

A SYSTEMATIC EXAMINATION OF PRACTICE AMOUNT IN CHILDHOOD
APRAXIA OF SPEECH (CAS) TREATMENT USING AN INTEGRAL
STIMULATION APPROACH

A Thesis
Submitted to
the Temple University Graduate Board

In Partial Fulfillment
of the Requirements for the Degree
MASTER OF ARTS

by
Mackenzie Welsh
August 2017

Thesis Approvals:

Dr. Edwin Maas, Advisory Chair, Communication Sciences and Disorders

Dr. Jamie Reilly, Communication Sciences and Disorders

Francine Kohen, M.A. CCC-SLP

ABSTRACT

The purpose of this study was to examine how a critical principle of motor learning, practice amount (high number of trials versus a low number of trials), affects speech motor learning in childhood apraxia of speech (CAS). It also sought to contribute to the literature base regarding using an integral stimulation approach for these children. Currently, a limited evidence base exists for decision-making regarding practice amount in CAS treatment. Using a single-case experimental design with two participants, three target sets of utterances (High Amount, Low Amount, and Control) received different amounts of treatment. Outcomes were compared in terms of retention. Targets were scored regarding perceptual (prosodic and segmental) accuracy. Effect sizes were computed to quantify the extent of treatment effects. For both participants, results show some evidence suggesting a higher amount of practice is advantageous and leads to greater learning. A low amount of treatment did not show clear differences compared to not receiving any treatment. Caution should be taken when interpreting these findings due to its small sample size and modest effects. Results suggest that the integral stimulation approach may only be effective if provided with a significantly high amount of practice. Further research is needed to examine how the principles of motor learning and the integral stimulation approach should be sensibly and systematically applied to promote best outcomes for this population.

This thesis is dedicated to my family, friends,
and mentors who have supported me
throughout graduate school.

ACKNOWLEDGMENTS

I would like to specifically acknowledge my thesis advisor, Dr. Edwin Maas for allowing me to take part in this invaluable learning experience and for his continued mentorship and academic support. I would also like to thank Dr. Jamie Reilly and Francine Kohen for their time and assistance as part of my thesis advisory committee. I am very grateful for Anne Van Zelst and her clinical expertise in performing assessment and treatment, as well as all of the research assistants who have greatly contributed to completing this larger study including Talia Irgangladden, Sarah Rosenberg, and Helen Vradelis. Additionally, I would like to thank the Childhood Apraxia of Speech Association of North America (CASANA) for funding and making this larger project possible.

Exploring the principles of motor learning would not have been the same experience without my dear graduate school friend and thesis pal, Nicolette Kovacs, by my side. A special thanks also goes out to all of my other close family members and friends for their constant encouragement and kindness.

TABLE OF CONTENTS

	Page
ABSTRACT	ii
DEDICATION	iii
ACKNOWLEDGMENTS	iv
LIST OF TABLES	vii
LIST OF FIGURES	viii
CHAPTER	
1. INTRODUCTION	1
Schema Theory	3
Practice Distribution	4
Practice Amount.....	5
Integral Stimulation Approach/Dynamic Temporal and Tactile Cueing.....	8
2. METHODS	11
Participants.....	11
Inclusionary and Exclusionary Criteria	11
Participant 002	15
Participant 005	18
Design	20
Target Selection	22
Procedures.....	24

Probes Procedures	24
Treatment Procedures	25
Data Analysis	27
3. RESULTS	30
Participant 002's Results.....	35
Participant 005	36
4. DISCUSSION.....	37
Clinical Implications.....	41
5. CONCLUSION.....	43
REFERENCES CITED.....	44
APPENDICES	
A. TREATMENT PROTOCOL	52
B. PROTOCOL FLOWCHART.....	54
C. PROBE SCORING INSTRUCTIONS	55

LIST OF TABLES

Table	Page
1. Summary of participant information.....	14
2. Participant 002's target sets.....	23
3. Participant 005's target sets.....	24
4. Target sets' average number of syllables and baseline accuracy.....	24
5. Effect sizes and percent changes between conditions for both participants.....	35
6. Segmental score criteria.....	56
7. Segmental error types (judged for each syllable) and explanations.....	56
8. Prosodic error types and explanations.....	59
9. Examples of scoring procedures with the target "please"	63

