. However, in the retention analysis there was a significant effect of sequence type on absolute timing error ( $F[1,12] = 13.66, p < .005$ ), which indicated greater error for 4S than for 4L sequences. The other effects were not significant.

**Table 6-7.** Absolute timing error (E) for single presses for each patient (collapsed across blocks), compared to controls.

Subject	Duration	Acquisition	t-value	p-value	Retention	t-value	p-value
AOS1	1L	57 (19)	0.930	0.185	132*	1.791	0.049
	1S	36 (17)	0.182	0.429	80*	2.972	0.006
			<i>0.466</i>	<i>0.650</i>		<i>0.749</i>	<i>0.468</i>
AOS2	1L	48 (14)	- 0.107	0.458	27	- 0.898	0.193
	1S	54 (8)*	2.326	0.019	62*	1.779	0.050
			<i>1.514</i>	<i>0.156</i>		<i>1.696</i>	<i>0.116</i>
AOS3	1L	55 (12)	0.699	0.249	101	0.997	0.169
	1S	45 (23)	1.254	0.117	67*	2.110	0.028
			<i>0.346</i>	<i>0.736</i>		<i>0.706</i>	<i>0.493</i>
AOS4	1L	57 (22)	0.930	0.185	105	1.100	0.147
	1S	49 (8)^	1.730	0.055	61^	1.712	0.056
			<i>0.499</i>	<i>0.627</i>		<i>0.389</i>	<i>0.704</i>
APH2	1L	59 (16)	1.160	0.134	145*	2.124	0.028
	1S	45 (14)	1.254	0.117	43	0.520	0.306
			<i>0.058</i>	<i>0.954</i>		<i>1.018</i>	<i>0.329</i>
CON (N=13)	1L	49 (8)			62 (38)		
	1S	34 (8)			35 (15)		

\*  $p < .05, df=12$ , one-tailed (Crawford & Howell, 1998).

^  $p < .10, df=12$ , one-tailed (Crawford & Howell, 1998).

Analysis of absolute timing error using the modified t-test indicated that for single presses, only AOS2 differed from the controls for the 1S response during acquisition and retention ( $t_s > 1.70, p_s < .05, df=12$ , one-tailed); at retention testing,

AOS1 and AOS3 also differed from controls in terms of absolute timing accuracy (see Table 6-7). For sequences, none of the patients with AOS differed from the control group during either acquisition or retention (all  $t_s < 1.10$ ,  $p_s > .05$ ;  $df=12$ , one-tailed) (see Table 6-8). The patient with aphasia did not differ from controls during acquisition, but he did show greater absolute timing error for the 1L response and for sequences at retention testing ( $t_s > 2$ ,  $p_s < .05$ ). During acquisition, the larger absolute timing error for sequences for APH failed to reach significance ( $t= 1.589$ ,  $p= .069$ ).

**Table 6-8.** Relative timing error (AE-prop) and absolute timing error (E) for sequences (collapsed across 4S and 4L, and across blocks).

Subject	Measure	Acquisition	t-value	p-value	Retention	t-value	p-value
AOS1	AE-prop	.449 (.060)*	3.702	0.002	.556*	4.583	0.000
	E	132 (36)	0.483	0.319	204	0.820	0.214
			2.636¶	0.022		3.577¶	0.004
AOS2	AE-prop	.506 (.047)*	4.730	0.000	.645*	5.783	0.000
	E	139 (23)	0.667	0.259	224	1.053	0.157
			3.308¶	0.006		4.444¶	0.001
AOS3	AE-prop	.450 (.033)*	3.716	0.001	.449*	3.127	0.004
	E	149 (36)	0.930	0.185	199	0.762	0.230
			2.286¶	0.041		2.278¶	0.042
AOS4	AE-prop	.123 (.033)*	- 2.264	0.021	.150	- 0.925	0.187
	E	123 (23)	0.245	0.405	169	0.414	0.343
			2.062^^	0.062		1.298	0.219
APH2	AE-prop	.383 (.080)*	2.489	0.014	.399*	2.450	0.015
	E	174 (51)^	1.589	0.069	347*	2.480	0.014
			0.743	0.472		0.029	0.978
CON (N=13)	AE-prop	.247 (.053)			.219 (.071)		
	E	114 (37)			133 (83)		

\*  $p < .05$ ,  $df=12$ , one-tailed (Crawford & Howell, 1998).

^  $p < .10$ ,  $df=12$ , one-tailed (Crawford & Howell, 1998).

¶  $p < .05$ ,  $df=12$ , two-tailed (Crawford & Garthwaite, 2005).

^^  $p < .10$ ,  $df=12$ , two-tailed (Crawford & Garthwaite, 2005).

Relative Timing. Results of individual patient comparisons for relative timing error (AE-prop) of sequences are included in Table 6-8. For the controls, the ANOVA on the acquisition data revealed a significant effect of Sequence Type ( $F[1,12] = 10.37, p < .01$ ), indicating greater relative timing error for the 4S sequences than for the 4L sequences, and a significant effect of Block ( $F[10,120] = 5.53, p < .0001$ ), indicating that relative timing error was greater in block 2 than in blocks 5-12. The interaction was not significant. Retention analysis indicated a significant effect of Sequence Type ( $F[1,12] = 9.68, p < .01$ ), with greater relative timing error for 4S than for 4L.

All four patients with AOS as well as the patient with aphasia differed significantly from the control group (all  $t_s > |2|, p < .05$ ; see Table 6-8), with all but one patient (AOS4) showing greater error than controls. Surprisingly, AOS4 had significantly smaller relative timing error than controls ( $t = -2.264, p < .021$ ). To further examine the discrepancy between absolute and relative timing error observed in the separate modified t-tests, formal testing of the differences was conducted using the RSDT (Crawford & Garthwaite, 2005). Three of the four patients with AOS demonstrated a significantly greater difference between absolute and relative timing error than the control group ( $t_s > 2, p_s < .05$ ; see Table 6-8), caused mainly by the larger relative timing error. The difference for AOS4 (who had smaller relative timing error than controls) just failed to reach significance ( $t = 2.062, p = .062$ ). In contrast, absolute and relative timing error did not dissociate for the patient with aphasia ( $t =$

.058,  $p > .10$ ), who showed greater relative and absolute timing error than controls, with a proportionate increase for both.

### **6.3. Discussion**

#### **6.3.1. Motor Programming in Younger Control Participants**

Before turning to the patient data, the results from the control participants will be discussed in relation to previous studies and to Experiment 1. Importantly, several key findings from previous studies were replicated in the present experiment. First, a sequence length effect was obtained for RT, consistent with other studies that have used these responses in the self-selection paradigm (e.g., Immink & Wright, 2001; Wright et al., 2004). This sequence length effect suggests that the button presses in a sequence were programmed as separate units, thus increasing SEQ load relative to single presses.

These RT findings differ from those in Experiment 1, in which no sequence length effect was found, suggesting that for speech, sequences of repeated gestures were programmed as single units. This difference between the programming of finger and speech movements is likely due to the extensive practice with speech over the course of a lifetime, making speech a highly overlearned motor skill with rapid and fluent execution as one of its key characteristics. In contrast, the finger movement responses in Experiment 2 represented a relatively novel motor task, making it less likely that pre-existing programs or routines are available to facilitate integration of button presses into a single unit. Interestingly, there is evidence to support the idea

that extensive practice results in formation and consolidation of multiple button presses into a single chunk (e.g., Klapp, 1995; Wright et al., 2004). For example, Wright et al. demonstrated that extensive random practice (but not blocked practice) results in long-term changes in programming these sequences, i.e. by recoding the sequence from the initial multiple-unit representation into a single multi-element unit.

Second, a sequence length effect was also present on ST, which could be due either to loading all unique units (Immink & Wright, 2001) or to loading all units (Magnuson et al., in press). Recent work by Magnuson et al. (in press) suggested that for button press responses, all units that make up a sequence are preprogrammed, although the sequence length effect on RT indicates that this preprogramming does not lead to the formation of a single unit. The ST effect for speech in the context of equal RT for single vs. multiple syllables may reflect the greater INT demand for preprogramming all syllables as separate units plus an additional integration process that creates a single unit, or it could reflect the greater INT demand for preprogramming a single unit of greater complexity.

In contrast to the effects of sequence length, there was no effect of single press duration on ST or RT, as in Experiment 1 but unlike previous studies using these finger movements sequences (e.g., Immink & Wright, 2001). Again, this lack of effect could not be attributed to a failure on the participants' part to differentiate between the short and long button presses. As discussed above, while Wright and colleagues (2004) also failed to obtain a significant duration effect, these authors did observe the effect in a reanalysis involving only the first 30 trials. Similar reanalysis in the present

study failed to reveal a duration effect; however, since we excluded the first block as a warm-up block, it is possible that the duration effect had dissipated by the second block (which was the first block analyzed here).

An alternative explanation discussed for Experiment 1, i.e. the feedback hypothesis which states that people fully program responses under conditions in which they receive detailed feedback, does not apply in Experiment 2 since feedback was provided here similar to previous studies. Another possibility is that some form of response conflict caused interference for either the INT process or the press of the space bar to end the ST interval (the responses as well as the end of the ST interval were performed by the left hand). This hypothesis accounts for the absence of a duration effect in Experiment 2, but does not apply to Experiment 1, where different modalities were used for the ST and RT intervals.

Finally, perhaps a more plausible alternative explanation is that the longer RT window used in the present study is responsible for this failure to replicate the duration effect. This hypothesis would account for both Experiment 1 and 2. Future studies could test this hypothesis by varying the RT window.

### **6.3.2. Motor Programming in Apraxia of Speech**

Experiment 2 was designed to address two primary questions: 1) whether AOS involves a central (i.e. not speech-specific) deficit, and 2) like Experiment 1 above, whether different motor programming processes dissociate for finger movements in AOS. The results are consistent with the hypothesis that AOS represents a central



(non-modular) deficit in the INT process, since the prediction of longer ST relative to controls was confirmed for all four patients with AOS in this nonspeech movement task. The prediction of a disproportionate complexity effect on ST for patients with AOS could not be assessed since there was no effect of single press duration on ST for the controls.

With respect to accuracy of timing, absolute timing error for single presses did not differ between patients with AOS and controls, which suggests that the longer ST for single presses in the patients with AOS cannot be attributed to a speed-accuracy trade-off (i.e. longer STs did not result in greater accuracy). Moreover, the longer STs for sequences in two of the patients with AOS (AOS2 and AOS3) were associated with increased relative timing error. Although there were few differences in absolute timing error for single button presses during acquisition, three of the patients with AOS showed greater relative timing error than controls for the sequences, further supporting the interpretation that AOS involves a deficit in INT, since similar (or reduced) accuracy was achieved with considerably longer preparation time.

To assess the integrity of the SEQ process, it is important to consider how a potential deficit in SEQ might be reflected in the data. There are two types of findings that could point to a deficit in SEQ. One of these would be a finding of overall increased RT. It should be noted however that longer RT by itself would also be consistent with problems in other processes occurring after the presentation of the go-signal, such as initiation (sending a burst of activation to the effector system) or unpacking a subprogram (Sternberg et al., 1978), rather than the search and retrieval

operations that characterize the SEQ process. The second type of finding that would suggest an SEQ deficit is a disproportionate sequence length effect on RT, in the context of a sequence length effect for controls.<sup>13</sup> In this case, initiation or unpacking explanations would be less plausible; rather, such a pattern would suggest that additional units in the buffer disproportionately tax the search and retrieval process.

Note that in Experiment 1, only the first type of finding could be used to assess the SEQ process, since the controls did not show a sequence length effect on RT. However, in Experiment 2 there was a sequence length effect on RT for the controls, allowing a more powerful test of the integrity of the SEQ process in AOS. In the present experiment, none of the patients demonstrated longer RT relative to controls (as was the case in Experiment 1), and three of the four patients showed a sequence length effect on RT of the same magnitude as the controls, supporting the hypothesis that the SEQ process was not impaired in these individuals. However, AOS2 did show a disproportionate sequence length effect on RT, suggesting that in this patient, the SEQ process may also be affected.

The dissociation between longer ST and normal RT for these patients with AOS cannot be explained on the basis of a generalized reduction in processing speed; instead, the slowing was confined to the preprogramming stage. This pattern of results is consistent with, and was predicted by, the hypothesis of a localized deficit in INT

---

<sup>13</sup> Note that a sequence length effect on RT for patients when controls do not show a sequence length effect would be more suggestive of a deficit in INT (e.g., Deger & Ziegler, 2002), with controls being able to integrate multiple elements of a sequence unlike patients, who would then rely on the (intact?) SEQ process to program the sequence. Thus, the entire pattern of results must be considered in the interpretation.

with intact movement initiation and SEQ (although there is some evidence to suggest an additional SEQ deficit in AOS2).

In contrast, the patient with aphasia did not show longer ST (or RT) than controls, suggesting that brain damage per se or the presence of aphasia is not responsible for the ST effect in the patients with AOS. However, the findings from this patient with aphasia should be interpreted with caution, for various reasons. First, the findings from a single individual may not generalize to a larger sample of patients with aphasia. Unfortunately, additional patients with aphasia were not available.

Second, APH2 was left-handed, and since the responses were carried out using the left hand for all participants, it is possible that both producing the responses and depressing the space bar to end the ST interval were facilitated due to greater proficiency using the left hand. Consistent with this view is the fact that AOS1, who was also left-handed, showed a smaller ST difference with controls than the other patients with AOS. However, AOS1 did show significantly longer ST than controls during retention testing, unlike APH2.

Third, APH2 produced a disproportionate effect of duration on RT (as did the other left-handed patient, AOS1). Within the INT/SEQ model, this is unexpected, since presumably the number of units is the same for a single button press regardless of duration, and SEQ is not thought to be sensitive to unit complexity. One possibility is that unit complexity affects the unpacking process as envisioned by Sternberg et al. (1978), and that unpacking a long button press takes longer than unpacking a short button press. However, this interpretation is undermined by the absence of duration

effects on RT for controls. An alternative view is that the INT process occurred during the RT interval in this patient. Since no overall increase in ST was observed for this patient (unlike AOS1), the reason for such INT programming during RT is likely not one of a preprogramming deficit. Instead, it may be that APH1 was unable to maintain a response in the buffer, necessitating the reprogramming of these responses after the go-signal. Consistent with this interpretation is the fact that APH2 had the fastest mean RTs among the patients, which might reflect a strategy of producing the response as soon as possible after programming it since it cannot be maintained in the buffer. The absence of a disproportionate duration effect for this patient in the speech experiment might then imply that such a putative buffer maintenance problem in APH1 affects a different memory system than that used in speech production.

However, given the potentially different measurement errors and thus sensitivities of button press duration and acoustic measures, and the fact that the speech experiment was conducted approximately one year after the finger experiment, the interpretation of the disproportionate duration effect on RT must remain purely speculative.

With respect to the domain-specificity of the deficit in AOS, the results of this experiment add to the growing body of evidence that individuals with AOS also show impairments on nonspeech motor control tasks (e.g., Ballard & Robin, in press; Clark & Robin, 1998; Hageman et al., 1994; McNeil et al., 1995). However, this study is the first to use the exact same paradigm with speech and non-speech movements involving the exact same temporal patterns. The central deficit in INT is considered to be present because of the longer ST for the finger movements, as well as by the greater

timing error for finger movement sequences and for speech. INT programming time for single finger presses was longer for three of the four patients with AOS but not for the patient with APH. It is possible that the longer ST for speech and longer ST for finger movements result from entirely separate impairments, and thus that the finger findings do not need to be explained in reference to a model of speech motor programming. However, it is interesting to note that the three patients with AOS who showed the strongest effects in this finger experiment were the same three that showed longer ST than controls in the speech experiment, suggesting that the deficit in these patients affected both speech and manual modalities.

This apparent effector-independent nature of the deficit in AOS does not follow from the view that speech is subserved by a separate, specialized neuromotor control system (e.g., Ziegler, 2003a,b), since there is no reason to expect deficits in the control or programming of limb movements. However, according to the view that speech is subserved by a neuromotor control system that also controls other motor skills that share properties with speech (e.g., Folkins, 1985; Ballard et al., 2003), deficits in limb motor control or programming should be evident when using tasks similar to speech.

Matching speech and nonspeech tasks on relevant features is not a trivial task (e.g., Shaiman et al., 2006), and has been argued to result in a problem of infinite regression (Weismer, 2006). In the present set of experiments, similarity was defined in terms of the temporal patterns of speech and nonspeech movements, and the findings supported the notion that these movement patterns were governed by a shared

motor control system. One could of course argue that the speech movements in Experiment 1 were in fact not speech at all, and therefore that the similarity between the findings of Experiment 1 and Experiment 2 arose from a shared *nonspeech* volitional motor control system. While this argument forces the requirement of a clear definition of speech as separate from speech-like movements, it may reduce the definition to one in terms of the properties that speech exhibits, essentially reducing the task-dependent model (Ziegler, 2003a,b) to the integrative model that seeks to understand speech in terms of its properties (Ballard et al., 2003).

The present findings contribute to our understanding of the nature of speech motor programming, and AOS, either by uncovering similarities between speech and nonspeech motor control or by forcing a refinement of the definition of speech. Further research is needed to resolve this issue; a promising, potentially powerful approach is to examine patterns of transfer across speech and nonspeech tasks (Weismer, 2006), and preliminary studies suggest that transfer from nonspeech to speech motor control does indeed occur (e.g., Shaiman et al., 2006).

On the assumption that the responses in Experiment 1 did engage the speech motor control system, the similarity of the findings for our patients with AOS narrows the field of competing accounts of AOS discussed in the previous chapter, or at least necessitate further elaboration or refinement of such accounts, in addition to elaborations needed to account for the speech findings themselves (see Discussion of Experiment 1). For example, the reduced buffer capacity hypothesis (Rogers & Storkel, 1999) would need to add the assumption that the buffer whose capacity is

limited in AOS is a domain-general motor buffer, and thus that the capacity limitation must be captured in some way other than by reference to speech-specific constructs such as syllables. Similarly, a deficit in verbal working memory cannot account for the observed finger motor programming difficulties in AOS, since this task presumably does not involve rehearsal of verbal material. The dual route hypothesis (Varley & Whiteside, 2001) was argued to be similar to the hypothesis of an INT deficit, with one difference being that the dual route hypothesis in its current formulation is specific to speech production and thus does not account for the finger programming deficits observed in the present experiment. The same argument applies to an account in terms of phonological encoding, though not to the stimulus-response mapping hypothesis, which would also affect ST in this experiment. As before however, this account does not account for the observed errors in absolute and relative timing, nor for the observed sequence length effect on ST.

One other account that was discussed in Chapter 5 that may apply to the findings from this finger experiment as well is that the longer ST is due to a more conservative estimate of one's self-judgment of readiness to respond, and that such judgments interact with perceived task difficulty. This explanation is not implausible for speech, given that speakers with acquired communication disorders may be less confident in speaking than unimpaired speakers, due to past experiences of speech/language difficulties. And although finger movements may also have resulted in problems for these patients since their stroke, this experiment involved the index finger of the left (intact) hand. Moreover, given that these patients presumably had

greater experience with using the left hand since their stroke than the controls, this might have resulted in greater confidence, which would have predicted shorter STs in this experiment. However, it remains possible that self-judgments of readiness are not related to confidence level or past experiences. Finally, this account does not account for errors of absolute and relative timing.

An interesting additional observation in terms of the timing accuracy for these patients is that while the patient with aphasia also showed reduced temporal accuracy of finger movements, his pattern was qualitatively different than the patterns of the patients with AOS. Specifically, whereas the patient with aphasia showed proportionate decreases in accuracy of relative and absolute timing, all patients with AOS showed a dissociation between absolute and relative timing accuracy of the sequences. These dissociations in AOS provide neuropsychological support for the theoretical distinction between generalized motor programs (GMPs) and parameters (Schmidt, 1975, 2003; Schmidt & Lee, 2005), and further build the case of a motor programming deficit in AOS. In three of the patients, relative timing was more impaired than absolute timing, whereas in the fourth patient (AOS4), relative timing was in fact significantly better than in controls. While this may seem surprising, this patient commented during the instructions for Experiment 2 that this task was very much like Morse code, with which he had considerable experience.