LIST OF FIGURES

Figure	Page
1. Design overview.....	21
2. Graphical representation of Participant 002's mean set scores.....	33
3. Graphical representation of Participant 005's mean set scores.....	34
4. Protocol Flowchart.....	54

CHAPTER 1

INTRODUCTION

Childhood apraxia of speech (CAS) is considered to be a speech sound disorder (SSD) characterized by difficulty in motor planning and/or programming articulatory gestures for speech in the absence of neuromuscular impairments (ASHA, 2007; Caruso & Strand, 1999). CAS researchers and clinical experts have shown the highest consensus for the following three segmental and suprasegmental features in establishing a diagnosis of CAS: inconsistent errors on consonants and vowels in repeated speech productions, lengthened and imprecise coarticulatory transitions, and inappropriate lexical and phrasal prosody (ASHA, 2007). It should be noted these three features may not be the most fundamental or sufficient markers of CAS (ASHA, 2007).

Recent SSD prevalence estimates range from 2% to 25% in children aged 5 to 7 (Law, Boyle, Harris, Harkness, & Nye, 2000). Over 50% of school-based speech-language pathologists' (SLP) caseloads regularly include children with CAS (ASHA, 2012). CAS is estimated to occur in 1 to 2 children per 1,000 children based on clinical referral data (Shriberg, Aram, & Kwiatkowski, 1997), and was also identified in 3.4% to 4.3% of children referred for speech disorders (Delaney & Kent, 2004). Children with CAS are reported to be at risk for reduced development of reading, spelling, and writing skills, which may result in decreased social communication skills and academic potential (Lewis, Freebairn, Hansen, Iyengar, & Taylor, 2004; Gillon & Moriarty, 2007). Considering children with CAS may require prolonged and intense treatment (ASHA,

2007; Davis & Velleman, 2000; Strand, 1995), identifying the amount of practice required for successful outcomes in this population is crucial when considering the limited resources of many clinical settings (e.g., high caseloads and limited third-party reimbursements).

CAS is thought to require intensive and extensive treatment (ASHA, 2007; Davis & Velleman, 2000; Strand 1995), but there is a limited evidence base in CAS treatment findings to support this notion. This is due to the lack of a standardized definition of CAS, problems with differential diagnosis, varying severity degrees of CAS and characteristics between children who have CAS, and varying intrapersonal CAS characteristics over time (ASHA, 2007). This study aimed to determine optimal treatment conditions for children with CAS. The principles of motor learning, typically applied with nonspeech motor skills, was applied in CAS treatment to differentially examine how these principles apply to speech motor learning. Specifically, this study examined how one of the principles, practice amount (large vs. small number of practice trials), influences the retention (maintenance of a skill over time) of speech motor learning. In addition, this study represents a replication to determine if implementing an integral stimulation approach for CAS treatment results in retention using a single-case experimental design.

The motor learning literature provides a framework for systematically examining how practice conditions should be applied in CAS treatment in order to obtain optimal outcomes. Within this literature, certain practice conditions have been suggested to enhance motor learning (retention and transfer) of trained targets over others. This set of

conditions is referred to as the *principles of motor learning* (Maas et al., 2008; Strand, Stoeckel, & Baas, 2006). For example, the nonspeech motor learning literature suggests variable over constant practice enhances learning (Lee, Magill, & Weeks, 1985; Wulf & Schmidt, 1997). Since practice conditions affect motor programming and planning (Wright, Black, Immink, Brueckner, & Magnuson, 2004), and this is the suspected area of deficit in CAS, incorporating these principles into CAS treatment has been recommended (Strand et al., 2006; Robin, Maas, Sandberg, & Schmidt, 2007). Recently, many studies have incorporated the principles of motor learning in CAS treatment (e.g., Maas & Farinella, 2012; Strand et al., 2006). Currently, only four have examined practice conditions in CAS treatment (Edeal & Gildersleeve-Neumann, 2011; Maas & Farinella, 2012; Maas, Butalla, & Farinella, 2012; Namasivayam et al., 2015) and their effect on outcome measures. Only two (Edeal & Gildersleeve-Neumann, 2011; Namasivayam et al., 2015) examined the effects of practice amount.

Schema Theory

Schema Theory is a theory of motor control and learning. It proposes storing and retrieving rapid and discrete movements involves motor programs that are derived from memory and adjusted to the requirements of a specific situation (Schmidt, 1975, 2003; Schmidt & Lee, 2005). For example, kicking a soccer ball involves a forward kicking motion governed by one *general motor program*, but the *parameters* (duration and amplitude of the overall movement) may depend on the requirement of how far the ball must travel. This requires knowledge of the relations among initial conditions (e.g.,

leg/feet position), generated motor commands (e.g., amplitude of muscle contraction), sensory consequences of the commands (e.g., tactile sensation of kicking the ball), and outcome of the movement (e.g., the distance the ball traveled). When applied to speech production, Schema Theory predicts a greater number of practice trials may allow for more opportunities to establish relations between these various types of information for speech movements and their overall memory stabilization.

Practice Distribution

Practice distribution refers to how amount of practice is distributed over time and is different yet empirically difficult to isolate from practice amount. Very few studies have examined both of these conditions, and some have confounded practice distribution and amount (Namasivayam et al., 2015; Allen, 2013). For example, Namasivayam et al. (2015)'s *More* group experienced 20 treatment sessions over 10 weeks. This condition received both massed practice and a higher amount of practice than the *Less* group, which only received 10 treatment sessions over 10 weeks. In order to determine how conditions under practice amount (high number of trials vs. low number of trials) and practice distribution (massed vs. distributed) differentially affect speech motor learning, they must vary separately. This thesis is part of a larger study differentially examining the effects of practice amount, practice distribution, and social validity outcomes measured by intelligibility ratings from unfamiliar listeners in connected speech. This specific single-case experimental design study will focus on the variable of practice amount without confounding practice distribution in two children with CAS.

Results of previous studies comparing practice amounts combined with the predictions of Schema Theory lead to predicting that higher amounts of practice will result in greater retention and transfer of speech targets in CAS treatment in the context of an integral stimulation approach.

This study sought to provide controlled evidence regarding the role of practice amount in speech-motor learning, specifically in CAS treatment, in order to contribute to the currently limited evidence-base on clinical decision-making for this population. Keeping the time period for comparison among conditions constant will control for maturation, and varying the number of targets will allow for control of practice distribution (in terms of frequency of sessions and number of trials per session). This study will also contribute to the evidence-base for the integral stimulation approach in CAS treatment, as supported by Murray, McCabe and Ballard's (2014) systematic review of single-case experimental CAS treatment studies.

Practice Amount

Practice amount refers to the number of practice trials for a given target within a treatment session. According to Schmidt and Lee (2005), it is potentially the most powerful principle of motor learning. The nonspeech motor literature suggests increased practice amount indicates greater motor learning, specifically in terms of retention and transfer in a variety of nonspeech motor tasks (Park & Shea, 2003, 2005; Shea & Kohl, 1991). Some evidence has been found to support similar effects with nonspeech motor skills in regards to disordered systems in the adult population. For example, Kwakkel et

al.'s (2004) systematic review of treatment intensity of augmented exercise therapy time (AETT) after stroke reported greater gains in activities of daily living (ADL) with more intensive AETT implementation within the first 6 months post-stroke.

Preliminary evidence has also been found to support similar effects for increased amounts of practice supporting greater motor learning in disordered speech/language systems in the adult population. Cherney, Patterson, Raymer, Frymark, and Schooling's (2008) systematic review of Constraint-Induced Language Therapy (CILT) for individuals with stroke-induced aphasia found modest evidence supporting greater positive treatment outcomes with increased treatment amounts. Varley et al. (2016) also examined report outcomes of self-administered computer therapy for post-stroke apraxia of speech, and reported a significant positive correlation between treatment dose and response.

The extent to which these effects apply to speech production in disordered speech systems in children warrants further exploration. Few studies have examined the effects of practice amount specifically in CAS. Results of Edeal and Gildersleeve-Neumann (2011) and Namasivayam et al. (2015) suggest higher amounts of practice support greater retention and transfer of speech motor skills and currently provide the strongest evidence-base towards this prediction.