The dissociation between relative timing and absolute timing accuracy is consistent with a study by Clark and Robin (1998), who also found that GMPs and parameters were differentially impaired in their patients with AOS using a nonspeech



oral motor control task. These authors noted that either the GMP or the parameters were impaired, and that these appeared to be in a trading relationship to each other, even within an individual (e.g. one patient showed larger GMP error on one occasion but larger parameter error on another occasion). In the present study, control of relative timing of finger movement sequences was selectively impaired in three of the patients with AOS, suggesting a deficit in the GMP rather than the parameters. Recall that in Experiment 1, two of the patients with AOS also showed a dissociation between absolute and relative timing error of sequences. However, in that case it was absolute timing that was more impaired. It is possible that these differences reflect differences between speech and nonspeech motor control, although Clark and Robin's (1998) suggestion of a trading relationship between GMP and parameter accuracy suggests that these differences do not necessarily stem from different motor control systems.

The fact that timing accuracy was reduced in the patients with AOS for both speech and nonspeech movements suggests an interpretation of the hypothesized INT deficit as one involving an impairment in the control of timing at an effector-independent level of processing. This view is consonant with the assumption in Schema Theory that the effector system that executes a movement is set by a parameter that is supplied to an abstract motor program that specifies the relative timing (and force) structure of a movement (Schmidt, 1975; Schmidt & Lee, 2005). In addition, there is evidence from motor learning studies that a sequence representation is independent from its effector system (e.g., Keele et al., 1995). If it is assumed that

the INT process includes activation and parameterization of GMPs, then an INT deficit might affect either of these subprocesses, independent of the effector system that will carry out the movement. Note that this does not mean that no effector-specific deficits can ever be found, since it is possible in this conceptualization of motor programming that there exist deficits in setting the appropriate effector parameters. Rather, the hypothesis offered here claims that in these patients with AOS, the deficit appears to be at the level of timing control, and that this deficit affects production of movements with both oral and manual effector systems.

This interpretation in turn suggests that the INT process involves a domain-general component that handles specification of temporal goals in an abstract, effector-independent manner (similar to the effector-independence of sequence information; Keele et al., 1995; Klapp, 2003). If INT deficits had only been found for one modality, this would have suggested effector-specific INT processes.

#### **6.4. Conclusions**

The present study provides evidence from a reaction time approach in support of the hypothesis that AOS represents an impairment of the INT process, a motor programming process that is responsible for organizing the internal spatiotemporal structure of a movement. Patients with AOS spent more time preprogramming movements, yet were generally less accurate in terms of timing than controls. In contrast, movement initiation and the SEQ process appeared to be intact. The fact that this INT deficit was not confined to speech motor programming (Experiment 1) but

was also evident in programming simple finger movements suggests that linguistic-based or speech-specific accounts of AOS are inadequate to capture the full extent of the deficit. Further specification of this hypothesized INT deficit was also offered. In particular, based on the fact that these experiments involved an explicit emphasis on timing accuracy and given that individuals with AOS were generally less accurate than control participants in terms of timing, an interpretation of the INT deficit in AOS as one of impaired programming of timing is plausible.

In turn, the observed dissociations between ST and RT in AOS provide additional, neuropsychological evidence to support the INT/SEQ model. Moreover, the similarity of the findings across speech and manual movements suggests that not only is the model relevant to understanding both speech and nonspeech motor programming, but also that the INT process involves shared components between these modalities, arguing against a strictly modular view of the organization of speech motor control.

### **Acknowledgements**

Chapter 6 is being prepared for publication in a manuscript entitled “Motor Programming in Apraxia of Speech”, by Edwin Maas, Donald A. Robin, David L. Wright, and Kirrie J. Ballard. The dissertation author is the primary investigator and author of the manuscript.

## **7.0. The Role of Rhythm in Speech Motor Programming**

The present study was designed to increase our understanding of speech motor programming, and more specifically, the role of rhythm in the programming of syllable sequences. Rhythm is defined here as “the distribution of different levels of stress across a series of syllables” (Kent, Kent, Weismer, & Duffy, 2000: 279). This factor has received relatively little attention in the study of speech production (Kent et al., 2000) yet forms an integral part of fluent speech, as well as many other serial motor skills (Sakai, Hikosaka, & Nakamura, 2004). Kent et al. suggested that rhythm may facilitate speech production by providing a temporal framework that allows the coordination of different sensory and motor aspects of an ongoing movement sequence. This suggestion indicates a potentially important role for rhythm during motor programming. The recent INT/SEQ model (Klapp, 1995, 2003) that is the focus of this dissertation has not yet delineated a clear role for rhythm during the programming of speech or finger movements, and it is possible that consideration of this factor may help resolve an outstanding issue regarding the nature of speech motor programming. The issue is why speakers sometimes appear to program a sequence of syllables as a single unit whereas in other situations they appear to program a sequence of syllables as separate units.

Recall that in Experiment 1, which used the self-selection paradigm (e.g., Immink & Wright, 2001), a sequence length effect was found on Study Time (ST) but not on Reaction Time (RT). That is, ST was longer for a sequence of four repeated syllables than for a single syllable, but no such difference was found for RT. Based on

the logic of the INT/SEQ model, this pattern of results suggests that the sequences and the single syllables did not differ in terms of the number of units maintained in the motor buffer, and thus that these sequences were preprogrammed as a single unit. Although this interpretation is consistent with evidence that speakers can preprogram multisyllabic sequences, be they words (e.g., Klapp, 1974; Klapp et al., 1973; Santiago et al., 2000) or nonwords (e.g., Klapp, 2003), it is at odds with previous studies that suggested that syllable sequences are not integrated when they consist of reiterations of the same syllable (nonwords: Deger & Ziegler, 2002; Klapp, 2003; words: Klapp et al., 1979; Sternberg et al., 1978). This discrepancy requires resolution, because it potentially undermines an important method of identifying units of speech motor programming, i.e. the examination of sequence length effects on simple RT (Klapp, 2003: 137).

The basic assumption of the model that SEQ (measured by simple RT) is not sensitive to the internal complexity of a unit can only be maintained by assuming different unit sizes for different sequences of the same number of syllables (e.g. /dadada/ is three units, /dabaga/ is one unit). However, to maintain the falsifiability of the model, an independent, principled account must be provided for why only sequences of different syllables become integrated into a single unit, while sequences of repeated syllables are programmed as separate units. Neither Klapp (2003), nor Deger and Ziegler (2002) provided a satisfactory explanation for this discrepancy.

Klapp (2003) suggested that sequences of different syllables have greater coarticulatory demands than do sequences of repeating syllables, and Deger and

Ziegler (2002) made the similar assumption that transitions between different syllables are more demanding than transitions between repeating syllables. The reasoning is that speakers choose to preprogram more demanding sequences, and that less demanding sequences (i.e. reiterations of the same syllable) are not preprogrammed. In other words, the motor programming demands associated with transitions between different syllables are shifted into preprogramming, where the INT process integrates the syllables into a single unit and as a result reduces the SEQ load. A related account is that syllables are integrated into a single unit only when this minimizes serial order errors. For non-repeating sequences, accidentally retrieving the incorrect unit from the buffer will result in a perceivable error of serial order, whereas the same error for a repeating syllable sequence would be indistinguishable from the target. All three accounts share the assumptions that 1) speakers will postpone programming until the SEQ stage *unless* there is good reason to preprogram (e.g., error avoidance, programming coarticulation), and 2) that integration depends on the composition of the syllable sequence (only different syllables are integrated). These three accounts will not be distinguished here and will be considered together as the Composition Hypothesis (CH).

An alternative hypothesis considered here, which we will call the Rhythm Hypothesis (RH), claims that syllable sequences are integrated only when there is a unifying rhythmic structure to the sequence that provides a framework for coordination of spatial and temporal aspects of speech movements (cf. Kent et al., 2000). Support for this hypothesis was provided by Sternberg and colleagues (1978)

who found that the slope of the simple RT function was affected by the number of stressed syllables, not by the total number of syllables. Sternberg et al. suggested that the stress group constituted the unit of speech motor programming; a stress group refers to a syllable with primary stress and associated unstressed syllables, and corresponds to what Levelt and colleagues refer to as a phonological word (Levelt et al., 1999).<sup>14</sup> A strong form of this Rhythm Hypothesis (RH) shifts the focus away from the composition of the syllable sequence to the rhythmic structure of the sequence, and claims that even a repeating-syllable sequence will be integrated if there is a unifying rhythmic structure. This would account for the findings from Experiment 1, where repeating-syllable sequences consisting of non-isochronous syllables produced a sequence length effect on ST (INT) but not on RT (SEQ). To account for the sequence length effect on simple RT for repeated syllables obtained by other investigators (e.g., Klapp, 2003), the RH must assume that such sequences represent separate stress groups. Since most words (at least in languages such as English and German) involve alternation between different syllables, it is reasonable to suppose that speakers in these studies considered the sequences of repeating syllables as repetitions of the same one-syllable item, which might induce a more list-like prosody with equal stress on each syllable. In Klapp (2003) and Deger and Ziegler (2002), the stress pattern of the target responses was not specified, unlike in our Experiment 1

---

<sup>14</sup> A phonological word is defined as the domain of syllabification (Levelt et al., 1999) and includes a stressed syllable and surrounding unstressed syllables. For example, the phrase “He saved it” would constitute a single phonological word (he.save.dit), whereas the phrase “He saved Igor” would constitute two phonological words (he.saved and i.gor) since there are two stressed syllables.

where the sequences had specific rhythmic structures that had to be learned and produced accurately.

The present study was designed to tease these hypotheses apart by examining repeating syllable sequences with and without a stress pattern (rhythmic vs. isochronous sequences) in the same experimental context. Given that the INT/SEQ model is based on results from both choice and simple RT paradigms, the present study includes both paradigms. Experiment 3 utilized the self-selection paradigm (e.g., Immink & Wright, 2001) as in Experiment 1, to determine whether syllable sequences become integrated as assessed by RT (SEQ). The self-selection paradigm also provides a measure of the INT process in its Study Time (ST) measure. However, given that some differences have been observed between ST and more traditional choice RT measures of the INT process (Klapp, 1995, 2003), especially with respect to the preprogramming of sequences, Experiment 4 involved a choice RT paradigm using the same responses. This way, the ST index of INT from Experiment 3 can be compared with the choice RT index of INT from Experiment 4 to provide cross-validation of both paradigms.

### **7.1. Experiment 3: INT and SEQ (Self-Selection Paradigm)**

The primary purpose of this experiment was to test the two hypotheses by assessing the SEQ process and the INT process within a single paradigm. The number of syllables (3 vs. 5 syllables) was factorially crossed with the temporal structure of the sequences (isochronous vs. rhythmic), resulting in four conditions; only repeating



syllables were used to examine the effects of rhythm separate from sequence composition. The two hypotheses make different predictions regarding INT and SEQ for these sequences.

According to the CH, only alternating syllables are integrated, and thus the sequences in this experiment should be programmed as separate units since they involve repeated syllables. If only the first unit (syllable) is preprogrammed (Klapp, 1995, 2003) or all unique units (Immink & Wright, 2001), then there should be no sequence length effect for INT, nor should there be effects involving rhythm when the first syllable is kept constant across conditions. However, if the self-selection paradigm's ST allows preprogramming of all units of a sequence (as was observed for finger movements; Magnuson et al., in press), then it is possible that a sequence length effect will emerge on ST. However, such a sequence length effect should be independent of the rhythmic structure of the sequence. Importantly, the CH predicts a sequence length effect on RT for both isochronous and rhythmic sequences; there should be no differences between isochronous and rhythmic sequences.

According to the RH, integration is facilitated by the presence of a rhythmic structure but does not occur for isochronous sequences, which are programmed as a series of separate units. On this view, the INT process should take longer for rhythmic sequences than for isochronous sequences if preprogramming occurs only for the first unit or all unique units, since the first unit for rhythmic sequences would be the entire sequence whereas for isochronous sequences the first unit would be the first syllable. Moreover, if a long rhythmic sequence is more complex than a short rhythmic

sequence, then there should be a sequence length effect on INT for these sequences, but not for the isochronous sequences. However, if the self-selection paradigm allows preprogramming of all units during the ST interval (Magnuson et al., in press), then it is possible that both rhythmic and isochronous sequences will produce a sequence length effect on ST. To differentiate between preprogramming of multiple syllables with integration and preprogramming of multiple syllables without integration, the pattern of RT will be critical: The RH predicts that a sequence length effect will only be present for isochronous sequences, not for rhythmic sequences. Furthermore, the RH predicts that rhythmic sequences should produce faster RTs than isochronous sequences, since rhythmic sequences constitute only a single unit whereas isochronous sequences consist of multiple units. Moreover, a non-trivial and counter-intuitive prediction following from this is that RT for the long rhythmic sequences should be faster than the short isochronous sequences.

Finally, this experiment also investigated whether any effects would be affected by presentation format. There is some evidence to suggest that the composition of the set of target responses affects motor programming, both for speech (e.g., Meyer et al., 2003) and for finger movements (e.g., Immink & Wright, 2001; Wright et al., 2004). For instance, Wright et al., using the self-selection paradigm involving the same finger movements as in our Experiment 2, observed a sequence length effect on RT and ST during acquisition only for participants in random practice conditions (in which multiple responses were practiced in each session, in random order), but not for participants in blocked practice conditions (in which each session

involved only a single response type). Wright et al. suggested that blocked-practice participants were able to quickly form an integrated representation of the sequences during acquisition, unlike random-practice participants (although at retention testing, the pattern was reversed, suggesting that only random practice resulted in formation of integrated representations of the sequences).

With respect to speech, Meyer et al. (2003), using a choice RT task that involved naming pictures of monosyllabic and disyllabic words, only observed a sequence length effect in pure blocks (blocks in which words were grouped by number of syllables), but not in mixed blocks (involving both monosyllabic and disyllabic words). Meyer et al. argued that speakers adapt their criterion to begin articulation as a function of the grouping of responses. That is, perhaps speakers chose to fully program the entire response in pure blocks, whereas in mixed blocks they chose to initiate speech as soon as the first syllable program was retrieved. Although Klapp (2003) did observe the sequence length effect on choice RT using random (mixed) presentation for nonwords consisting of alternating syllables, this was not the case for repeating-syllable sequences. Based on the findings by Meyer et al., it is possible that a sequence length effect would be observed for such sequences when they are presented grouped by sequence length.

It is important to point out that although the notions of random vs. blocked presentation are not equivalent to the notions of pure and mixed blocks, blocked practice does entail a pure condition. Thus, the finding that blocked practice in finger movements produced no sequence length effects (Wright et al., 2004), whereas pure

conditions were the only ones to produce a sequence length effect in speech production (Meyer et al., 2003) led us to include both a mixed and a pure condition in this study to determine in which condition, if any, a sequence length would emerge when using the self-selection paradigm applied to speech.

## **7.1.2. Methods**

### **7.1.2.1. Participants**

Participants were eleven individuals who participated for course credit. Three were excluded due to bilingualism or neurological history. The remaining eight participants (all female) were all monolingual speakers of English, were right-handed, and had normal or corrected-to-normal vision and hearing. Mean age was 19.5 years (range: 18-21). All participants read and signed an informed consent form in accordance with local IRB guidelines.

### **7.1.2.2. Materials and Design**

Materials consisted of syllable sequences forming a factorial set crossing Rhythm (rhythmic vs. isochronous sequences) and Number (3 vs. 5 syllables)<sup>15</sup>, resulting in four conditions. Target patterns were demonstrated using auditory response models (see below), and were created by varying syllable duration to yield the four temporal patterns presented in Table 7-1.

---

<sup>15</sup> Sequences of 3 and 5 syllables were chosen (rather than 2 and 4) in order to allow control for both total response duration and duration of the initial element, which is not possible with even-numbered sequences.

**Table 7-1.** Overview of materials for Experiment 3 and 4.

Length	Rhythm	Syllable	Response	Cue	Pattern	Duration
3	No	/da/ /gi/	/dadada/ /gigigi/	+3+ #3#	200-200-200	600
	Yes	/da/ /gi/	/dadada: /gigigi:/	@3@ ^3^	200-125-275	600
5	No	/da/ /gi/	/dadadadada/ /gigigigigi/	+5+ #5#	200-200-200-200-200	1000
	Yes	/da/ /gi/	/dadada:dada/ /gigigi:gigi/	@5@ ^5^	200-125-275-200-200	1000

In order to avoid anticipation of the first syllable (which may abolish motor programming effects due to the possibility of maintaining the first syllable in the buffer), each condition and each block contained sequences of the syllable /da/ and sequences of the syllable /gi/, resulting in a total of eight different responses throughout the experiment. Since the influence of rhythm on the integration of repeating syllables was important, the same phonemes and syllables were used in the rhythmic and isochronous conditions.

An auditory model was created for each response, as follows. A speaker with a native-like accent (EM) produced each of the two syllables in a carrier phrase (“It’s a \_\_\_\_ a day”) with syllable duration of approximately 1500 ms, recorded at 22 kHz. The target syllables were then spliced from the carrier phrase using Adobe Audition v1.5 software, and used to create the syllables for the sequences. Each syllable consisted of a 50-msec voiced stop-gap, and duration of the vowel was varied to create syllables of different durations. Durations were chosen that would result in responses

within the normal speaking rate range of approximately 5 syllables per second (e.g., Yorkston, Beukelman, Strand, & Bell, 1999). Syllables were then edited together to form the rhythmic and isochronous sequences in Table 7-1. Rhythmic structure was defined in terms of duration of the same syllable token to ensure that there were no unintended differences in the vowel or consonant; this way, the sequences truly consisted of repeated syllables. Furthermore, one goal of the present study was to further explore the lack of a sequence length effect on RT in Experiment 1, where only duration was varied as well.

Finally, each response was paired with a visual symbol (cue) that served as the imperative signal. The cues consisted of a number (3 or 5) with a non-alphabetic character on either side; throughout the experiment, Courier New font (20 pts.) was used to ensure equal size of the cues (see Table 1). The symbol-naming task has been used in previous studies on speech production (e.g., Levelt & Wheeldon, 1994) and was used here to avoid the inherent confounds with visual stimulus processing and reading when using orthography to cue responses of different sequence length.

Each of the four conditions above was tested in pure and mixed conditions; all factors were tested within participants. In the pure condition, responses were grouped by sequence length (3 or 5 syllables), whereas in the mixed condition there were two 3-syllable sequences and two 5-syllable sequences. There were two pure phases and two mixed phases, each with a total of four responses (two /da/ and two /gi/). Pure and mixed phases alternated, and the order of phases was counterbalanced across participants so that every possible order occurred once across all participants.

### 7.1.2.3. Task and Procedures

The experimental setup involved a desktop computer, keyboard and monitor, a button box (SRBox, Psychology Software Tools, Inc.), a DAT recorder (Sony model TCM-M1), two omnidirectional lapel microphones (one for the voice key in the button box and one for the DAT recorder), and a padded adjustable headset that was used to keep the microphones at a constant mouth-to-microphone distance of approximately 5 cm. The experiment was controlled by E-Prime software (version 1.1; Psychology Software Tools, Inc.).

Each participant was tested individually. After reading and signing the consent form and filling out a questionnaire, the general nature of the experiment was explained by the experimenter. Participants were told that the experiment focused on how well people can produce novel speech utterances, and that they would have to learn associations between visual symbols and the responses. The words “word” and “syllable” were carefully avoided during the instructions; instead the words “utterance” and “response” were used. Participants were also informed that their responses would be recorded onto digital audio tape. Once the headset was in place, participants were told to press the space bar to retrieve written instructions on the screen; they were encouraged to ask questions if the written instructions were unclear.

The experiment consisted of four phases, in each of which participants practiced a set of four responses. The number of response options was limited to four in all phases, so that each trial involved a choice between four different responses (2 /da/ and 2 /gi/, 2 isochronous and 2 rhythmic sequences). At the beginning of each

phase the visual cues were presented along with orthographic representations of the associated responses (e.g., @3@: dadadaaa). Once participants had studied the cues and utterances, each cue was presented four times along with the auditory response model. This procedure was included to familiarize participants both with the specific temporal pattern of the target responses as well as with the cue-response pairings.

Following the presentation of the models, there was a matching test that was included to provide an additional opportunity to establish the cue-response pairing. After a warning signal (“Listen”) that was presented for 500 ms, a blank screen followed for 500 ms, after which one of the four auditory targets was presented, followed by one of the visual cues presented in the center of the screen with a fixed stimulus onset asynchrony (SOA) of 1500 ms. Participants had to decide as quickly and as accurately as possible whether the cue matched the auditory target by pressing the “v” (yes) or “n” (no) on the keyboard with the index and middle fingers of their left hand. Order of presentation was random, and feedback was provided immediately and consisted either of the message “Good!” in teal-colored font or “Wrong” in red. Each cue appeared in both matching and mismatching conditions; mismatches were created by pairing the cue with the incorrect stress pattern (same phonemes) for the pure conditions and with the wrong phonemes (same number of syllables) for the mixed conditions. Incorrect trials were rerun until each of the eight conditions (four responses, match and mismatch) had been answered correctly twice, or until a total of 64 trials had been presented.



After the matching task, instructions appeared on the screen about the self-selection paradigm (the primary task of interest). Participants were informed of the trial events and procedures, and were told that there would be three blocks of trials. They were also informed that the duration of the blocks and the entire experimental session depended on their performance (speed and accuracy), because incorrect trials would be rerun at the end of each block until they reached criterion (see below).

In the self-selection paradigm, participants first saw the word “ready” in white letters (Courier New, 20 pts) in the center of a black screen, for 1000 ms. After a blank black screen for 500 ms, the cue was presented in the center of the screen. At this point, participants prepared the utterance as much as possible and pressed the space bar with the index finger of their left hand when they were ready to respond. This removed the cue from the screen and initiated an unpredictable variable delay interval lasting between 800 and 1200 ms, during which they were to maintain their readiness to respond. After the delay, the word “Go!” was presented in the center of the screen, accompanied by a 2000 Hz tone, which prompted participants to respond as quickly and as accurately as possible. Immediately upon completion of the utterance, the experimenter judged the perceptual accuracy of the response.

If the response was correct, the next trial started after an intertrial interval of 800 ms. If the response was incorrect, the cue appeared in green letters, accompanied by the auditory model of the correct response. Incorrect responses were defined as responses that did not match the cue, responses with the wrong number of syllables, responses with phonemic errors, and responses that did not match the intended stress

pattern as judged perceptually by the experimenter. Additional error types were premature responses (responses initiated before the go-signal, during the unpredictable delay interval) and slow start errors (responses initiated more than 1000 ms after the go-signal). Both of these error types elicited an error message (“Too early” or “Too slow”) in magenta-colored font for 1000 ms, accompanied by a 3000 Hz tone of 500 ms. Error messages were followed by an 800 ms intertrial interval before the next trial started. Incorrect trials were rerun at the end of each block until 5 correct trials were collected for each of the four responses in a block, or until a total of 60 trials had been run (thus, a block contained between 20 and 60 trials). Participants were informed of this procedure to encourage fast and accurate responses.

Trials were presented in random order, and at the end of each block, the average percentage correct and the average reaction time for that block were presented on the screen (white letters on teal background). This feedback was presented to encourage accurate and fast responses by allowing participants to gauge their performance. Participants then pressed the space bar when they were ready for the next block. After three blocks of the self-selection paradigm, the next phase started with presentation of the cues and responses. The total duration of the experimental session (including paperwork and rest intervals) ranged from approximately 55 minutes to about 80 minutes.

#### 7.1.2.4. Design, Analysis, and Predictions

The primary dependent variables of interest were reaction time (RT) as an index of the SEQ process, and study time (ST) as an index of the INT process. The first self-selection block of each phase was considered a warm-up block and excluded from analysis. Analyses involved 2 (Condition: pure vs. mixed presentation) x 2 (Sequence Length: 3 vs. 5 syllables) x 2 (Rhythm: rhythmic vs. isochronous sequences) repeated measures ANOVAs with repeated measures on all factors, and were performed on the mean per subject per condition (20 observations per participant per condition). Follow-up testing was conducted using Tukey tests.

To reiterate the predictions from the main hypotheses for RT (SEQ), the CH assumes that none of the sequences in this experiment (which involved only repeated syllable sequences) should be integrated, predicting a main effect of sequence length on RT (i.e. longer RT for sequences of five syllables than for sequences of three syllables, regardless of stress). On the other hand, the RH assumes that rhythmic sequences are integrated but isochronous sequences are not, and thus predicts an interaction between sequence length and rhythm, with a sequence length effect present only for isochronous sequences. In addition, the RH predicts that rhythmic sequences should result in faster RTs than isochronous sequences, since only the latter consist of multiple units, and furthermore, that rhythmic sequences of five syllables (one unit) should be initiated faster than isochronous sequences of three syllables (three units).

For ST (INT), the predictions were as follows. According to the CH, if only a single unit is fully preprogrammed (Klapp, 2003), or only all unique units (Immink &

Wright, 2001), then there should be no sequence length effect on ST, since all sequences consisted of the same syllable. However, if all units are preprogrammed (Magnuson et al., in press), then there should be a sequence length effect on ST, for both rhythmic and isochronous sequences, but there should be no difference between rhythmic and isochronous sequences. In contrast, the RH predicts that if only the first unit is preprogrammed, then ST should be longer for integrated, rhythmic sequences than for segregated, isochronous sequences. In addition, one would expect a sequence length effect for rhythmic sequences (on the assumption that programming a five-syllabic unit is more complex than a three-syllabic unit), but not for isochronous sequences if only the first unit (syllable) or all unique units are preprogrammed (if all syllables are preprogrammed then one would also expect a sequence length effect for isochronous sequences, as above).

Finally, if the size of the programming unit depends on presentation condition, then any potential sequence length effects should interact with presentation condition. Specifically, if speakers program only the first unit in mixed blocks but the entire response in pure blocks (e.g., Meyer et al., 2003), then sequence length effects on ST should emerge only in pure blocks, and sequence length effects on RT should emerge only in mixed blocks.

### **7.1.3. Results**

#### **7.1.3.1. Matching Task**

Data from Experiment 3 are presented in Table 7-2. For the matching task, accuracy was calculated based on the number of rerun trials per participant per

condition and expressed as a percentage correct of all trials. A 2 (Condition) x 2 (Rhythm) x 2 (Sequence Length) repeated measures ANOVA revealed no significant effects (all  $F_s < 1$ , except Condition:  $F[1,7] = 1.09$ ,  $p > .25$ ; and Condition and Rhythm ( $F[1,7] = 3.60$ ,  $p = .0997$ ). For the analysis of matching task RT, only correct matching trials were included. Analysis of the matching task was based on log-transformed means to meet the normality assumption; however, untransformed means are presented for interpretability. The 2 (Condition) x 2 (Rhythm) x 2 (Sequence Length) ANOVA revealed a significant main effect of Sequence Length in favor of the long sequences ( $F[1,7] = 6.36$ ,  $p = .0397$ ) and a marginal effect of Rhythm ( $F[1,7] = 4.68$ ,  $p = .0674$ ). The Rhythm x Sequence Length interaction was also significant ( $F[1,7] = 7.27$ ,  $p = .0308$ ) and indicated that RT for the long rhythmic sequence was faster than for all other conditions ( $p_s < .05$ ), which did not differ from each other.

#### **7.1.3.2. Self-selection Task**

Percentage accuracy scores for the self-selection task were calculated based on the rerun trials, per condition per participant. Error data were subjected to 2 (Condition) x 2 (Rhythm) x 2 (Sequence Length) x 2 (Block) ANOVAs, which revealed no significant effects, although there was a marginal Condition x Sequence Length interaction ( $F[1,7] = 4.69$ ,  $p = .0671$ ), suggesting greater accuracy for short sequences than for long sequences in the pure condition (98% vs. 91% accuracy) but not in the mixed condition (both long and short 95% accuracy).

**Table 7-2.** Means (SDs) from Experiment 3.

Task	Condition	Rhythmic			Isochronous		
		3	5	$\Delta$	3	5	$\Delta$
Accuracy (%)	Pure	97 (4)	92 (11)		99 (2)	90 (11)	
	Mixed	95 (4)	96 (6)		96 (5)	93 (7)	
	<i>mean</i>	<i>96 (4)</i>	<i>94 (9)</i>		<i>97 (4)</i>	<i>92 (10)</i>	
ST (INT)	Pure	946 (469)	999 (426)	53	1051 (527)	1174 (569)	123
	Mixed	980 (461)	965 (302)	-15	1153 (479)	1068 (492)	-85
	<i>mean</i>	<i>968 (509)</i>	<i>981 (443)</i>	<i>13</i>	<i>1102 (498)</i>	<i>1121 (526)</i>	<i>19</i>
RT (SEQ)	Pure	506 (73)	532 (79)	26	527 (85)	534 (69)	7
	Mixed	517 (65)	516 (81)	-1	508 (74)	523 (82)	15
	<i>mean</i>	<i>512 (67)</i>	<i>524 (78)</i>	<i>12</i>	<i>518 (78)</i>	<i>529 (73)</i>	<i>11</i>
Matching Acc. (%)	Pure	72 (28)	78 (25)		77 (29)	81 (24)	
	Mixed	87 (21)	84 (23)		83 (20)	83 (22)	
	<i>mean</i>	<i>79 (26)</i>	<i>81 (24)</i>		<i>80 (25)</i>	<i>82 (22)</i>	
Matching RT	Pure	1412 (647)	1025 (484)	-383	1509 (565)	1259 (340)	-251
	Mixed	1579 (591)	1099 (220)	-480	1400 (558)	1401 (498)	1
	<i>mean</i>	<i>1495 (605)</i>	<i>1062 (365)</i>	<i>-432</i>	<i>1455 (546)</i>	<i>1330 (418)</i>	<i>-125</i>

Study Time (INT). STs less than 100 ms or greater than 5 seconds were removed as invalid trials, resulting in a total data loss of 1.6%. Analyses were performed on log-transformed means to achieve normal fit; however, untransformed means are presented for interpretability. Data for /da/ and /gi/ were collapsed as there were no main effects or interactions involving this factor. Since there were effects involving block (2 vs. 3), this factor was included in the ANOVA. This 2 (Condition) x 2 (Rhythm) x 2 (Sequence Length) x 2 (Block) ANOVA revealed a significant effect of Block ( $F[1,7] = 9.93$ ,  $p = .0161$ ,  $\eta_p^2 = .587$ ), which indicated faster ST in block 3 (mean = 1009, SD = 482) than in block 2 (mean = 1076, SD = 449), and a marginal effect of Rhythm ( $F[1,7] = 5.09$ ,  $p = .0587$ ,  $\eta_p^2 = .421$ ), which suggested faster ST for

rhythmic sequences (mean = 973, SD = 410) than for isochronous sequences (mean = 1112, SD = 508). In addition, there was a significant Condition x Block x Rhythm interaction ( $F[1,7] = 9.93$ ,  $p = .0161$ ,  $\eta_p^2 = .587$ ), which indicated that in the mixed condition, the rhythm effect (rhythmic < isochronous) was only significant in block 3, whereas in the pure condition, the rhythm effect was only significant in block 2. No other effects were significant (all  $F_s < 1$ , except Block x Sequence Length:  $F[1,7] = 1.44$ ,  $p = .2692$ , and Block x Rhythm x Sequence Length:  $F[1,7] = 2.10$ ,  $p = .1906$ ).

Reaction Time (SEQ). Analyses were conducted on the means of correct trials only. Since no main effects or interactions were observed with block (2 and 3 in each phase) or syllable (/da/ and /gi/), data were collapsed across these factors, resulting in 20 observations per condition per participant. The 2 (Condition) x 2 (Rhythm) x 2 (Sequence Length) ANOVA on RT revealed no significant effects (all  $F_s < 1$ , except Condition x Rhythm:  $F[1,7] = 2.49$ ,  $p = .1589$ ).

#### **7.1.4. Discussion**

Experiment 3 was designed to assess the INT process and the SEQ process for sequences of repeated syllables as a function of rhythm. No sequence length effects were observed either on RT (SEQ) or on ST (INT). In contrast, there was weak evidence for an effect of rhythm on ST, with rhythmic sequences resulting in shorter ST than isochronous sequences in block 2 (pure condition) and block 3 (mixed condition). No effects of rhythm were observed for RT.

Starting with the SEQ process, the absence of a sequence length effect on RT for the isochronous sequences was unexpected, since both hypotheses predicted a sequence length effect on RT in this case. Based on the logic underlying simple RT, this would suggest that all sequences were programmed as a single unit. It is possible that participants programmed isochronous sequences as a single unit as well, perhaps due to their occurrence in each block with rhythmic sequences or the emphasis placed on timing accuracy (in addition to speed). Informal perceptual analysis during the experiment indicated that participants were generally able to distinguish between the different sequences. In order to accurately differentiate between the rhythmic and isochronous sequences, a useful strategy would be to encode the specific temporal pattern for each sequence and produce the sequence as a single unit.

However, this interpretation runs into difficulties because it predicts a sequence length effect on ST (as observed in Experiment 1): the internal complexity of a unit (defined in terms of the number of syllables) should affect the INT process. No such sequence length effect on ST was observed on ST however, making this interpretation less plausible. One could argue that there was in fact a sequence length effect on ST for rhythmic sequences once visual stimulus processing time is taken into account, since the matching task indicated faster RTs for longer rhythmic sequences but not for isochronous sequences. However, this interpretation is speculative since the matching task always preceded the self-selection task and participants were still learning the mapping between cue and response.



It is possible that ST is not a sufficiently sensitive measure of preprogramming complexity, or that a sequence length effect on ST would have been obtained with a larger sample size or a larger number of trials. However, ST did reveal an effect of rhythm, suggesting that larger effects can be detected, and thus perhaps that if there were a sequence length effect, it was too small to be detected. In order to address this issue further, Experiment 4 was conducted using the exact same target responses in a choice RT paradigm.

The rhythm effect on ST suggests that additional processing was required for isochronous sequences during preprogramming, but this additional processing did not appear to relate to the total number of syllables (or the length of the presumed unit). Neither the RH nor the CH can account for this rhythm effect in their current formulations. The CH fails because it makes no reference to rhythmic structure; the RH fails because it predicted that rhythmic sequences should take longer than isochronous sequences, not shorter. If INT involves preprogramming all units in a sequence, then a sequence length effect should be seen on ST for the isochronous sequences. If INT involves preprogramming of the first unit or all unique units, then the additional processing seen for isochronous sequences compared to rhythmic sequences must relate to the temporal structure of the sequences. Perhaps a temporal frame that includes slots of equal duration is more complex than a temporal frame that specifies different durations for its slots. While it has been proposed in the limb literature that isochronous patterns are actually less complex than rhythmic patterns (Wright & Shea, 2001), it may be that the inherently rhythmic nature of speech makes

it more difficult to produce isochronous sequences, especially for longer sequences where it may be more challenging to determine when to stop speaking (e.g., Klapp et al., 1979). Indeed, although no formal error analysis was undertaken, observations during the experiment suggested that the primary error for long isochronous sequences was an incorrect number of syllables.

Another possibility is that the relatively long sequences used in this experiment imposed a large load on the motor buffer, such that long sequences might not have been loaded completely thereby reducing the SEQ difference. The rhythm effect on ST might then suggest that rhythmic sequences do not impose such a load since they are a single unit. However, this account does not explain why RTs to isochronous sequences were no longer than RTs to rhythmic sequences. One could assume that the rhythmic sequences were in fact not programmed as a single unit but as two units. Recall that the rhythmic sequences were created by prolongation of the third syllable and reduction of the second syllable (to maintain a constant total sequence duration and first syllable duration). This may have led speakers to program the rhythmic sequences with stress on the first and third syllables; however, the attachment of the second syllable (to the first or third syllable) to create a phonological word may have varied between participants, thus making comparisons based on experimenter-imposed target structures invalid. Given that phonological theory assumes that syllables are grouped into binary feet (either trochaic or iambic) (see Gerken & McGregor, 1998), perhaps the use of sequences consisting of an odd number of syllables induced confusion in the speakers regarding the attachment of the short, second syllable. This confusion

may have led some speakers to group the first two syllables together, whereas other speakers may have grouped the second syllable with the third, thereby confounding the results (though note that this explanation does not account for the isochronous sequences).

The possibility of different rhythmic structures imposed on movement sequences is consistent with findings in the motor learning literature that subject-specific rhythms (chunkings) emerge during skill acquisition (Sakai et al., 2004). The present data cannot speak to this possibility, since only perceptual judgments were used to determine whether a response contained the target rhythmic pattern or not. Perhaps a detailed acoustic analysis might help address this issue by identifying chunks on the basis of pause distributions. In addition, acoustic analysis may also help determine whether speakers manipulated factors other than syllable duration to achieve the rhythmic patterns. It is possible that motor programming demands differ for manipulations of pitch vs. duration vs. loudness, and thus if perceptual judgments were insensitive to these distinctions then this may have led to the inclusion in the analysis of productions that failed to meet the task requirements. Unfortunately, acoustic data were not available to address these issues.

In sum, the results from this experiment are puzzling in that the sequence length effects reported in the literature were not replicated. Instead, rhythm appeared to facilitate preprogramming, although the nature of this facilitation remains unclear. There appears to be no satisfactory explanation for these data, and it is possible that methodological aspects of this experiment (e.g., small sample size, small number of

trials) were responsible for the absence of clear effects. One possibility is that the ST measure is not an appropriate measure of preprogramming in this case. Specifically, since this measure has been demonstrated to also be sensitive to the number of units in a sequence (Immink & Wright, 2001; Magnuson et al., in press), it cannot be determined whether the rhythm effect in this experiment (which involved only multisyllabic sequences) is one relating to the complexity of a unit or to the number of units. In Experiment 4, the same target responses were examined using a choice RT paradigm, which has been shown to be sensitive to unit complexity (e.g., Klapp, 1995) but insensitive to the number of units (e.g., Khan et al., 2006; Klapp, 1995).

## **7.2. Experiment 4: INT (Choice RT)**

Experiment 4 was designed to assess the INT process using a standard choice RT paradigm. Experiment 3 failed to find sequence length effects on RT, which might indicate that all sequences were programmed as a single unit, contrary to prediction. If this were the case, a sequence length effect should have emerged on ST, which was not the case. In order to determine whether differences in preprogramming were not detected due to insensitivity of the ST measure, Experiment 4 examined the same sequences as in Experiment 3 using the more traditional choice RT paradigm.

## **7.2.1. Methods**

### **7.2.1.1. Participants**

Eleven college students from the same pool as in Experiment 3 participated in this experiment to obtain course credit. None of these individuals participated in Experiment 3. Three participants were excluded due to bilingualism or neurological history. The remaining eight participants (2M, 6F; mean age = 19, range = 18-21) were monolingual English speakers and reported no learning disabilities, neurological history, or uncorrected visual or auditory impairments. All participants read and signed an informed consent form in accordance with local IRB guidelines.

### **7.2.1.2. Materials and Design**

The materials and design were identical to those in Experiment 3.

### **7.2.1.3. Task and Procedures**

The experimental set-up and procedures were identical to those in Experiment 3, including the matching task. The only difference was the primary task used: in Experiment 4, we employed a choice RT paradigm instead of the self-selection paradigm. In the choice RT task, each trial started with a warning signal consisting of the word “ready” in white letters in the center of a black screen for 1000 ms. This was followed by a blank screen for 1000 ms, after which the cue was presented along with a 75-ms 2000 Hz tone that prompted the participant to produce the target indicated by the cue. Speech onset latency was measured using the voicekey of the SRBox

(Psychology Software Tools, Inc.), and the cue disappeared when the voicekey was triggered. Immediately following the response, the experimenter pressed a button on the button box to indicate whether the response was correct or not. If the response was correct, the next trial started after an 800-msec intertrial interval. If the response was incorrect, the cue appeared in green letters and the response model was played for the participant before the next trial was presented after the intertrial interval. Incorrect responses were defined as in Experiment 1 (except that premature responses did not occur in choice RT). Incorrect responses were rerun at the end of each block until 5 correct responses had been collected for each response or until a total number of 60 trials had been run.