Edeal and Gildersleeve-Neumann (2011) used a single-case experimental design to examine accuracy measures on randomly assigned targets from a moderate-frequency treatment condition (30 to 40 trials/15 minutes) and a high production frequency treatment condition (100+ trials/15 minutes) using an integral stimulation approach in

two children diagnosed with CAS. One child's sessions were 3 times per week for 11 weeks, and the other's were 2 times per week for 5 weeks. Each session was approximately 40 to 50 minutes long. Probes were administered for 5 minutes at the end of each condition administration and maintenance data were collected 2 weeks post-treatment. Results showed greater speech motor learning outcomes on target sounds under the high production frequency condition. This study currently provides the highest level of evidence supporting greater gains for higher amounts of practice; however, it only examined 2 participants and outcomes reflected only accuracy of the target sound, not the whole words in which the sounds were embedded. There may have also been within-session distribution effects since the number of trials varied within the same period of time for each treatment. Further research is needed to examine amount of practice and the integral stimulation approach and their effect on a greater number of participants with CAS as well as whole-word accuracy outcomes, while controlling for within-session distribution effects.

Namasivayam et al. (2015) compared two groups of children with CAS in a pre- and post-treatment design (*More*: 20 sessions/10 weeks, n=21; *Less*: 10 sessions/10 weeks, n=12) on outcome measures related to articulation, functional communication, and intelligibility. A Motor Speech Treatment Protocol (MSTP) was employed in the treatment approach. Results of the *More* group showed statistically significant greater gains on articulation and functional communication measures, and a generally larger effect size on most variables in comparison to the *Less* group; however, neither group

improved significantly in intelligibility and the study confounded practice distribution and amount (i.e., the *More* group also had more frequent treatment sessions).

Other studies have also examined practice amount but the relevance to CAS is uncertain. Allen (2013) performed a phonological intervention study for children with SSDs using a multiple oppositions approach to examine the effect of practice amount between two experimental groups and one control group. One experimental group received treatment 3 times per week for 8 weeks (P3), the other received it once per week for 24 weeks (P1). A control group received a separate intervention for 8 weeks (C). Measures were taken after 8 weeks across the three groups. The P3 condition had significantly greater performance on accuracy measures after 8 weeks than the P1 condition. A medium effect size was reported between the P3 and P1 conditions, and a large effect size was noted between the P3 and C conditions. At the 8-week measure, the P3 condition received 24 treatment sessions in 8 weeks, whereas P1 received 8 treatment sessions in 8 weeks. In this case, practice amount and practice distribution both varied. P3 received massed practice and three times the amount of practice as P1, whereas P1 received less practice that was relatively distributed. Results showed greater gains under the P3 condition, but it cannot be determined whether and to what extent practice amount or distribution contributed to these outcomes.

Integral Stimulation Approach/Dynamic Temporal and Tactile Cueing

Practice amount was examined under the integral stimulation approach/dynamic temporal and tactile cueing (DTTC). Integral stimulation is a motor-based approach that

involves hierarchical speech treatment, moving from the initial level at which clients are most successful to levels with increasing length and complexity. The approach begins with clinician imitation (“watch me, listen to me, say what I say”), moves to delayed imitation, and works towards independent production using multimodal cues (tactile, auditory, and visual) tailored to the child’s accuracy level, response, and motivation. Targets may begin at the syllable, word, and sentence levels. Cues are added or faded depending on the current level and progression in the treatment hierarchy (Rosenbek, Lemme, Ahern, Harris, & Wertz, 1973). Speech rate may also be reduced in order to provide the highest level of support and increase motor planning duration and promote awareness of tactile, proprioceptive, and kinesthetic feedback (Strand et al., 2006). As speech output accuracy increases, speech rate may be increased in order for the participant to eventually sustain a typical conversational rate.