#### **7.2.1.4. Design, Analysis, and Predictions**

The dependent variable was choice RT. The first choice RT block of each phase was considered a warm-up block and excluded from analysis. Analyses involved 2 (Condition: pure vs. mixed) x 2 (Sequence Length: 3 vs. 5 syllables) x 2 (Rhythm: rhythmic vs. isochronous) repeated measures ANOVAs with repeated measures on all factors, and were performed on the mean per subject per condition (20 observations per condition per participant). Follow-up tests were conducted using Tukey tests.

## 7.2.2. Results

### 7.2.2.1. Matching Task

Data from Experiment 4 are presented in Table 7-3. For the matching task, analysis on percent accuracy revealed no significant effects. For the RT analysis, average RTs were computed for each participant per condition, after removal of RTs greater than 5000 ms or less than 100 ms. Only correct responses to matching trials were included, and analyses were performed on log-transformed means to meet the normality assumption (untransformed means are presented for interpretability). A 2 (Condition) x 2 (Rhythm) x 2 (Sequence Length) ANOVA on matching RT revealed no significant effects.

**Table 7-3.** Means (SDs) from Experiment 4.

Task	Condition	Rhythmic			Isochronous		
		3	5	$\Delta$	3	5	$\Delta$
Matching Acc. (%)	Pure	73 (21)	79 (31)		62 (24)	76 (23)	
	Mixed	80 (18)	84 (20)		82 (18)	81 (24)	
	<i>mean</i>	77 (20)	80 (27)		71 (23)	80 (22)	
Matching RT (ms)	Pure	1286 (399)	1177 (315)	-109	1314 (413)	1202 (360)	-112
	Mixed	1149 (217)	1141 (444)	-8	1173 (261)	1308 (511)	35
	<i>mean</i>	1218 (318)	1159 (372)	-59	1243 (342)	1255 (431)	12
Accuracy (%)	Pure	92 (8)	89 (19)		93 (8)	92 (11)	
	Mixed	85 (12)	90 (15)		90 (9)	87 (9)	
	<i>mean</i>	89 (11)	89 (17)		92 (9)	90 (10)	
Choice RT (ms)	Pure	758 (60)	742 (52)	-16	747 (85)	753 (60)	6
	Mixed	755 (64)	743 (74)	-12	733 (61)	786 (66)	53
	<i>mean</i>	757 (60)	742 (62)	-15	740 (72)	769 (63)	29

### 7.2.2.2. Choice RT Paradigm

Percentage accuracy scores were calculated per condition per participant, based on rerun trials. Analysis of the accuracy data revealed only a marginal effect of block ( $F[1,7] = 4.32$ ,  $p = .0763$ ,  $\eta_p^2 = .382$ ) suggesting greater accuracy in block 3 (mean = 93%, SD = 8%) than in block 2 (mean = 89%, SD = 12%), and a significant Block x Rhythm x Sequence Length interaction ( $F[1,7] = 9.37$ ,  $p = .0183$ ,  $\eta_p^2 = .572$ ), which indicated that only the long rhythmic sequences improved significantly ( $p = .0092$ ) in accuracy from block 2 (mean = 87%, SD = 20%) to block 3 (mean = 95%, SD = 8%).

For choice RT, data for the two blocks in each phase were collapsed since there were no main effects or interactions involving this factor; similarly, data from /da/ and /gi/ were collapsed since there were no main effects or interactions involving this factor. The 2 (Condition) x 2 (Rhythm) x 2 (Sequence Length) ANOVAs on the choice RT data (Table 2) revealed no main effects or interactions (all  $F_s < 1$ ; except Rhythm:  $F[1,7] = 1.10$ , Condition x Sequence Length:  $F[1,7] = 1.06$ , and Condition x Rhythm x Sequence Length:  $F[1,7] = 2.08$ , all ns), except for a significant interaction of Rhythm x Sequence Length ( $F[1,7] = 8.90$ ,  $p = .0204$ ,  $\eta_p^2 = .560$ ), which indicated that the rhythmic and isochronous sequences differed in magnitude of the sequence length effect (rhythmic: -15 ms; isochronous: +29 ms). Tukey tests detected a trend of a sequence length effect for isochronous ( $p = .093$ ) but not for rhythmic sequences ( $p = .544$ ).



### 7.2.3. Discussion

Experiment 4 examined the INT stage of speech motor programming by manipulating sequence length and rhythm in repetitions of the same syllable. The results showed no overall sequence length effect but a significant interaction of rhythm and sequence length, in which rhythm facilitated processing for long sequences but not for short sequences. In addition, there was a marginal sequence length effect for the isochronous sequences but not for the rhythmic sequences.

Given the significant interaction between Rhythm and Sequence Length, it appears that rhythmic sequences are programmed differently than isochronous sequences. These findings are inconsistent with a strong form of the CH, since this hypothesis predicted no effects of rhythm, or of sequence length if only the first syllable (or all unique syllables) are preprogrammed. If it is assumed that all syllables are preprogrammed, then one would expect a sequence length effect, but for both isochronous and rhythmic sequences. The RH did predict an interaction between sequence length and rhythm; however, the RH predicted the opposite pattern, namely a sequence length effect for the rhythmic sequences only, on the assumption that rhythmic sequences form a single unit whose complexity varies with the number of syllables.

Recall that Klapp (2003) also found no choice RT effect of sequence length for repeating syllable sequences; however, he did obtain a sequence length effect on simple RT, unlike our findings in Experiment 3. Thus, the absence of sequence length effects cannot be interpreted in this case. Instead, these findings suggest that these

experiments were not sufficiently powerful to detect effects, perhaps due to the relatively small sample size, or an insufficient number of trials, which may be especially relevant factors in a task such as symbol naming that is relatively novel to the speakers and may have led to increased variability.

### **7.3. General Discussion**

The present study was designed to contrast two hypotheses about the nature of speech motor programming, framed within a contemporary model that distinguishes two broad processing stages (Klapp, 1995, 2003). We were interested in the distribution of programming cost to these two stages in relation to the nature of units of speech motor control. Specifically, we examined whether imposing a stress pattern on a sequence of repeating syllables would lead speakers to integrate these syllables into a single unit or whether speakers would choose to process the syllables as a sequence of units. The Composition Hypothesis (CH) states that syllable sequences are only recoded as a single unit when the sequence is composed of different syllables, not when the same syllable is reiterated. The Rhythm Hypothesis (RH) states that a sequence of syllables is recoded into a single unit when the sequence is rhythmic as opposed to isochronous, regardless of the composition of the sequence.

The results from Experiment 3 were surprising, in that no sequence length effect on (simple) RT was observed for isochronous repeated syllables, contrary to findings by Klapp (2003) and Deger and Ziegler (2002), and contrary to predictions by both the CH (which predicted a sequence length effect for isochronous and rhythmic

sequences alike) and the RH (which predicted a sequence length effect only for isochronous sequences).

Based on the assumption that simple RT only reflects the SEQ process and thus the number of units that reside in the motor buffer before initiation, the absence of a sequence length effect would imply that short and long sequences consisted of the same number of units, regardless of rhythmic structure. The absence of a sequence length effect for rhythmic sequences was predicted by the RH and is consistent with earlier work using real words (Sternberg et al., 1978) as well as repeated-syllable pseudowords (see Experiment 1). However, to make a persuasive argument that these syllable sequences are preprogrammed as a single unit, one needs to demonstrate a sequence length effect prior to the SEQ process, i.e. during INT. Thus, for rhythmic sequences, the RH predicted a sequence length effect on ST, the self-selection paradigm's index of the INT process (e.g. Immink & Wright, 2001; Magnuson et al., in press). The CH on the other hand predicted no sequence length effect on ST, unless all units in a sequence are activated and specified during INT (cf. Magnuson et al., in press). Since no sequence length effects were observed on ST nor on RT, the results of this experiment cannot be interpreted in relation to the hypotheses under investigation. In order to determine whether a sequence length effect would emerge using a more traditional and perhaps more sensitive measure of the INT process, i.e. choice RT, Experiment 4 was run. However, again there were no sequence length effects, suggesting that these experiments were not capable of detecting effects given the increased variability associated with this relatively novel task, the small sample size,

and the relatively small number of trials. In addition, it is possible that the specification of a rhythm in terms of temporal parameters alone created difficulty for the participants, either in terms of perception or in terms of attempts to reproduce the patterns. Thus, perhaps using different manipulations to create the rhythmic patterns (i.e. pitch and loudness variation) would have resulted in results more in line with the predictions from either hypothesis, and future work aimed at specifying the role of rhythm and stress in motor programming should include such manipulations as well.

It is possible that speakers parsed the sequences in different ways (e.g., Sakai et al., 2004), resulting in different unit sizes for individual participants. Moreover, these chunkings could vary as a result of practice. These possibilities suggest that a detailed analysis of the produced responses might shed light on the RT findings, if subgroups of speakers could be identified based on how they group the syllables in these sequences. However, these data are not available and therefore, this must remain speculation.

Finally, the finding that an advantage was found for rhythmic sequences relative to isochronous sequences on ST and choice RT, suggests that rhythm does play a role during speech motor preprogramming. In Experiment 4, this advantage of rhythm was observed only for long sequences, suggesting that a rhythmic structure is especially beneficial when relatively long sequences must be produced, perhaps because it reduces the reliance on a counting strategy to determine when to stop speaking (cf. Klapp et al., 1979). However, given the absence of sequence length

effects, which would be predicted if the rhythmic sequences were produced as a single unit, the exact nature of this benefit is unclear.

### **Acknowledgements**

Chapter 7 is being prepared for publication, under the running title “The Role of Rhythm in Speech Motor Programming, by Edwin Maas and Donald A. Robin. The dissertation author is the primary investigator and author of the manuscript.

## **8.0. Discussion and Conclusions**

The focus of this dissertation is on the motor processes and representations involved in production of fluent, articulate speech at the planning and programming levels and in particular breakdowns in motor programming as a result of apraxia of speech from neurological origin. Given the exceptionally high level of demand generated by the coordination of speech muscles and their subsequent movement involved in even the simplest utterance, it is remarkable that speech production is automatic in the adult speaker. As noted in Chapter 1, in order to understand speech motor control and its breakdowns, it is necessary to take a broad view and incorporate knowledge of motor control of other movement systems and information pertaining to speech motor learning as integral components of any theory of speech motor control. In addition, models of unimpaired speech motor control and learning should be amenable to explaining a variety of speech disorders. Thus, evidence from speech disorders may provide important constraints on the development of cognitive models of speech motor control and learning. In turn, these models should be applied to understanding speech production disorders, as well as guide assessment and intervention. The remainder of this discussion will be divided into four parts, each of which discusses the findings from this dissertation in relation to the two main aims outlined in Chapter 1, namely 1) the nature of speech motor programming in the intact system, and 2) speech motor programming in motor speech disorders, esp. AOS. An overview of the results is provided in Table 8-1.

**Table 8-1.** Overview of the results from the four experiments. Numbers 1, 3, 4, and 5 refer to number of syllables; S = short duration (150 ms); L = long duration (450 ms); R = Rhythmic sequence; I = Isochronous sequence; AOS = apraxia of speech group; APH = patients with aphasia; CON = control participants.

<b>Hypothesis</b>	<b>Prediction</b>		<b>Results</b>		<b>Interpretation</b>
	<i>ST (INT)</i>	<i>RT (SEQ)</i>	<i>ST (INT)</i>	<i>RT (SEQ)</i>	
<b><i>Experiment 1 – speech</i></b>					
Unit complexity affects INT, not SEQ	S < L	S = L	S = L	S = L	Syllable duration does not affect INT (or SEQ)
H1: Syllables are units H2: Sequences are units	H1: 1 < 4 H2: 1 < 4	H1: 1 < 4 H2: 1 = 4	1 < 4	1 = 4	Sequences are units
AOS = INT deficit, not SEQ deficit	AOS > CON	AOS = CON	AOS > CON	AOS = CON	INT deficit, not SEQ deficit
APH = INT and SEQ intact	APH = CON	APH = CON	APH = CON	APH = CON	No INT or SEQ deficit
<b><i>Experiment 2 – finger</i></b>					
Unit complexity affects INT, not SEQ	S < L	S = L	S = L	S = L	Press duration does not affect INT or SEQ; failure to replicate perhaps due to methodological differences
H1: Presses are units H2: Sequences are units	H1: 1 < 4 H2: 1 < 4	H1: 1 < 4 H2: 1 = 4	1 < 4	1 < 4	Presses are units
H3: AOS = central INT-deficit H4: AOS = speech-specific INT-deficit	H3: AOS > CON H4: AOS = CON	H3: AOS = CON H4: AOS = CON	AOS > CON	AOS = CON	INT deficit also in nonspeech finger movements; AOS is not speech-specific
H5: APH = INT and SEQ intact	APH = CON	APH = CON	APH = CON	APH = CON	No evidence for motor programming deficits; but N=1 (lefthanded)

**Table 8-1 (Continued).** Overview of the results from the four experiments. Numbers 1, 3, 4, and 5 refer to number of syllables; S = short duration (150 ms); L = long duration (450 ms); R = Rhythmic sequence; I = Isochronous sequence; AOS = apraxia of speech group; APH = patients with aphasia; CON = control participants.

<b>Hypothesis</b>	<b>Prediction</b>		<b>Results</b>		<b>Interpretation</b>
<i>Experiment 3 – rhythm</i>	<i>ST (INT)</i>	<i>RT (SEQ)</i>	<i>ST (INT)</i>	<i>RT (SEQ)</i>	
H1: sequences only programmed as single unit when syllables differ	3 = 5 R = I	H1: 3 < 5 R = I	3 = 5 R < I	3 = 5 R = I	All sequences programmed as single unit; rhythmic sequences easier to preprogram than isochronous sequences; however, methodological concerns
H2: sequences integrated only with rhythmic structure	3 < 5 for R only R > I	3 < 5 for I only R < I			
<i>Experiment 4 – rhythm</i>	<i>Choice RT (INT)</i>		<i>Choice RT (INT)</i>		
H1: sequences only programmed as single unit when syllables differ	3 = 5 R = I		3 < 5 for I only R < I for 5 only		Rhythm facilitates preprogramming for long sequences only; however, methodological concerns
H2: sequences integrated only with rhythmic structure	3 < 5 for R only R > I				



### **8.1. Speech Motor Programming in the Intact System**

Ever since Lashley (1951), the assumption that control of serial behavior involves a hierarchical architecture in which a unit at one level controls multiple units at a lower level has been widely accepted and incorporated into models of motor behavior, including speech (e.g. Sternberg et al., 1978; Klapp, 1995; Schmidt, 1975; Levelt et al., 1999). The main charge for such models is to identify and specify the putative processing levels and the units presumed to operate at each of these levels.

With respect to levels of processing, fundamental questions relate to their independence and temporal organization relative to each other. One issue that was raised at several points throughout the dissertation was the widely-accepted distinction between phonological encoding and speech motor programming. Most models implicitly or explicitly distinguish between language and speech, but it has been difficult to demonstrate conclusively the relative contributions of these levels to the results in any given experiment (Rogers & Storkel, 1998).

Thus, for instance, Klapp (2003) described his model entirely in terms of motor programming constructs such as motor programs, spatiotemporal structure, and temporal frames, but it may be that the effects can be captured equally well in phonological terms. Similarly, Santiago et al. (2000) described their findings of longer choice RT for bisyllabic words relative to monosyllabic words in terms of phonological constructs such as syllable structure and phonemes, but the findings might just as well be accounted for in terms of motor level constructs.

Experiment 1 attempted to circumvent potential confounds with phonological encoding by manipulating a variable that can be localized to the motor control level, namely time. The assertion here is that introducing a change at the phonological level will have ramifications for the subsequent motor stages, and thus that defining motor programming in terms of strictly phonological terms is fundamentally inadequate to tease these putative stages of processing apart. In contrast, if phonological encoding and motor programming are separate, serial, and strictly feedforward, then manipulating a variable at the motor level should produce effects that can be ascribed to the motor level alone. In Experiment 1, the motoric variable of timing was chosen as one main independent variable, in part because the INT/SEQ model that was the focus of the study (Klapp, 1995; 2003) was based on effects involving duration of responses, and in part because control of timing is not a unitary construct but can be further divided into relative timing and absolute timing, each of which is hypothesized to be controlled by fundamentally different mechanisms (e.g., generalized motor programs vs. parameters; Schmidt, 1975, 2003; Schmidt & Lee, 2005), and each of which may be differentially impaired (e.g., Clark & Robin, 1998) and respond differentially to various treatment conditions (see Maas et al., in preparation, for review).

The control speakers in Experiment 1 failed to produce reliable effects of single-syllable duration on Study Time, the measure of the INT process. The fact that speakers were quite able to produce the intended target durations and clearly differentiated the short and long syllables suggests that the null effect was not due to a

failure to perform the task adequately, which might have resulted from the absence of detailed and reliable feedback about the timing accuracy on single syllable duration. An alternative explanation is that programming of syllable duration is relatively easy for adults who have had extensive practice producing syllables with a variety of durations, and thus, that this manipulation did not substantially (measurably) tax the INT process in speech. Perhaps effects of syllable duration would be obtained for children, for whom speech is presumably a less well-developed motor skill.

However, the fact that no duration effect was found in Experiment 2 for the finger movements on which the model was originally based, suggests that the absence of a duration effect may not be related to the highly skilled status of speech. In addition, this finding argues against an explanation in terms of failure to achieve the task requirements due to absence of feedback, since in Experiment 2 the exact same feedback schedule and type were provided as those in the original self-selection studies (e.g., Immink & Wright, 2001; Wright et al., 2004).

Rather, the reason for this null effect of duration may be methodological in nature. One critical difference with previous work relates to the duration of the RT window after the go-signal. Recall that in order to ensure that the clinical participants would be able to respond during this interval without eliciting “slow start” errors (especially relevant for the AOS participants given the often-reported initiation difficulties in this population, e.g., Dabul, 2000; McNeil et al., 2000), the RT window increased to 1000 ms (compared to 400 ms in Immink & Wright, 2001; Wright et al.,

2004). Thus, it is possible that with shorter RT windows effects of syllable duration will be found.

In contrast to the duration findings (rather, absence thereof), both Experiment 1 and Experiment 2 produced reliable sequence length effects. However, the pattern differed between finger movements and speech movements. For speech, the sequence length effect was confined to the ST interval and did not appear on RT, whereas for finger movements, a sequence length emerged on both ST and RT, replicating previous work using this paradigm (e.g., Immink & Wright, 2001; Wright et al., 2004). Based on the logic underlying the model and the methodology, these findings suggest that for finger movements, button presses were programmed as separate units. The sequence length effect on ST indicated that more than one button press was preprogrammed (presumably all button presses, cf. Magnuson et al., in press), but that participants were unable to integrate these multiple units into a single coherent processing unit and load them into the motor buffer as a single item.

For speech, the findings suggest that repeated syllable sequences are preprogrammed as a single unit, as no RT difference was observed here. Whether the sequence length effect on ST reflects the delivery of a single unit to the INT process from a preceding (phonological) level, or whether it reflects the delivery of multiple, separate syllables that are subsequently integrated into a single unit during INT cannot be determined based on these data. In any case, the sequence length did not appear on RT, suggesting that for these speech sequences, the load on SEQ did not differ between single syllables and sequences of four syllables, and thus, that these

sequences, rather than individual syllables, constituted units of speech motor programming.

Several points are worth making in regard to these findings. Importantly, these results suggest a difference between finger and speech movements. Although this could be viewed as support for the task-dependent model (Ziegler, 2003a, b), recent findings from the motor learning literature show that with extensive random practice, a sequence of units may become permanently recoded or chunked as a single unit (e.g., Klapp, 1995; Sakai et al., 2004; Verwey, 1999, 2003; Wright et al., 2004; see Rhodes et al., 2004, for a review). Thus, the difference between the finger and speech findings most likely reflects differences in the amount of practice rather than qualitative differences between the two motor skills. Virtually all hearing adults are highly skilled in sequencing syllables due to a lifetime of speaking; in contrast, the sequencing of finger movements is presumably much less-practiced in the typical adult, except perhaps in skilled musicians. It would be interesting to see whether sequencing of novel (e.g. non-native) syllables would produce a sequence length effect on RT, and whether this effect disappears with extensive practice.

Second, it is possible that the observed sequence length effect on ST arises at a level of processing preceding INT, for example phonological encoding. Thus, this study suffers from the same ambiguity in attributing effects to one or the other level as most other studies on phonological encoding or speech motor programming (Rogers & Storkel, 1998), and it is possible that sequence length affects both phonological and motor programming levels due to their intimate connection. If we had found that

patients with AOS demonstrate a disproportionate sequence length effect, this would have supported a motor programming interpretation; similarly, a disproportionate effect for the patients with phonological impairments would have suggested a phonological interpretation. However, no such effects were found for either the patients with aphasia or the patients with AOS, perhaps due to the relatively small sample sizes, the large variability in the control group, and/or the relatively mild impairments in these patients. As a result, the sequence length effect in the controls cannot be ascribed with confidence to one or the other level of processing. Note however that a phonological explanation does not account for the sequence length effect in finger movements.

Third, the interpretation that sequences of repeated syllables are programmed as a single unit is in contrast with earlier findings (e.g., Deger & Ziegler, 2002; Klapp, 2003; Sternberg et al., 1978), who observed that repeated syllables produced a sequence length effect on RT. One possible explanation is that the generally longer simple RTs observed in Experiment 1 relative to those reported in the literature is responsible for this absence of a sequence length effect. Klapp (2003) used a qualification procedure in which speakers were not run on the experiment if they failed to meet a RT criterion of less than 350 ms. However, the absence of a sequence length effect on RT is nonetheless surprising given that the difference in number of syllables was relatively large (1 vs. 4 syllables).

One critical difference between Experiment 1 and previous studies was the temporal structure of the sequences. In particular, Experiment 1 involved responses

with specific target syllable durations and relative timing patterns, whereas previous studies have not specified either the target temporal patterns or the produced temporal patterns of the responses. Based on the fact that most real words consist of combinations of different syllables rather than the same syllable, it is plausible that speakers in previous studies processed the syllable repetitions as reiterations of the same word, leading them to produce each syllable as stressed, i.e. with a list-like prosody. It is possible that the rhythmic (temporal) structure imposed on the target sequences in Experiment 1 encouraged speakers to process the syllables as a single GMP or scalable response structure (Shea & Wulf, 2005), in order to achieve a stable relative timing pattern across the sequence. Thus, the use of a single GMP that governs the relative properties of these rhythmic sequences would eliminate any differences in terms of SEQ load; a sequence would involve a more complex unit than a single syllable, but still only one unit.

The proposal that the rhythmic (temporal) structure of the sequences in Experiment 1 facilitated the programming of the syllables as a single unit was tested further in Experiments 3 and 4, which were designed to assess the distribution of processing cost to INT vs. SEQ as a function of rhythmic structure and sequence length for repeated syllables. Experiment 3 also failed to produce a sequence length effect on RT, but rather than a sequence length effect on ST, an effect of rhythmic structure emerged in which rhythmic sequences were prepared faster than isochronous sequences. Thus, unlike the findings in Experiment 1, where a sequence length effect was indeed observed but only on ST, the absence of sequence length effects in

Experiment 3 does not allow us to conclude whether the sequences were programmed as a single unit or as a sequence of units.

However, the benefit of a rhythmic structure over an isochronous structure must be explained. One interpretation would be that all sequences were programmed as a single unit (contrary to Klapp's [2003] suggestion that repeated syllables are not integrated) and that isochronous sequences constitute more complex units than rhythmic sequences. Although in the motor learning literature, isochronous sequences have been considered simpler than rhythmic sequences (e.g., Wright & Shea, 2001), it could be that the rhythmic nature of speech assists in the coordination of various aspects of the movement (e.g., Kent et al., 2000) and therefore makes rhythmic sequences easier than isochronous sequences. One benefit of a rhythmic structure for the relatively long sequences studied here is that it appears to minimize confusion about when to terminate production of the sequence (e.g., Klapp et al., 1979). Although no formal error analysis was undertaken, observations during the experiment suggested that a prevalent error type for the long isochronous sequences (but not for the long rhythmic sequences) was the addition or omission of a syllable. These observations might suggest that the response structures for isochronous sequences are less stable than those for rhythmic sequences, and as such require longer time to activate.

An alternative explanation for these findings is that only the first syllable was preprogrammed for all sequences, as was proposed by Klapp (2003). Recall that the initial syllable was kept constant across conditions, in order to be able to isolate any



effects to the temporal structure of the sequence. While we would then expect a sequence length effect on RT, it has been noted before that the sequence length effect is less robust for longer sequences (Klapp et al., 1979). In order to account for the faster ST for rhythmic as opposed to isochronous sequences, this account would have to assume that loading a temporal frame for a rhythmic sequence is less demanding than loading a temporal frame (or multiple single-slot frames) for isochronous sequences.

In an effort to determine whether all sequences were indeed integrated or whether only the first syllable were preprogrammed, Experiment 4 examined the same responses using the more traditional choice RT paradigm. In this experiment, a benefit of rhythmic structure (relative to isochrony) was again observed, but this time the rhythmic structure interacted with sequence length in that the rhythm benefit occurred only for long sequences but not for short sequences. If it is assumed that all sequences constitute a single unit, then this would suggest that long isochronous sequences are more complex than the other sequences. On the other hand, if it is assumed that repeated syllable sequences are programmed as a series of separate units (Klapp, 2003), then these findings suggest that choice RT includes SEQ (as in the original version of the model, Klapp, 1995) and further, that the SEQ process can take longer than the INT process (contrary to the assumption in the original model).

While the results from Experiments 3 and 4 remain puzzling, they do show that the rhythmic nature of the sequence affects the programming of syllable sequences, and specifically, that a rhythmic (temporal) structure appears to facilitate the

programming of such sequences. No firm conclusions can be drawn from these experiments however, due to the failure to replicate the predicted sequence length effect on RT or ST. It is possible that the small sample size and number of trials, in combination with the proposal that person-specific rhythms emerge for sequential movements and that these rhythms change with practice (Sakai et al., 2004) obscured any potential effects of interest. No data related to the actual responses produced were available.

The most robust findings of this dissertation emerged from the experiments on AOS (i.e. Experiments 1 and 2). A neuropsychological approach to understanding normal motor programming was employed. Though the sample size was small, particularly for the participants with aphasia, statistical analysis comparing each subject's performance to the mean of the control group using a relatively new method that has been shown to control alpha level despite small sample sizes and departures from normality (e.g., Crawford & Garthwaite, 2005; Crawford & Howell, 1998) clearly demonstrates that three of the four subjects with AOS differed from unimpaired controls, whereas only one of three individuals with aphasia differed from the unimpaired group. The finding that patients with AOS demonstrated a longer ST but a normal RT was interpreted as a deficit in the INT process with the SEQ process being intact. The fact that a similar pattern emerged for nonspeech finger movements (Experiment 2) constrained the interpretation of the ST difference in Experiment 1, in that the INT deficit appears to affect motor stages of processing rather than phonological stages or processes related to verbal working memory. While it remains

a possibility that the longer STs in speech are unrelated to the longer STs in finger movements and represent independent impairments, this seems less plausible (or at least less parsimonious) in light of the finding that it was the same three patients with AOS in each experiment who demonstrated the effects.

Thus, a specific deficit during the ST interval without deficits during the RT interval was predicted by the hypothesis that AOS involves an INT-specific impairment, and is consistent with the assumption that INT and SEQ are separate processes (Klapp, 2003). In addition, the finding that speakers with AOS were impaired for both speech and finger movements suggests that speech and nonspeech motor skills are controlled by overlapping neuromotor control systems (e.g. Ballard et al., 2003). According to an integrative view of the motor system (Ballard et al., 2003), speech is not controlled by a separate speech motor control module (Ziegler, 2003) but rather by a motor system that controls specific aspects of movements that are relevant to speech but may also be relevant in other motor tasks. For example, speech involves repeated opening-closing movements of the jaw in which peak velocity occurs in the middle of movements (Folkins et al., 1995), production of sound, high velocities, coordination of multiple structures, constraints on absolute and relative timing of movements, etc. In Experiments 1 and 2, we demonstrated that speech movements and finger movements (that produced sound) with identical absolute and relative timing requirements were associated with increased preprogramming time in the same individuals with AOS, supporting the contention that these movements share neural and cognitive control mechanisms.

Finally, the observed dissociations between absolute and relative timing in individuals with AOS provided further support for the idea that these aspects of a movement are controlled by different mechanisms, as proposed by the schema theory of motor control (Schmidt, 1975; Schmidt & Lee, 2005). Again, neuropsychological evidence converges with evidence from other sources such as the differential effects of motor learning variables on absolute and relative aspects of movements (see Maas et al., submitted, for review). Furthermore, the application of measures of absolute and relative timing used in the motor learning literature to speech suggest that the concepts of generalized motor programs (or scalable response structures; Shea & Wulf, 2005) and parameters have utility in understanding speech motor control and its disorders.

## **8.2. Speech Motor Programming in Apraxia of Speech**

While the general consensus among researchers and clinicians seems to be that AOS represents a speech motor programming disorder (e.g., Aichert & Ziegler, 2004; Deger & Ziegler, 2002; Darley et al., 1975; McNeil et al., 1997, 2000; Wambaugh et al., 2006), the precise nature of this motor programming disorder remains relatively poorly understood. The work presented here represents an effort to further our understanding of the nature of the deficit in AOS, by applying a time course approach to the study of speech motor programming in AOS. Such an approach may be better suited to localizing the deficit than approaches based on analysis of final output data alone, by virtue of the fact that the time course approach allows one to tap into different stages as they unfold in time.

Framed within the INT/SEQ model, the primary hypothesis was that AOS reflects a disruption at the preprogramming (INT) stage, whereas the SEQ stage was thought to be intact. Using the self-selection paradigm, this hypothesis was supported in that three out of four speakers with AOS (but only one out of three individuals with aphasia) demonstrated significantly longer INT processing than controls, whereas there was no slowing of the SEQ process in AOS (nor in aphasia). These results are interesting not only because they further localize the deficit to a specific processing stage, but also because they suggest that the frequently reported initiation difficulties observed in AOS may in fact be difficulties in preprogramming. That is, when given time to prepare an utterance, there is no difference in the ability to initiate movement by sending a go-signal to the muscles. Thus, the results from this study indicate that programming can be empirically distinguished from initiation (sending an electric pulse to the muscles).

Alternative accounts for the observed findings were considered in more detail in Chapters 5 and 6. While several of these could account for certain aspects of the data, there were several problems with most of them. For instance, the reduced buffer capacity hypothesis (Rogers & Storkel, 1999) can account for the absence of a sequence length effect on RT but would have to be modified to account for the longer ST for finger movements, and does not account easily for the observed longer ST for single syllables (which should be within buffer capacity) nor the sequence length effect on ST (which suggests that multiple syllables are being programmed).

Similarly, an explanation in terms of verbal working memory could account for the absence of a sequence length effect on RT by assuming that items cannot be maintained in the buffer. The ST effect could be explained if it is assumed that failure to maintain the sequence in a buffer results in continuous reprogramming (although one might then expect the sequence length to re-emerge on RT, contrary to our findings). Although we did not obtain measures of working memory on these patients, the exact relationship between working memory and speech production is still a matter of debate. For example, Waters et al. (1992) have shown that patients with AOS are impaired on measures of verbal working memory. However, it seems more likely that the speech motor programming deficit is the source of the working memory problem rather than the other way around. Furthermore, this hypothesis does not account for longer ST in finger movements.

The dual route hypothesis (Varley & Whiteside, 2001a, b) could account for the observed findings, as the distinction between the direct and indirect routes was argued to be similar to the distinction between INT and SEQ, respectively. However, since the dual route hypothesis is restricted to speech motor programming, it also fails to account for the longer ST in the finger movement study.

The ST findings for the AOS speakers in Experiment 1 could in principle reflect a phonological level deficit. However, since there was no evidence for phonological level impairments during assessment (e.g., no serial order errors) this seems unlikely. Moreover, the fact that the two patients who did have positive evidence for phonological impairment did not demonstrate longer ST relative to

controls argues against this hypothesis. Finally, the longer ST for finger movements cannot be explained in terms of a phonological deficit.

An alternative hypothesis was that the longer ST in AOS reflects a problem in visual stimulus processing; this account fails because it predicts longer RTs as well. And although an account in terms of a stimulus-response mapping deficit could account for the longer ST in both speech and finger experiments, this hypothesis does not explain the presence of absolute or relative timing errors in the responses produced by the patients. However, this hypothesis cannot be ruled out on the basis of these data, and future experiments should include a task or condition in which the stimulus-response mapping stage is manipulated (e.g., by varying the number of response alternatives or the transparency of the mapping).

Finally, a hypothesis was considered that explains the longer ST for the patients in terms of self-monitoring or self-judgment of readiness, where it is assumed that individuals with AOS are less confident about their speech (e.g., due to their experience with speech difficulties since their stroke), and as a result need more time to feel ready to respond. While this hypothesis cannot be ruled out, it is less plausible to assume that a similar need for longer preparation time existed for finger movements. In addition, this hypothesis suffers from a difficulty in operationally defining the constructs of self-judgment and confidence. More importantly, the hypothesis would have to assume that this judgment interacts somehow with response characteristics, in order to account for the observed sequence length effect on ST.

While most of these alternative hypotheses can be adapted to account for the data, the hypothesis of an INT deficit predicted these results, and can accommodate the nonspeech findings based on the assumption that motor programming involves a stage at which movements and movement sequences are programmed in relatively abstract (effector-independent) terms (e.g., Keele et al., 1995; Schmidt & Lee, 2005). Thus, these findings appear most compatible with a deficit at the preprogramming stage of processing in AOS.

Another way in which the results from Experiments 1 and 2 further the understanding of the deficit in AOS is through the comparison of speech and nonspeech motor programming. Two different theoretical positions exist on the relation between speech motor control and nonspeech motor control. According to the task-dependent model (Ziegler, 2003a,b; cf. also Weismer, 2006), speech is subserved by a separate and specialized speech motor control module; according to the integrative model (e.g., Ballard et al., 2003) speech and nonspeech motor control share neuromotor control systems to the extent that the tasks share properties. In Experiments 1 and 2, the same speakers with AOS participated in a speech experiment and in a nonspeech (finger movement) experiment. According to the task-dependent model, there is no reason to expect problems with finger movement programming; according to the integrative model, similar problems are expected since the tasks share many properties. The fact that the same three patients who demonstrated impairment in INT for speech were also the same three who showed an INT impairment for finger



movements supports the idea that speech and nonspeech motor skills may be governed by an overlapping neuromotor control system.

The inclusion of motor execution level measures (i.e. absolute and relative timing error) allowed a further refinement of the nature of the deficit. Specifically, absolute and relative timing error dissociated in AOS (but not in our patients with aphasia), suggesting that there was not some generalized deficit in timing control, but rather that each can be impaired independently. This extends the findings from Clark and Robin (1998) from a visuomotor tracking task to timing control in speech production. However, there was a difference between the speech and finger findings for these patients. In particular, for speech movements, there was evidence for disproportionate impairment of absolute timing in two of the four patients with AOS, whereas for the finger movements, three of the patients with AOS showed disproportionate impairment of relative timing.

It is possible that these findings reflect different control systems for speech and nonspeech tasks, as would be expected from the viewpoint of the task-dependent model (Ziegler, 2003). However, it should be kept in mind that the experiments were separated by approximately one year, which means that the relative emphasis placed on absolute or relative timing accuracy by these individuals may have changed over time. Clark and Robin also observed that either relative timing, or absolute timing, but not both were impaired in their patients with AOS, whereas no such dissociations occurred for their patients with conduction aphasia. They suggested that control of relative timing and control of absolute timing may draw on the same pool of resources,

and that patients with AOS may choose to optimize either one at the expense of the other. Moreover, they observed that the two different patterns occurred in a single patient at different times. It is thus possible that the patients in Experiments 1 and 2 changed their strategy in terms of optimizing relative timing accuracy and absolute timing accuracy in between the two experiments. A learning experiment that examines patterns of transfer between speech and limb movements might shed more light on this potential speech-nonspeech difference.

### **8.3. Limitations**

Limitations of the present work include, first, the relatively small number of participants tested, and the limited range of severity of our patients. This necessarily restricts the generalizability of the findings to a wider population, and as such, the findings from these studies must be interpreted with caution. Naturally, these studies require replication, and should be extended to include other populations in order to further determine whether the observed dissociation in these patients with AOS can be substantiated and supplemented by a reverse dissociation.

Second, no detailed lesion data was available for the patients in this study. Though the diagnosis of AOS is based on behavioral data and the underlying neural substrates are debated (e.g., Dronkers, 1996; Hillis, Work, Barker, Jacobs, Breese, & Maurer, 2004), such data may help further delineate the nature of the impairments in these patients. The present study was a behavioral study, based largely on the fact that AOS is a behaviorally-defined disorder. However, it is recognized that subtypes of

AOS may exist (e.g., Duffy, 2005) and that lesion information may contribute to our understanding of the disorder, if differences in lesion site or extent can be related in a meaningful way to observed behaviors in AOS. Future work will include both behavioral and neural data. In addition, we did not obtain measures of working memory for these patients, and it is possible that reductions in working memory capacity may account for some aspects of the data. Although most behavioral measures of working memory (e.g., digit span, word span) cannot clearly distinguish between a true working memory problem and a problem in speech motor programming which disrupts the maintenance of information in working memory, lesion data may also help dissociate these possibilities.

Third, the reaction time approach taken in this dissertation is limited in its ability to specify the exact nature of the processing that occurs during the various RT intervals. Although careful manipulation of only those factors of interest can shed light on the ongoing processes, ultimately the findings from RT methodologies must be integrated with findings from other measures, including for instance acoustic analyses, kinematic measures, and neural imaging. Recently, initial attempts at using neuroimaging to further specify the processes in these types of RT paradigms have recently been made (e.g., Bohland & Guenther, in press).

Fourth, and relatedly, the range of speech motor programming variables studied here was limited. Although the choice to manipulate syllable duration as a factor presumably free from phonological import was based on the finger movement literature, this manipulation did not affect either ST or RT. It is possible that other

variables thought to affect the speech motor programming process (e.g., speaking with a biteblock, specification of overall pitch and loudness parameters) would have resulted in measurable effects. The failure to replicate the duration effect disallowed an assessment of disproportionate complexity effects in our patients. To tease apart phonological from motor level processes, future studies should include a range of response types defined in terms of putative phonological and motoric variables and examine their interaction; however, this was not the primary goal of this dissertation.

Fifth, only pseudowords were used in these experiments, and it is possible that real words are programmed differently than nonwords. For example, in the task-dependent view of speech motor control (Ziegler, 2003a,b), nonwords might be viewed as nonspeech (Ziegler, 2006), and thus may be processed by the novel oral motor control system rather than by the speech motor control system. Although this view would have to be elaborated to account for the fact that people can and do learn new words throughout their life (suggesting that the same motor response is somehow shifted from the novel motor control system into the speech motor control system), and also fails to explain how treatment targeting nonwords can transfer to the production of real words (e.g., Maas et al., 2002; Schneider & Frens, 2005), future studies should also include real words to study speech motor programming.

Sixth, we did not observe any effects on RT for the speech experiments, only for the finger experiments. While we did expect to see sequence length effects on RT in Experiments 1 and 3, it is unclear whether this failure represents a problem in the methodology or an indication that sequences were preprogrammed. It should be noted

that effects of rhythm were obtained in Experiment 4, suggesting that this absence of sequence length effects is not due to equipment malfunction. Future work should include a condition in which multiple words are included, in order to demonstrate the ability of RT to register effects.

#### **8.4. Conclusion and Future Directions**

In conclusion, keeping the limitations of the studies in mind, this dissertation contributes to our understanding of speech motor programming, and of AOS. In addition, the applicability of the self-selection paradigm to the study of speech production received moderate support, in that some but not all previously reported effects were replicated. The paradigm merits further investigation to determine the optimal methodological parameters to apply this paradigm to speech, since unlike the comparison of simple and choice RT, the self-selection paradigm allows for the within-trial, within-participant assessment of distinct processing stages involved in motor programming. The present work represents only the first extension of this paradigm to speech, and explored the effect of factors that had been found to affect motor programming of nonspeech movements. The fact that for example the duration effect was not replicated does not necessarily that the paradigm is flawed; rather, it may be that specification of absolute duration within the range examined here did not sufficiently tax the motor system, or that a smaller RT window is required to encourage speakers to rely on preprogramming.