Murray et al. (2014) rated integral stimulation/DTTC to have “preponderant evidence” according to Smith’s (1981) level of certainty of treatment effects. Several studies have also shown replicated effects when using integral stimulation approaches for intervention in CAS (Strand & Debertine, 2000; Strand et al., 2006; Baas, Strand, Elmer, & Barbarresi, 2008; Edeal & Gildersleeve-Neumann, 2011; Maas & Farinella, 2012; Maas et al., 2012). This approach is also the only CAS intervention approach supported by independent research groups (Baas et al., 2008; Edeal & Gildersleeve-Neumann, 2011; Maas et al., 2012; Maas & Farinella, 2012; Strand & Debertine, 2000; Strand et al., 2006). To date, studies have almost exclusively examined the effect of integral stimulation on direct perceptual measures of speech sound accuracy (e.g., whole-word

accuracy, accuracy of target sounds in words) rather than outcome measures related to social validity (e.g., activity/participation, personal and environmental factors; cf. World Health Organization, 2001; see Baas et al., 2008, for exception). Lastly, this study may further contribute to the evidence-base for this approach and to its candidacy for Phase III research (Murray et al., 2014).

CHAPTER 2

METHODS

Participants

Inclusionary and Exclusionary Criteria

Two children diagnosed with moderate to moderate-to-severe CAS participated in this study, which was part of a larger study involving six children. Participants were required to be from a monolingual English or English-dominant household and educational setting. Recruitment took place through the Temple University Speech-Language-Hearing Center, local educational settings, and health agencies.

Comprehensive speech and language evaluations were performed for the participants using formal and informal measures appropriate to each participant's age and level of communication (described further below). To be included in the study, participants had to have (1) a speech sound disorder (based on a score below the 10th percentile on the *Goldman Fristoe Test of Articulation*; GFTA-2, Goldman & Fristoe, 2000), (2) a diagnosis of CAS (see below), (3) a verbal output of 50 words or more, and (4) communicative intent at the start of the study as determined by parent report and the SLP. Exclusionary criteria were (1) co-morbid diagnoses of other neurobehavioral disorders that significantly affect communication and/or social interactions (e.g., autism spectrum disorders) as determined based on referral diagnosis, (2) significant impairments of hearing or vision as determined by parent report, and (3) significant

impairments of oral structure (e.g., cleft palate) as judged by the clinician based on an oral motor examination (Robbins & Klee, 1987).

Currently, there is a lack of validated diagnostic markers for CAS (ASHA, 2007) and the gold standard for diagnosis is expert opinion. To increase confidence in diagnosis for this study, the diagnosis of CAS required meeting a pre-defined set of three stringent criteria. These criteria were designed to reflect the three features generally agreed upon among researchers and supported by ASHA (2007): inconsistent vowel and consonant errors during repeated speech productions, difficulty transitioning between articulatory placements, and abnormal prosody.

First, children had to receive a diagnosis of CAS by four independent clinical experts. Initial diagnoses were made by a certified SLP with extensive expertise in diagnosing CAS, based on observations from a range of speaking tasks. Three additional clinically experienced SLPs in CAS diagnosis and treatment confirmed the diagnoses for each participant independently from video recordings of the assessments. Each SLP assigned a score to the participant based on their assessment observations using a three-point scale, in which 0 = no CAS, 1 = possible CAS, and 2 = definitely CAS. To be included, children had to receive an average score greater than 1 and could not receive a score of 0 from any SLP.

Second, children had to demonstrate inconsistent errors as determined by a score over 40% on the Inconsistency subtest of the *Diagnostic Evaluation of Articulation and Phonology* (DEAP; Dodd, Hua, Crosbie, Holm, & Ozanne, 2006). This subtest involves naming 25 words three times each. Each target word is scored as consistent (similar

production on all trials) or inconsistent (different productions during at least two trials). This test yields an Inconsistency Score, and provides a standardized and operationalized measure of inconsistency in production, considered one of the consensus features of CAS (ASHA, 2007). This study followed the DEAP test manual in considering 40% to represent significant inconsistency across articulatory productions (Dodd et al., 2006).