With respect to the observed sequence length effects on ST but not RT in Experiment 1, these suggest that the number of syllables is a relevant factor in early processes but not in later-occurring processes. This is consistent with the INT/SEQ model, which assumes that unit complexity affects the preprogramming stage (INT) but not the buffer scanning and retrieval operations (SEQ). Although previous studies had found sequence length effects on simple RT for repeated syllables, these studies did not specify the temporal patterns of their target utterances, and it is possible that the presence of a rhythmic structure (as opposed to an isochronous structure) facilitates processing a sequence as a single unit. Unfortunately, a direct test of this possibility failed to produce reliable sequence length effects, despite demonstrating a benefit of a rhythmic structure relative to isochrony. Thus, these findings cannot be easily interpreted, and do not necessarily speak to the findings of Experiment 1, in which a clear sequence length effect was obtained for ST.

It should be noted that this sequence length effect could in principle also have arisen at the level of phonological encoding, since the number of syllables presumably affects phonological encoding as well as speech motor programming. The present data do not allow us to distinguish between a phonological encoding account and a motor programming (INT) account (although the explanation in terms of INT would naturally accommodate the sequence length effect for finger movements, whereas the phonological account would have to postulate an additional, separate mechanism for that finding). In order to distinguish these accounts, future work may cross sequence

length with unambiguously motoric factors (such as overall pitch or loudness levels, speaking with a biteblock) and determine if these factors are additive or interactive.

With respect to the nature of AOS, the findings further delineated the nature of this motor programming disorder, at least in the patients studied here, in that an impairment was found localized to a specific stage of processing, i.e. the preprogramming stage. Our hypothesis was that AOS reflects a localized but not speech-specific impairment of one stage of motor programming (i.e. INT) whereas another stage of motor programming (i.e. SEQ) was hypothesized to be intact. Several predictions emanating from this hypothesis were confirmed, namely: 1) longer ST than age-matched controls, 2) no differences in RT relative to controls, 3) reduced timing error (at least for some of the patients), and 4) the similarity of the ST and RT patterns for speech and nonspeech movements. Although these findings were predicted by our hypothesis, alternative explanations were also discussed in Chapters 5-6 that could account for some aspects of the data (though most do not account for the nonspeech findings). Thus, although the hypothesis of an INT deficit predicted the pattern of results, and was not falsified (e.g., by abnormal RT effects), further research is clearly needed to further elaborate the exact nature of processing that occurs during this stage of preprogramming, especially as it relates to speech production.

Future directions within this research program will address several issues. With respect to unimpaired speech motor programming, one major question relates to the independence and temporal organization of phonological and motor stages of processing. One way to approach this question is by looking for interactions between

putative phonological and motoric variables. Another approach would be to use a partial precue paradigm, in which one aspect of the response is specified in advance while another aspect of the response is specified by the imperative stimulus. The logic is that if the two aspects are independent, then advance specification of a response attribute should allow speakers to preprogram this attribute, abolishing its effects on RT. However, if the specification of an attribute depends on specification of the other, then precuing should not lead to preprogramming and thus RT should reflect the effects of both factors. For example, Klapp (1977) used this partial precue approach to demonstrate that programming of the duration of a response can be completed even when the specific effector that will execute the response is not yet known. With respect to the phonology vs. motor distinction, one possibility would be to contrast syllable structure complexity (phonological variable) with pitch level (motor variable). If syllable structure is created during prosodification (as part of phonological encoding) and pitch level is specified at a subsequent motor programming level, then precuing of syllable structure but not pitch level should produce only pitch effects on RT (i.e. syllable structure can be preprogrammed in the absence of pitch specifications). However, if pitch level can only be specified once a syllable program has been activated, then precuing pitch level should not reduce the pitch effect on RT.

One critical issue that permeates all areas of the research program, and the study of cognition in general, is the size and content of the units at each level of processing. It is generally assumed that practice leads to formation of larger units or “chunks” (e.g., Rhodes et al., 2004; Klapp, 1995; Verwey, 1999; Wright et al., 2004),



but the size and content of such chunks in speech remains a matter of debate (e.g., Varley et al., 2005). The application of motor learning principles that have been shown to facilitate chunking (e.g., random practice, Wright et al., 2004) provides an interesting approach to this issue. In particular, by examining patterns of transfer of learning, we should be able to determine what exactly becomes specified in such units, in other words, how generalized the motor programs are.

With respect to the nature of motor speech disorders, future work should address the neural underpinnings of speech motor impairment. A recent model of speech production makes specific predictions about the contributions of different neural regions in the control of speech (Guenther, Ghosh, & Tourville, 2006; Guenther, Hampson, & Johnson, 1998), and this model may have great utility in understanding various motor speech disorders and in improving assessment and intervention. Future work will also extend the application of the reaction time methodology to study a range of clinical populations, to determine its potential for contributing to differential diagnosis and detailed assessment of underlying processes.

Systematic, model-driven investigation of speech production at the post-lexical stages of processing will lead to a better understanding of speech motor control and speech motor learning, as well as to a deeper understanding of motor speech disorders. It is my hope that ultimately, knowledge of the intricate processes involved in speech production will translate into more refined assessment methods, as well as into more effective interventions for motor speech disorders, in order to improve the quality of life for individuals with a motor speech disorder and those close to them.

**Table A-1.** Results from the Apraxia Battery for Adults – 2 (ABA-2; Dabul, 2000), the Auditory Word Discrimination subtest from the Psycholinguistic Assessment of Language Processing in Aphasia (PALPA; Kay, Lesser, & Coltheart, 1992), and the Boston Naming Test (BNT, Kaplan, Goodglass, & Weintraub, 1983).

	AOS1		AOS2		AOS3		AOS4		APH1		APH2		APH3		APH4	
	<i>Score</i>	<i>Sev.</i>	<i>Score</i>	<i>Sev.</i>	<i>Score</i>	<i>Sev.</i>	<i>Score</i>	<i>Sev.</i>	<i>Score</i>	<i>Sev.</i>	<i>Score</i>	<i>Sev.</i>	<i>Score</i>	<i>Sev.</i>	<i>Score</i>	<i>Sev.</i>
<b>ABA-2</b>																
1	12	Mi	13	Mi	10	Mi	18	Mi	28	N	19	Mi	43	N		
2A	9	S	4	Mi	5	Mod	9	S	0	N	6	Mod	0	N		
2B	4	Mod	9	S	9	S	1	N	0	N	7	S	0	N		
3A	50	N	50	N	49	N	48	N	50	N	49	N	50	N	50	N
3B	50	N	45	N	47	N	49	N	42	Mi	41	Mi	50	N	48	N
4	12	N	37	Mi	10	N	16	Mi	10	N	10	N	10	N		
5	6	Mod	21	Mi	17	Mi	19	Mi	28	N	30	N	30	N		
6	9		7		5		9		4		4		0			
<b>PALPA</b>	36/36		36/36		36/36		36/36		36/36		36/36		36/36		36/36	
<b>BNT</b>	58/60	90-100	-		55/60	80-90	58/60	90-100	-		13/15	80-90	-		12/15	

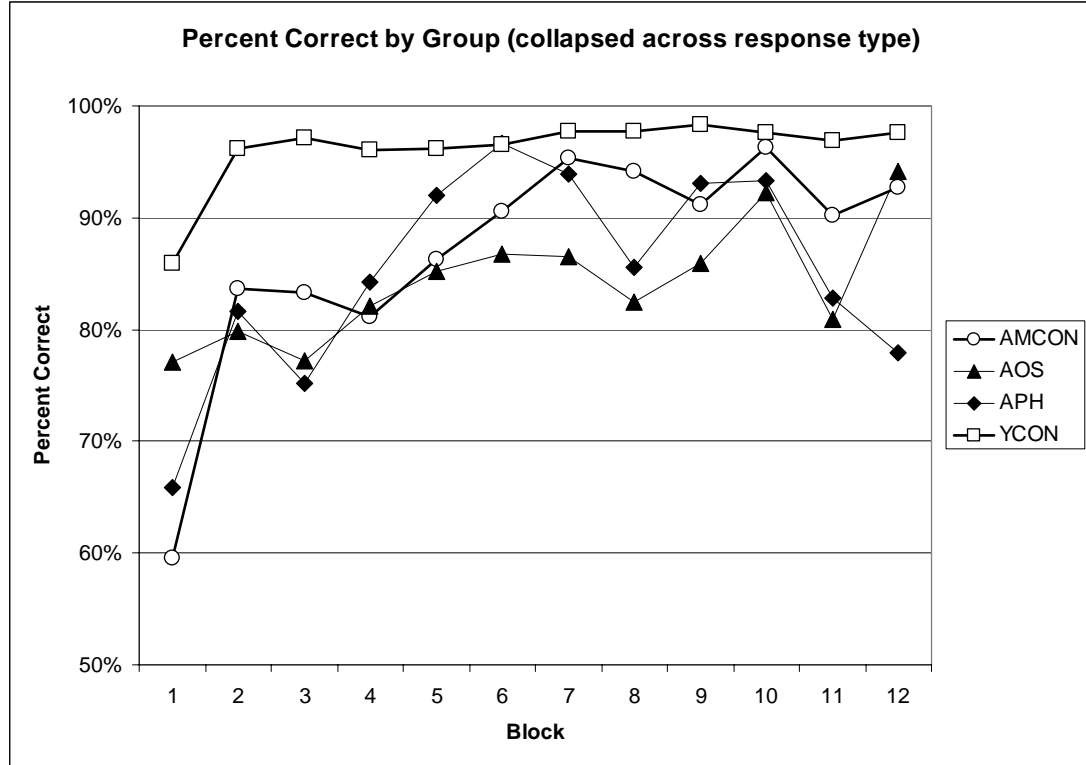
Legend: ABA-2 subtests: 1 = diadochokinetic rate; 2A and 2B = increasing word length; 3A = limb apraxia; 3B = oral apraxia; 4 = utterance time for polysyllabic words; 5 = repeated trials; 6 = inventory of articulation characteristics (5 or more suggests AOS); Sev. = severity; N = None; Mi = Mild; Mod = Moderate; S = Severe.

**Table A-2.** Results from the Western Aphasia Battery (WAB; Kertesz, 1982) and the Boston Diagnostic Aphasia Examination (BDAE, 3<sup>rd</sup> edition; Goodglass, Kaplan & Barresi, 2001).

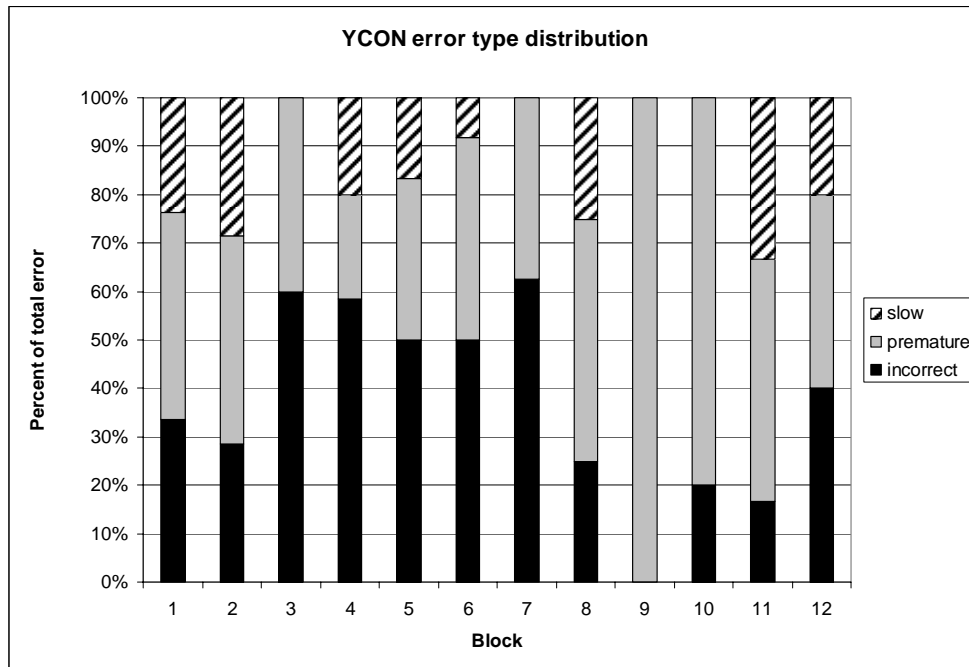
<b>WAB</b>	<b>AOS2</b>	<b>AOS3</b>	<b>AOS4</b>	<b>APH1</b>	<b>APH3</b>	<b>APH4</b>	<b>BDAE (3<sup>rd</sup> edition)</b>	<b>AOS1</b>		<b>APH2</b>	
Aphasia Quotient (AQ)	<b>69.3</b>	<b>95.5</b>	<b>97.2</b>	<b>89.6</b>	<b>96.9</b>		Competency	<b>81</b>			
<b>Spontaneous Speech</b>	12	19	19	18	19		<b>Fluency</b>	<i>score</i>	<i>%ile</i>	<i>score</i>	<i>%ile</i>
Information Content	8	10	10	9	10		Phrase length	7/7	100	7/7	100
Fluency	4	9	9	9	9		Melodic line	5/7	40	6/7	60
<b>Comprehension</b>	153	197	200	186	197	192	Grammatical form	6/7	70	5/7	50
Yes/no Questions	54	57	60	60	57	60	Articulatory agility	3/7	20	6/7	70
Aud. Word Recogn.	45	60	60	59	60	60	<b>Auditory Comp.</b>				
Sequential Commands	54	80	80	67	80	72	Word comprehension	37/37	100	15.5/16	70
<b>Repetition</b>	72	97	98	86	96		Commands	14/15	60	10/10	100
<b>Naming</b>	78	92	98	89	100		Complex Material	10/12	80	6/6	100
Object Naming	56	60	60	60	60		<b>Repetition</b>				
Word Fluency	10	12	18	9	20	20	Single words	10/10	100	4/5	60
Sentence Completion	4	10	10	10	10	10	Sentences	10/10	100	2/2	100
Responsive Speech	8	10	10	10	10	10	<b>Naming</b>				
<b>Reading</b>	-	100	100	-	100	98	Responsive naming	20/20	100	9/10	70
Reading comp/Sentences	-	40	40	40	40	40	Special categories	12/12	100	12/12	100
Reading commands	-	20	20	-	20	19	<b>Reading</b>				
Word - Object matching	-	-	-	-	-	-	Matching cases-scripts	8/8	100	4/4	100
Word - Picture matching	-	-	-	-	-	-	Number matching	12/12	100	4/4	100
Picture - word matching	-	-	-	-	-	-	Picture-word matching	10/10	100	3/4	40
Spoken - written word	-	-	-	-	-	-	Word reading	30/30	100	15/15	100
Letter Discrimination	-	-	-	-	-	-	Sent. reading	-	-	2/5	60
Spelled Word Recogn.	-	-	-	-	-	-	Sent. read./comp	9/10	80	3/3	100
Spelling	-	-	-	-	-	-	Comp. paragraphs	10/10	100	4/4	100
<b>Writing</b>	-	-	87	-	93	95	<b>Writing</b>				
Writing on Request	-	-	6	3	6	6	Dictated words	16/16	100	3/4	30
Written Output	-	-	29	33	32	32	Written pic. naming	11/12	90	1/4	30
Writing to Dictation	-	-	8.5	-	8.5	9.5	Narrative writing	10/11	90	9/11	80

**Appendix B. Supplementary Data for Experiment 1 (Speech).**

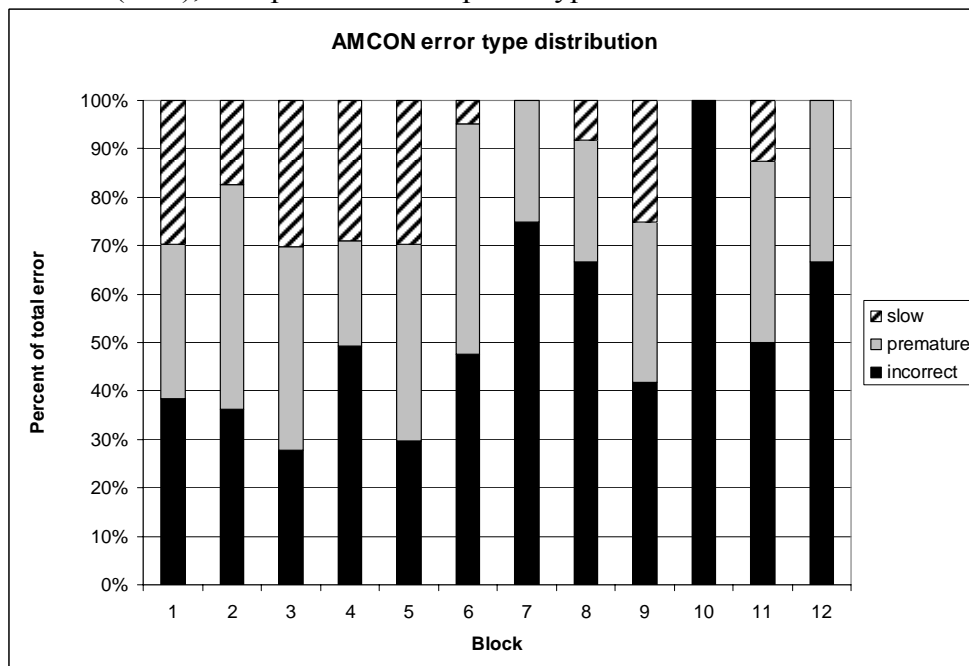
**Graph B-1.** Total percentage correct by group and block, collapsed across response type (note that y-axis starts at 50%).



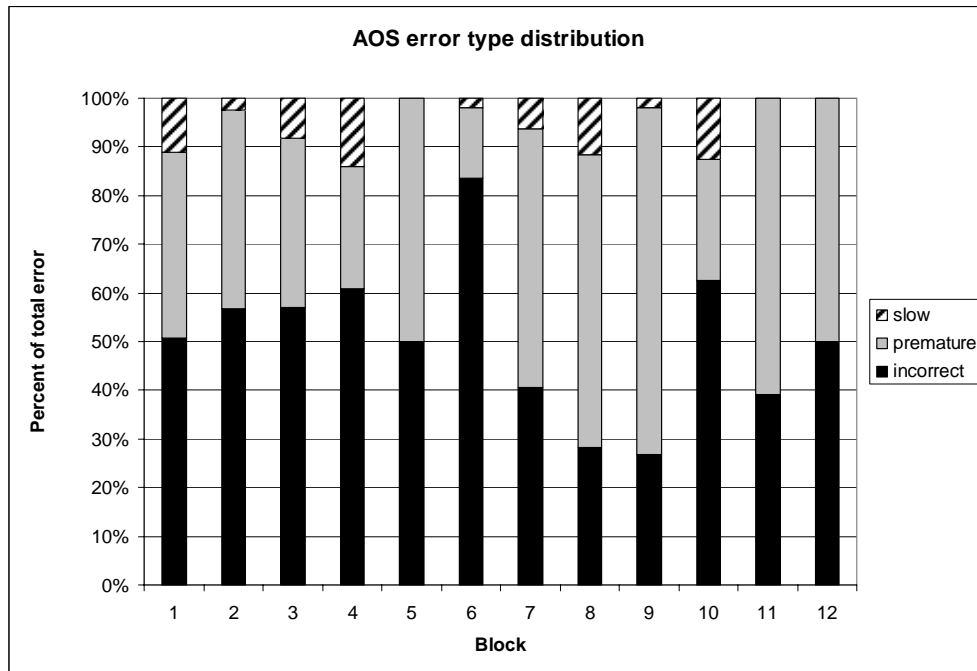
**Graph B-2.** Distribution of error types as a percentage of all errors, for younger controls (N=12), collapsed across response types.



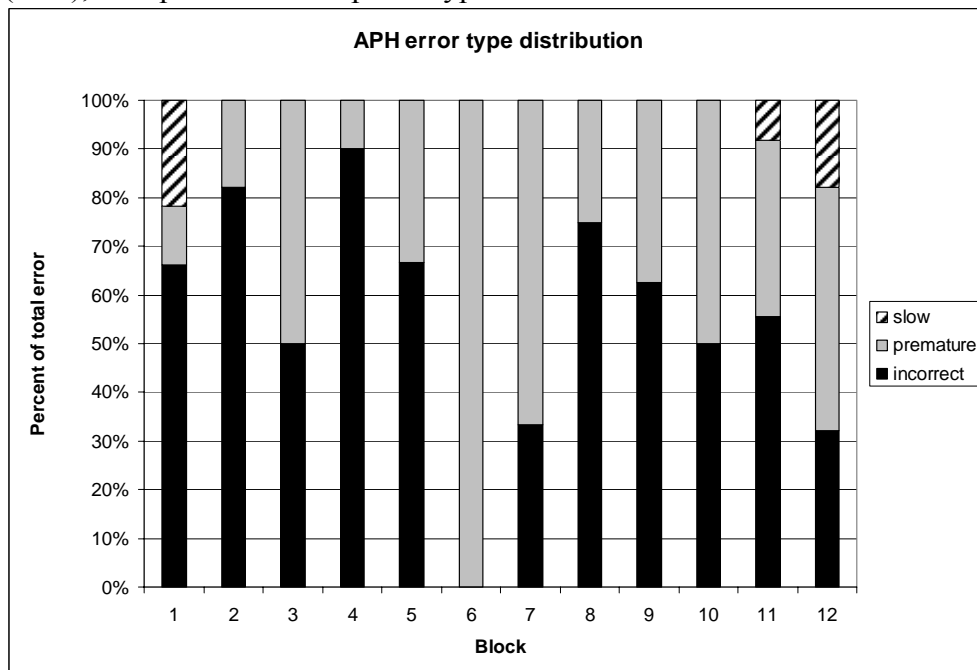
**Graph B-3.** Distribution of error types as a percentage of all errors, for age-matched controls (N=5), collapsed across response types.



**Graph B-5.** Distribution of error types as a percentage of all errors, for AOS group (N=4), collapsed across response types.



**Graph B-5.** Distribution of error types as a percentage of all errors, for APH group (N=3), collapsed across response types.



**Table B-1.** Percent correct by response type (collapsed across blocks) for each patient against age-matched control group (N=5). Standard deviations (SD) based on means of 12 practice blocks.

Patient	Response	Correct (SD)	<i>t</i> -value	<i>p</i> -value
AOS1	1L	91.1 (13.7)	- 0.054	0.480
	1S	93.3 (9.8)	0.048	0.482
	4L	91.1 (13.7)	0.191	0.429
	4S	50 (23)	- 0.913	0.206
AOS2	1L	92.2 (12)	0.012	0.496
	1S	95 (9)	0.186	0.431
	4L	86.2 (18.9)	- 0.077	0.471
	4S	49.7 (33.7)	- 0.923	0.204
AOS3	1L	94.6 (9.9)	0.155	0.442
	1S	87.8 (13.6)	- 0.396	0.356
	4L	79.8 (19.8)	- 0.426	0.346
	4S	90.3 (15.4)	0.513	0.317
AOS4	1L	88.3 (15.1)	- 0.221	0.418
	1S	88.9 (12.2)	- 0.307	0.387
	4L	86.1 (13.2)	- 0.082	0.469
	4S	82.8 (14.8)	0.248	0.408
APH1	1L	85 (14.2)	- 0.418	0.349
	1S	85.3 (14.7)	- 0.598	0.291
	4L	87.5 (16.5)	- 0.005	0.498
	4S	68.1 (24)	- 0.272	0.399
APH2	1L	98.3 (5.8)	0.376	0.363
	1S	100 (0)	0.590	0.294
	4L	95.6 (10.8)	0.437	0.342
	4S	60.3 (21.1)	- 0.548	0.306
APH3	1L	92 (16.9)	0.000	0.500
	1S	95.6 (10.8)	0.234	0.413
	4L	92.3 (18.1)	0.257	0.405
	4S	62.3 (31.6)	- 0.488	0.325
APH4 <sup>a)</sup>	1L	100 (0)	0.477	0.329
	1S	90 (11)	- 0.218	0.419
	4L	72.8 (33.4)	- 0.809	0.232
	4S	34.1 (37.4)	- 1.475	0.107
AMCON (N=5)	1L	92.0 (15.3)		
	1S	92.7 (11.3)		
	4L	87.6(16.7)		
	4S	75.8 (25.8)		
YCON (N=13)	1L	98.0 (7.3)		
	1S	98.1 (6.9)		
	4L	97.3 (7.5)		
	4S	91.2 (16.2)		

<sup>a)</sup> Data for APH4 are based on 6 acquisition blocks only.

**Table B-2.** Syllable duration for single syllables for each patient (collapsed across blocks), compared to young controls (N=8).

Subject	Duration	Acquisition	t-value	p-value	Retention	t-value	p-value
AOS1	1L	484 (42)	0.943	0.189		N/A	
	1S	185 (43)	0.459	0.330			
			<i>0.951</i>	<i>0.373</i>			
AOS2	1L	568 (64)*	1.921	0.048		N/A	
	1S	195 (109)	0.713	0.249			
			<i>2.304^^</i>	<i>0.055</i>			
AOS3	1L	487 (145)	0.978	0.180	379	0.108	0.459
	1S	191 (29)	0.612	0.280	224	1.033	0.168
			<i>0.722</i>	<i>0.494</i>		<i>1.026</i>	<i>0.339</i>
AOS4	1L	512 (28)	1.269	0.123	453	0.772	0.233
	1S	178 (61)	0.280	0.394	107	- 1.088	0.156
			<i>1.907^^</i>	<i>0.098</i>		<i>2.027^^</i>	<i>0.082</i>
CON	1L	403 (81)			367 (105)		
(N=8)	1S	167 (37)			167 (52)		

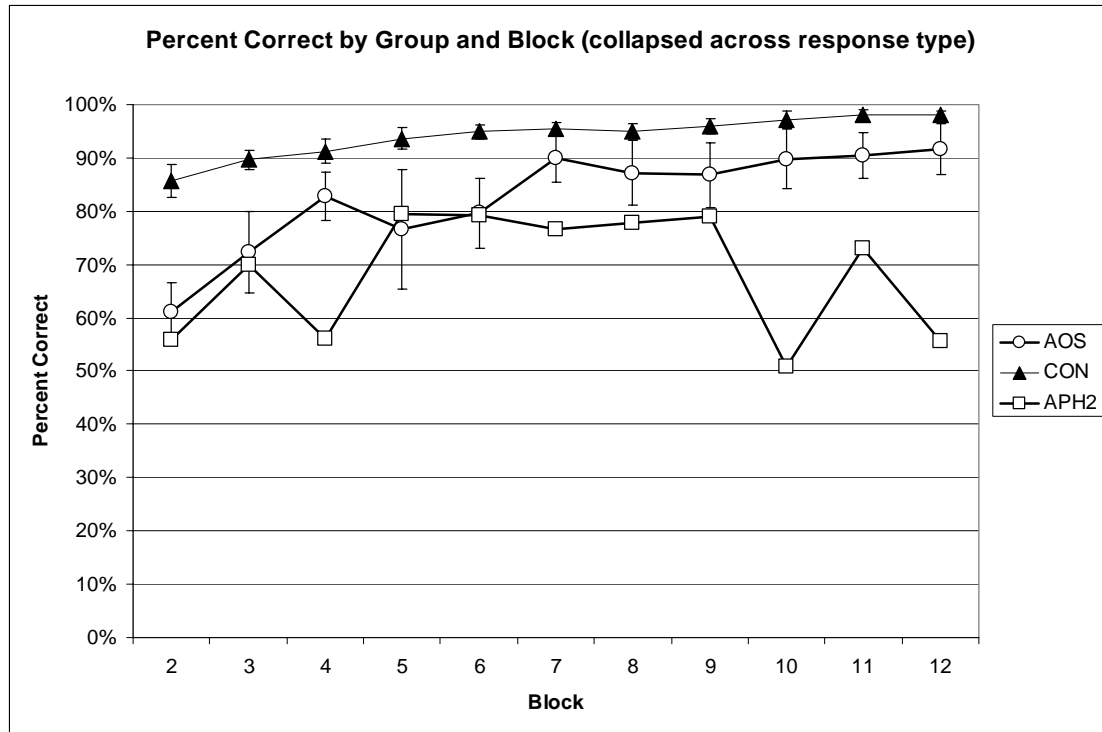
\*  $p < .05$ ,  $df=7$ , one-tailed (Crawford & Howell, 1998).

^^  $p < .05$ ,  $df=7$ , two-tailed (Crawford & Garthwaite, 2005).



**APPENDIX C. Supplementary data for Experiment 2 (Finger movements).**

**Graph C-1.** Accuracy (percent correct) by block for Experiment 2 (collapsed across response types), for controls (CON, N=13), patients with apraxia of speech (AOS, N=4), and one patient with aphasia (APH).



**Table C-1.** Percent correct by response type (collapsed across blocks) for each patient against the control group (N=13). Standard deviations (SD) based on means of 11 practice blocks.

Patient	Response	Percent correct (SD)	<i>t</i> -value	<i>p</i> -value
AOS1	1L	76.2 (20.3)	- 1.054	0.156
	1S	73.3 (14.1)	- 2.427	0.016*
	4L	73.7 (31.6)	- 1.065	0.154
	4S	46.5 (20.7)	- 2.090	0.029*
AOS2	1L	77.3 (16.7)	- 0.976	0.174
	1S	86.7 (17.8)	- 1.034	0.161
	4L	79.3 (32.2)	- 0.721	0.242
	4S	82.1 (27.1)	- 0.306	0.383
AOS3	1L	72.4 (19.7)	- 1.322	0.105
	1S	72.4 (19.7)	- 2.524	0.013*
	4L	77.8 (26.5)	- 0.814	0.216
	4S	82.9 (27.2)	- 0.262	0.399
AOS4	1L	81.2 (29.3)	- 0.539	0.300
	1S	91.5 (23.8)	- 0.440	0.333
	4L	88.3 (20.7)	- 0.165	0.436
	4S	92.8 (13.5)	0.232	0.410
APH2	1L	80.0 (22.1)	- 0.779	0.226
	1S	92.8 (13.5)	- 0.408	0.345
	4L	65.4 (29.7)	- 1.575	0.071 <sup>^</sup>
	4S	21.8 (16.3)	- 3.333	0.003*
CONTROLS (N=13)	1L	91.0 (13.6)		
	1S	96.7 (9.3)		
	4L	91.0 (15.7)		
	4S	88.2 (19.2)		

\*  $p < .05$ ,  $df=12$ , one-tailed (Crawford & Howell, 1998).

<sup>^</sup>  $p < .10$ ,  $df=12$ , one-tailed (Crawford & Howell, 1998).

## References

- Abbs, J.H., Gracco, V.L. & Cole, K.J. (1984). Control of multimovement coordination: sensorimotor mechanisms in speech motor programming. *Journal of Motor Behavior*, *16*(2), 195-231.
- Aichert, I., & Ziegler, W. (2004). Syllable frequency and syllable structure in apraxia of speech. *Brain and Language*, *88*, 148-159.
- Austermann Hula, S.N., Robin, D.A., Maas, E., Ballard, K.J., & Schmidt, R.A. (submitted). Effects of feedback manipulations in treatment for apraxia of speech. *Journal of Speech, Language, and Hearing Research*.
- Baddeley, A.D. (1986). *Working Memory*. Oxford, England: Clarendon Press.
- Ballard, K.J. (2001). Response generalization in apraxia of speech treatments: taking another look. *Journal of Communication Disorders*, *34*, 3-20.
- Ballard, K.J., Barlow, J.A., & Robin, D.A. (2001). The underlying nature of apraxia of speech: A critical evaluation of Varley and Whiteside's dual route speech encoding hypothesis. *Aphasiology*, *15*(1), 50-58.
- 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. (in press). Treating control of voicing in apraxia of speech with variable practice. *Aphasiology*.
- Ballard, K.J., & Robin, D.A. (in press). Influence of continual biofeedback on jaw pursuit tracking in healthy adults and adults with apraxia plus aphasia. *Journal of Motor Behavior*.
- Ballard, K.J., Robin, D.A., & Folkins, J.W. (2003). An integrative model of speech motor control: a response to Ziegler. *Aphasiology*, *17*(1), 37-48.
- Bates, E., Appelbaum, M., Salcedo, J., Saygin, A.P., & Pizzamiglio, L. (2003). Quantifying dissociations in neuropsychological research. *Journal of Clinical and Experimental Neuropsychology*, *25*(8), 1123-1158.
- Bohland, J.W., & Guenther, F.H. (in press). An fMRI investigation of syllable sequence production. *NeuroImage*.
- Boliek, C.A., Hixon, T.J., Watson, P.J., & Morgan, W.J. (1996). Vocalization during breathing during the first year of life. *Journal of Voice*, *10*, 1-22.

- Boliek, C.A., Hixon, T.J., Watson, P.J., & Morgan, W.J. (1997). Vocalization during breathing during the second and third years of life. *Journal of Voice, 11*, 373-390.
- Bullock, D. (2004). Adaptive neural models of queuing and timing in fluent action. *Trends in Cognitive Sciences, 8*(9), 426-433.
- Canic, M.J., & Franks, I.M. (1989). Response preparation and latency in patterns of tapping movements. *Human Movement Science, 8*, 123-139.
- Carter, M.C., & Shapiro, D.C. (1984). Control of sequential movements: Evidence for generalized motor programs. *Journal of Neurophysiology, 52*(5), 787-796.
- Cholin, J., Levelt, W.J.M., & Schiller, N.O. (2005). Effects of syllable frequency in speech production. *Cognition, 99*(2), 205-235.
- 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.
- Clark, H.M., Robin, D.A., McCullagh, G., & Schmidt, R.A. (2001). Motor control in children and adults during a non-speech oral task. *Journal of Speech, Language, and Hearing Research, 44*, 1015-1025.
- Code, C. (1998). Models, theories and heuristics in apraxia of speech. *Clinical Linguistics and Phonetics, 12*(1), 47-65.
- Crawford, J.R., & Garthwaite, P.H. (2005). Testing for suspected impairments and dissociations in single-case studies in neuropsychology: Evaluation of alternatives using Monte Carlo simulations and revised tests for dissociations. *Neuropsychology, 19*(3), 318-331.
- Crawford, J.R., Garthwaite, P.H., Azzalini, A., Howell, D.C., & Laws, K.R. (in press). Testing for a deficit in single-case studies: Effects of departures from normality. *Neuropsychologia*.
- Crawford, J.R., Garthwaite, P.H., & Gray, C.D. (2003). Wanted: Fully operational definitions of dissociations in single-case studies. *Cortex, 39*, 357-370.
- Crawford, J.R., & Howell, D.C. (1998). Comparing an individual's test score against norms derived from small samples. *Clinical Neuropsychologist, 12*, 482-486.

- Croot, K. (2001). Integrating the investigation of apraxic, aphasic and articulatory disorders in speech production: A move towards sound theory. *Aphasiology*, 15(1), 58-62.
- Dabul, B.L. (2000). *Apraxia Battery for Adults - 2*. Austin, TX: Pro-Ed.
- Darley, F.L., Aronson, A.E., & Brown, J.R. (1975). *Motor Speech Disorders*. Philadelphia: W.B. Saunders, Inc.
- Deger, K., & Ziegler, W. (2002). Speech motor programming in apraxia of speech. *Journal of Phonetics*, 30, 321-335.
- Dell, G.S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, 93(3), 283-321.
- Den Ouden, D.B. (2002). *Phonology in Aphasia: Syllables and Segments in Level-Specific Deficits*. Doctoral dissertation, University of Groningen, The Netherlands.
- Dronkers, N.F. (1996). A new brain region for coordinating speech articulation. *Nature*, 384, 159-161.
- Duffy, J. (2005). *Motor Speech Disorders: Substrates, Differential Diagnosis, and Management (2<sup>nd</sup> edition)*. St. Louis, MO: Mosby-Year Book, Inc.
- Duffy, J.R., Peach, R., & Strand, E.A. (2006). Progressive apraxia of speech as a sign of motor neuron disease. Paper presented at the 13<sup>th</sup> Biennial Conference on Motor Speech: Motor Speech Disorders, Austin, TX, March 23-26, 2006.
- Folkins, J.W., Moon, J.B., Luschei, E.S., Robin, D.A., Tye-Murray, N., & Moll, K.L. (1995). What can nonspeech tasks tell us about speech motor disabilities? *Journal of Phonetics*, 23, 139-147.
- Friedmann, N., & Gvion, A. (2006). Is there a relation between working memory limitation and sentence comprehension? A study of conduction and agrammatic aphasia. Paper presented at the 36<sup>th</sup> Clinical Aphasiology Conference, Ghent, Belgium, May 29 – June 2, 2006.
- Garrett, M.F. (1982). Production of speech: observations from normal and pathological language use. In A.W. Ellis (Ed.), *Normality and Pathology in Cognitive Functions*. London: Academic Press. (pp. 19-76).

- Gerken, L., & McGregor, K.K. (1998). An overview of prosody and its role in normal and disordered child language. *American Journal of Speech-Language Pathology, 7*, 38-48.
- Goodglass, H., Kaplan, E., & Barresi, B. (2001). *Boston Diagnostic Aphasia Examination (3<sup>rd</sup> edition)*. New York: Lippencott, Williams, and Williams.
- Gracco, V.L., & Abbs, J.H. (1986). Variant and invariant characteristics of speech movements. *Experimental Brain Research, 65*, 156-166.
- 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(3)*, 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.
- Hageman, C., Robin, D.A., Moon, J.B., & Folkins, J.W. (1994). Visuomotor tracking abilities of speakers with apraxia. *Clinical Aphasiology, 22*, 219-229.
- Henry, F.M., & Rogers, D.E. (1960). Increased response latency for complicated movements and a "memory drum" theory of neuromotor reaction. *Research Quarterly, 31*, 448-458.
- Heuer, H., Schmidt, R.A., & Ghodsian, D. (1995). Generalized motor programs for rapid bimanual tasks: a two-level multiplicative-rate model. *Biological Cybernetics, 73*, 343-356.
- Hillis, A.E., Work, M., Barker, P.B., Jacobs, M.A., Breese, E.L., & Maurer, K. (2004). Re-examining the brain regions crucial for orchestrating speech articulation. *Brain, 127*, 1479-1487.
- Immink, M.A., & Wright, D.L. (2001). Motor programming during practice conditions high and low in contextual interference. *Journal of Experimental Psychology: Human Perception and Performance, 27(2)*, 423-437.
- Itoh, M., Sasanuma, S., & Ushijima, T. (1979). Velar movements during speech in a patient with apraxia of speech. *Brain and Language, 7*, 227-239.
- Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *Boston Naming Test*. Philadelphia, PA: Lea & Febiger.

- Kay, J., Lesser, R., & Coltheart, M. (1992). *Psycholinguistic Assessment of Language Processing in Aphasia*. Hove, UK: Lawrence Erlbaum Associates.
- Keele, S.W. (1968). Movement control in skilled motor performance. *Psychological Bulletin*, *70*, 387-403.
- Keele, S.W., Cohen, A., & Ivry, R. (1990). Motor programs: Concepts and issues. In M. Jeannerod (Ed.), *Attention and Performance XIII: Motor Representation and Control*. Hillsdale, NJ: Lawrence Erlbaum Associates. Chapter 3 (pp. 77-110).
- Keele, S.W., Jennings, P., Jones, S., Caulton, D., & Cohen, A. (1995). On the modularity of sequence representation. *Journal of Motor Behavior*, *27*(1), 17-30.
- Kelso, J.A.S., Saltzman, E.L., & Tuller, B. (1986). The dynamical perspective on speech production: Data and theory. *Journal of Phonetics*, *14*, 29-59.
- Kelso, J.A.S., & Tuller, B.H. (1981). Toward a theory of apractic syndromes. *Brain & Language*, *12*, 224-245.
- Kent, R.D., Adams, S.G., & Turner, G.S. (1996). Models of speech production. In N.J. Lass (Ed.), *Principles of Experimental Phonetics*. St. Louis, MO: Mosby. Chapter 1 (pp. 3-45).
- Kent, R.D., Duffy, J.R., Slama, A., Kent, J.F., & Clift, A. (2001). Clinicoanatomic studies in dysarthria: review, critique, and directions for research. *Journal of Speech, Language, and Hearing Research*, *44*, 535-551.
- Kent, R.D., Kent, J.F., Weismer, G., & Duffy, J.R. (2000). What dysarthrias can tell us about the neural control of speech. *Journal of Phonetics*, *28*, 273-302.
- Kent, R.D., & Rosenbek, J.C. (1983). Acoustic patterns of apraxia of speech. *Journal of Speech and Hearing Research*, *26*, 231-249.
- Kerr, B. (1978). Task factors that influence selection and preparation for voluntary movements. In G.E. Stelmach (Ed.), *Information Processing in Motor Control and Learning*. New York: Academic Press. Chapter 3 (pp. 55-69).
- Kertesz, A. (1982). *Western Aphasia Battery*. New York: Grune & Stratton.
- Khan, M.A., Lawrence, G.P., Buckolz, E., & Franks, I.M. (2006). Programming strategies for rapid aiming movements under simple and choice reaction time conditions. *Quarterly Journal of Experimental Psychology*, *59*(3), 524-542.

- Klapp, S.T. (1974). Syllable-dependent pronunciation latencies in number naming: A replication. *Journal of Experimental Psychology*, *102*(6), 1138-1140.
- Klapp, S.T. (1977). Response programming, as assessed by reaction time, does not establish commands for particular muscles. *Journal of Motor Behavior*, *9*(4), 301-312.
- Klapp, S.T. (1981). Motor programming is not the only process which can influence RT: Some thoughts on the Marteniuk and MacKenzie analysis. *Journal of Motor Behavior*, *13*(4), 320-328.
- Klapp, S.T. (1995). Motor response programming during simple and choice reaction time: the role of practice. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(5), 1015-1027.
- 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), 135-150.
- Klapp, S.T., Abbott, J., Coffman, K., Greim, D., Snider, R., & Young, F. (1979). Simple and choice reaction time methods in the study of motor programming. *Journal of Motor Behavior*, *11*(2), 91-101.
- Klapp, S.T., Anderson, W.G., & Berrian, R.W. (1973). Implicit speech in reading, reconsidered. *Journal of Experimental Psychology*, *100*(2), 368-374.
- Klapp, S.T., & Erwin, C.I. (1976). Relation between programming time and duration of the response being programmed. *Journal of Experimental Psychology: Human Perception and Performance*, *2*(4), 591-598.
- Klapp, S.T., & Greim, D.M. (1981). Technical considerations regarding the short (dit) - long (dah) key press paradigm. *Journal of Motor Behavior*, *13*(1), 1-8.
- Klapp, S.T., & Wyatt, E.P. (1976). Motor programming within a sequence of responses. *Journal of Motor Behavior*, *8*(1), 19-26.
- Lai, Q., Shea, C.H., Wulf, G., & Wright, D.L. (2000). Optimizing generalized motor program and parameter learning. *Research Quarterly for Exercise and Sport*, *71*, 10-24.
- Lashley, K.S. (1951). The problem of serial order in behavior. In L.A. Jeffress (Ed.), *Cerebral Mechanisms in Behavior. The Hixon Symposium*. New York: John Wiley & Sons, Inc. (pp. 112-146).



- Levelt, W.J.M. (1989). *Speaking: From Intention to Articulation*. Cambridge, MA: MIT Press.
- Levelt, W.J.M. (1999). Models of word production. *Trends in Cognitive Sciences*, 3(6), 223-232.
- Levelt, W.J.M., Roelofs, A., & Meyer, A.S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22, 1-75.
- Levelt, W.J.M., & Wheeldon, L. Do speakers have access to a mental syllabary? *Cognition*, 50, 239-269.
- Maas, E., Barlow, J., Robin, D., & Shapiro, L. (2002). Treatment of sound errors in aphasia and apraxia of speech: Effects of phonological complexity. *Aphasiology*, 16 (7/8/9), 609-622.
- Maas, E., Robin, D.A., Austermann Hula, S., Freedman, S.E., Wulf, G., Ballard, K.J., & Schmidt, R.A. (in preparation). Principles of motor learning in treatment of motor speech disorders. *American Journal of Speech-Language Pathology*.
- MacNeilage, P.F. (1998). The frame/content theory of evolution of speech production. *Behavioral and Brain Sciences*, 21, 499-546.
- Magnuson, C.E., Robin, D.A., & Wright, D.L. (in press). Sequencing multiple elements of the same duration: Detailing the INT process in Klapp's two-process account of motor programming. *Quarterly Journal of Experimental Psychology*.
- Marteniuk, R.G., & MacKenzie, C.L. (1981). Methods in the study of motor programming: Is it just a matter of simple vs. choice reaction time? A comment on Klapp et al. (1979). *Journal of Motor Behavior*, 13(4), 313-319.
- Martin, A.D. (1974). Some objections to the term apraxia of speech. *Journal of Speech and Hearing Disorders*, 39(1), 53-64.
- Max, L., & Gracco, V.L. (2005). Coordination of oral and laryngeal movements in the perceptually fluent speech of adults who stutter. *Journal of Speech, Language, and Hearing Research*, 48, 524-542.
- McNeil, M.R., Caliguiri, M., & Rosenbek, J.C. (1989). A comparison of labiomandibular kinematic durations, displacements, velocities and dysmetrias in apraxic and normal adults. *Clinical Aphasiology*, 17, 173-193.

- McNeil, M.R., Doyle, P.J., & Wambaugh, J. (2000). Apraxia of speech: a treatable disorder of motor planning and programming. In S.E. Nadeau, L.J. Gonzalez-Rothi, & B. Crosson (Eds.), *Aphasia and Language. Theory to Practice*. New York: The Guilford Press. Chapter 9 (pp. 221-266).
- McNeil, M.R., Odell, K.H., Miller, S.B., & Hunter, L. (1995). Consistency, variability, and target approximation for successive speech repetitions among apraxic, conduction aphasic and ataxic dysarthric speakers. *Clinical Aphasiology*, 23, 39-55.
- McNeil, M.R., Robin, D.A., & Schmidt, R.A. (1997). Apraxia of speech: definition, differentiation, and treatment. In: M.R. McNeil (Ed.), *Clinical Management of Sensorimotor Speech Disorders*. New York: Thieme. Chapter 14 (pp. 311-344).
- McNeil, M.R., Weismer, G., Adams, S., & Mulligan, M. (1990). Oral structure nonspeech motor control in normal, dysarthric, aphasic and apraxic speakers: isometric force and static position control. *Journal of Speech and Hearing Research*, 33, 255-268.
- Meyer, A.S., Roelofs, A., & Levelt, W.J.M. (2003). Word length effects in object naming: the role of a response criterion. *Journal of Memory and Language*, 48, 131-147.
- Milenkovic, P. (2000). Time-Frequency Analysis for 32-bit Windows. Madison, WI (revised version October 10, 2003).
- Moore, C.A., Caulfield, T.A., & Green, J.R. (2001). Relative kinematics of the rib cage and abdomen during speech and nonspeech behaviors of 15-month old children. *Journal of Speech, Language, and Hearing Research*, 44, 80-94.
- Mowrey, R.A., & MacKay, I.R.A. (1990). Phonological primitives: electromyographic speech error evidence. *Journal of the Acoustical Society of America*, 88, 1299-1312.
- Nickels, L. (1997). *Spoken Word Production and its Breakdown in Aphasia*. Hove, UK: Psychology Press.
- Nijland, L., Maassen, B., Van der Meulen, S., Gabreëls, F., Kraaimaat, F.W., & Schreuder, R. (2003). Planning of syllables in children with developmental apraxia of speech. *Clinical Linguistics & Phonetics*, 17(1), 1-24.
- Park, J.H., & Shea, C.H. (2003). Effect of practice on effector independence. *Journal of Motor Behavior*, 35(1), 33-40.

- Park, J.H., & Shea, C.H. (2005). Sequence learning: Response structure and effector transfer. *Quarterly Journal of Experimental Psychology*, *58A*(3), 387-419.
- Rapp, B., & Goldrick, M. (2000). Discreteness and interactivity in spoken word production. *Psychological Review*, *107*(3), 460-499.
- Rapp, B., & Goldrick, M. (2004). Feedback by any other name is still interactivity: A reply to Roelofs (2002). *Psychological Review*, *111*(2), 573-578.
- 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*, 683-730.
- Robin, D.A. (1992). Developmental apraxia of speech: Just another motor problem. *American Journal of Speech-Language Pathology*, *1*, 19-22.
- Robin, D.A., Bean, C., & Folkins, J.W. (1989). Lip movement in apraxia of speech. *Journal of Speech and Hearing Research*, *32*, 512-523.
- Robin, D.A., Moon, J.B., Clark, H.M., Woodworth, G., & Folkins, J.W. (submitted). Visuomotor tracking abilities of speakers with apraxia of speech or conduction aphasia.
- Robin, D.A., Solomon, N.P., Moon, J.B., & Folkins, J.W. (1997). Nonspeech assessment of the speech production mechanism. In M.R. McNeil (Ed.), *Clinical Management of Sensorimotor Speech Disorders*. New York: Thieme. Chapter 3 (pp. 49-62).
- Roelofs, A. (2004). Error biases in spoken word planning and monitoring by aphasic and non-aphasic speakers: A comment on Rapp & Goldrick (2000). *Psychological Review*, *111*(2), 561-572.
- Roelofs, A. (2002). Syllable structure effects turn out to be word length effects: A comment on Santiago et al. (2000). *Language and Cognitive Processes*, *17*(1), 1-13.
- Roelofs, A., & Meyer, A.S. (1998). Metrical structure in planning the production of spoken words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 922-939.
- Rogers, M.A., & Spencer, K.A. (2001). Spoken word production without assembly: Is it possible? *Aphasiology*, *15*(1), 68-74.

- Rogers, M.A., & Storkel, H.L. (1998). Reprogramming phonologically similar utterances: The role of phonetic features in pre-motor encoding. *Journal of Speech, Language, and Hearing Research*, 41, 258-274.
- 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(9/10/11), 793-805.
- Rosenbek, J.C. (2001). Darley and apraxia of speech in adults. *Aphasiology*, 15(3), 261-273.
- Sakai, K., Hikosaka, O., & Nakamura, K. (2004). The emergence of rhythm during motor learning. *Trends in Cognitive Science*, 8 (12), 547-553.
- Santiago, J., MacKay, D.G., & Palma, A. (2002). Length effects turn out to be structure effects: Response to Roelofs (2002). *Language and Cognitive Processes*, 17(1), 15-29.
- Santiago, J., MacKay, D.G., Palma, A., & Rho, C. (2000). Sequential activation processes in producing words and syllables: Evidence from picture naming. *Language and Cognitive Processes*, 15(1), 1-44.
- Schiller, N.O., Fikkert, P., & Levelt, C.C. (2004). Stress priming in picture naming: an SOA study. *Brain and Language*, 90(1-3), 231-240.
- Schmidt, R.A. (1975). A schema theory of discrete motor skill learning. *Psychological Review*, 82(4), 225-260.
- Schmidt, R.A. (1976). The schema as a solution to some persistent problems in motor learning theory. In G.L. Stelmach (Ed.), *Motor Control. Issues and Trends*. New York: Academic Press. Chapter 2 (pp. 41-65).
- Schmidt, R.A. (2003). Motor schema theory after 27 years: reflections and implications for a new theory. *Research Quarterly for Exercise and Sport*, 74(4), 366-375.
- Schmidt, R.A., & Lee, T.D. (2005). *Motor Control and Learning: A Behavioral Emphasis (4<sup>th</sup> edition)*. Champaign, IL: Human Kinetics.
- Schneider, D.M., & Schmidt, R.A. (1995). Units of action in motor control: Role of response complexity and target speed. *Human Performance*, 8(1), 27-49.

- Schneider, S.L., & Frens, R.A. (2005). Training four-syllable CV patterns in individuals with acquired apraxia of speech: Theoretical implications. *Aphasiology, 19*(3/4/5), 451-471.
- Schönle, P.W., Hong, G., Benecke, R., & Conrad, B. (1986). Aspects of speech motor control: programming of repetitive versus non-repetitive speech. *Neuroscience Letters, 63*, 170-174.
- Seddoh, S.A.K., Robin, D.A., Sim, H.S., Hageman, C., Moon, J.B., & Folkins, J.W. (1996). Speech timing in apraxia of speech versus conduction aphasia. *Journal of Speech and Hearing Research, 39*, 590-603.
- Semjen, A., & Gottsdanker, R. (1990). Rapid serial movements: Relation between the planning of sequential structure and effector selection. In M. Jeannerod (Ed.), *Attention and Performance XIII: Motor Representation and Control*. Hillsdale, NJ: Lawrence Erlbaum Associates. Chapter 13 (pp. 409-427).
- Shaiman, S., McNeil, M.R., Szuminsky, N.J., Meigh, K., & Botler, J.B. (2006). Motor learning of volitional nonspeech oral movements: Intraoral pressure and articulatory kinematics. Poster presented at the 13<sup>th</sup> Biennial Conference on Motor Speech: Speech Motor Control, Austin, TX, March 23-26, 2006.
- Shattuck-Hufnagel, S. (1979). Speech errors as evidence for a serial order mechanism in sentence production. In W.E. Cooper, & E.C.T. Walker (Eds.), *Sentence Processing: Psycholinguistic Studies presented to Merrill Garrett*. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers. Chapter 10 (pp. 295-342).
- Shattuck-Hufnagel, S. (1992). The role of word structure in segmental serial ordering. *Cognition, 42*, 213-259.
- Shea, C.H., Lai, Q., Wright, D.L., Immink, M., & Black, C. (2001). Consistent and variable practice conditions: effects on relative and absolute timing. *Journal of Motor Behavior, 33*(2), 139-152.
- Shea, C.H., & Park, J.H. (2003). The independence of response structure and element production in timing sequences. *Research Quarterly for Exercise and Sport, 74*(4), 401-420.
- Shea, C.H., & Wulf, G. (2005). Schema theory: a critical appraisal and reevaluation. *Journal of Motor Behavior, 37*, 85-101.
- Shea, C.H., Wulf, G., Park, J.H., & Gaunt, B. (2001). Effects of an auditory model on the learning of relative and absolute timing. *Journal of Motor Behavior, 33*(2), 127-138.

- Sheridan, M.R. (1981). Response programming and reaction time. *Journal of Motor Behavior*, 13(3), 161-176.
- Smiley-Oyen, A.L., & Worringham, C.J. (1996). Distribution of programming in a rapid aimed sequential movement. *Quarterly Journal of Experimental Psychology*, 49A(2), 379-397.
- Smith, A., Johnson, M., McGillem, C., & Goffman, L. (2000). On the assessment of stability and patterning of speech movements. *Journal of Speech, Language, and Hearing Research*, 43, 277-286.
- Spencer, K.A., & Rogers, M.A. (2005). Speech motor programming in hypokinetic and ataxic dysarthria. *Brain and Language*, 94, 347-366.
- Sternberg, S., Monsell, S., Knoll, R.L., & Wright, C.E. (1978). The latency and duration of rapid movement sequences: Comparisons of speech and typewriting. In G.E. Stelmach (Ed.), *Information Processing in Motor Control and Learning*. New York: Academic Press. Chapter 6 (pp. 117-152).
- 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*. New York: Thieme. Chapter 1 (pp. 1-25).
- Varley, R., & Whiteside, S.P. (2001a). What is the underlying impairment in acquired apraxia of speech? *Aphasiology*, 15(1), 39-49.
- Varley, R., & Whiteside, S.P. (2001b). Reply: exploring the enigma. *Aphasiology*, 15(1), 78-84.
- Varley, R., Whiteside, S., Windsor, F., & Fisher, H. (2006). Moving up from the segment: A comment on Aichert and Ziegler's Syllable frequency and syllable structure in apraxia of speech, *Brain and Language*, 88, 148-159, 2004. *Brain and Language*, 96(2), 235-239.
- Verwey, W.B. (1999). Evidence for a multi-stage model of practice in a sequential movement task. *Journal of Experimental Psychology: Human Perception and Performance*, 25(6), 1693-1708.
- Verwey, W.B. (2003). Effect of sequence length on executing familiar keying sequences: Lasting segmentation and preparation? *Journal of Motor Behavior*, 35, 343-354.

- Wadman, W.J., Denier van der Gon, J.J., Geuze, R.H., & Mol, C.R. (1979). Control of fast goal directed arm movements. *Journal of Human Movement Studies*, 5, 3-17.
- Wambaugh, J.L., Duffy, J.R., McNeil, M.R., Robin, D.A., & Rogers, M. (2006). Treatment guidelines for acquired apraxia of speech: A synthesis and evaluation of the evidence. *Journal of Medical Speech Language Pathology*.
- Waters, G.S., Rochon, E., & Caplan, D. (1992). The role of high-level speech planning in rehearsal: Evidence from patients with apraxia of speech. *Journal of Memory and Language*, 31, 54-73.
- Weismer, G. (2006). Philosophy of research in motor speech disorders. *Clinical Linguistics & Phonetics*, 20(5), 315-349.
- Weismer, G., & Liss, J. M. (1991). Reductionism is a dead-end in speech research: perspectives on a new direction. In C. Moore, K.M. Yorkston, & D.R. Beukelman (Eds.), *Dysarthria and Apraxia of Speech: Perspectives on Management*. Baltimore: Paul H. Brookes. (pp. 15-27).
- Wertz, R.T., LaPointe, L.L., & Rosenbek, J.C. (1984). *Apraxia of Speech in Adults: The Disorder and its Management*. Orlando, FL: Grune & Stratton.
- Whiteside, S.P., & Varley, R.A. (1998). A reconceptualisation of apraxia of speech: A synthesis of evidence. *Cortex*, 34, 221-231.
- Wright, D.L., Black, C.B., Immink, M.A., Brueckner, S., & Magnuson, C. (2004). Long-term motor programming improvements occur via concatenating movement sequences during random but not blocked practice. *Journal of Motor Behavior*, 36(1), 39-50.
- 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 and Sport*, 72(1), 32-38.
- Wulf, G., & Lee, T.D. (1993). Contextual interference in movements of the same class: differential effects on program and parameter learning. *Journal of Motor Behavior*, 25(4), 254-263.
- 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. (1996). Average KR degrades parameter learning. *Journal of Motor Behavior*, 28(4), 371-381.
- Wulf, G., Schmidt, R.A., & Deubel, H. (1993). Reduced frequency feedback enhances generalized motor program learning but not parameterization learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19(5), 1134-1150.
- Young, D.E., & Schmidt, R.A. (1990). Units of motor behavior: Modifications with practice and feedback. In M. Jeannerod (Ed.), *Attention and Performance XIII: Motor Representation and Control*. Hillsdale, NJ: Lawrence Erlbaum Associates. Chapter 28 (pp. 763-795).
- Yorkston, K.M., Beukelman, D.R., Strand, E., & Bell, K.R. (1999). *Clinical Management of Motor Speech Disorders in Children and Adults (2<sup>nd</sup> edition)*. Austin, TX: Pro-Ed.
- Ziegler, W. (2001). Apraxia of speech is not a lexical disorder. *Aphasiology*, 15(1), 74-77.
- Ziegler, W. (2002). Task-related factors in oral motor control: speech and oral diadochokinesis in dysarthria and apraxia of speech. *Brain and Language*, 80, 556-575.
- Ziegler, W. (2003a). Speech motor control is task-specific. Evidence from dysarthria and apraxia of speech. *Aphasiology*, 17(1), 3-36.
- Ziegler, W. (2003b). To speak or not to speak: distinctions between speech and nonspeech motor control. A reply to Ballard, Robin & Folkins. *Aphasiology*, 17(2), 99-105.
- Ziegler, W. (2006). Perspectives on apraxia of speech. Keynote presentation at the 36<sup>th</sup> Annual Clinical Aphasiology Conference, Ghent, Belgium, May 29 – June 2, 2006.