Finally, children were required to receive a CAS score of 2 on the *Maximum Performance Task Protocol* (MaxPT; Thoonen et al., 1996, 1999; see also Rvachew, Hodge, & Ohberg, 2005). This protocol involves tasks such as vowel and fricative prolongation and repeating syllable sequences (e.g., 'pataka') at maximal rates, and yields two independent scores: a Dysarthria score and a CAS score (each with 3 levels: 0 = disorder not present, 1 = undetermined/disorder possibly present, 2 = disorder present). Combining scores on these tasks differentiates CAS from other SSDs with 100% sensitivity and 91% specificity (Thoonen et al., 1999).

Additional descriptive measures, not used to determine inclusion/exclusion, were also obtained, and included receptive language subtests of the *Clinical Evaluation of Language Fundamentals* (CELF-4; Semel, Wiig, & Secord, 2003), the *Expressive Vocabulary Test* (EVT-2; Williams, 2007), the *Peabody Picture Vocabulary Test* (PPVT-4; Dunn & Dunn, 2007), nonverbal cognition subtests of the *Reynolds Intellectual Assessment Scales* (RIAS; Reynolds & Kamphaus, 2003), and the phonological awareness subtests of the *Comprehensive Test of Phonological Processing* (CTOPP-PA; Wagner, Torgesen, & Rashotte, 1999). In addition, a case history was obtained and parents completed two questionnaires regarding communicative function, namely the

Intelligibility in Context Scale (ICS; McLeod et al., 2012, 2015) and the *Functional Outcomes of Communication Under Six* (FOCUS; Thomas-Stonell et al., 2012). The FOCUS-34 is a parent questionnaire designed to evaluate a child’s speech and communication skills through communicative participation and interaction in everyday situations. The ICS is a parent questionnaire designed to measure a child’s intelligibility across a variety of communication partners (e.g., parents, immediate family, teachers, strangers, etc.). A summary of evaluation results for both children is provided in Table 1.

Table 1		
<i>Summary of participant information</i>		
Participant Information	Participant 002	Participant 005
Age	7;11	6;0
Sex	M	M
Dx	CAS; Articulation disorder; Phonological error patterns	CAS; Dysarthria; Receptive-Expressive language disorder
CAS score (4 expert average rating)	1.88	
GFTA-2 SS (%ile)	55 (< 3 rd)	45 (< 1 st)
DEMSS		
Articulatory accuracy	40%	40%
Distorted vowels	25%	15%
Atypical prosody	18%	0%
Inconsistency	54%	38%
DEAP Inconsistency	80%	64%
MaxPT		
Maximum Phonation Duration (MFD) (Mean of normative sample = 14.98)	6.43 sec.	3.18 sec.
Maximum Fricative Duration (MFD) (Mean of normative sample = 15.01)	7.74 sec.	2.32 sec.
Max. Rep. Rate – monosyllables (MRRmono) (Mean of normative sample = 5.27)	4.59 syll./sec.	3.05 syll./sec.

Table 1		
<i>Continued</i>		
Participant Information	Participant 002	Participant 005
MaxPT (continued)		
Max. Rep. Rate – trisyllables (MRRtri) (Mean of normative sample = 5.33)	4.04 syll./sec.	unable
CAS Score	2	2
Dysarthria Score	0	2
CELF-4 Receptive Language Index	105	(could not compute)
PPVT-4 SS (%ile)	89 (23 rd)	75 (5 th)
EVT-2 SS (%ile)	88 (21 st)	60 (0.4 th)
CTOPP-PA SS (%ile)	88 (21 st)	(could not compute)
RIAS Nonverbal T-score (%ile)	42 (21 st)	45 (30 th)
FOCUS-34 Average Score (7 max.)	2.0	3.15
ICS Average Total Score (5 max.)	3.2	3.2

Participant 002

At the time of the evaluation, Participant 002 was a 7-year, 11-month old male. Per case history report, Participant 002 lived with his parents, one sibling, and attended an elementary school. He does not present with a family history of speech sound disorders or any feeding/swallowing issues. His hearing and vision were reported as within normal limits. He was diagnosed with CAS at age 3 prior to this study, but does not present with any previous learning or neurological diagnoses.

On the GFTA-2, Participant 002 received a standard score of 55 and a percentile rank of 3. His articulation errors were primarily characterized by substitution and omission errors. Voicing errors were also noted. He produced instances of possible phonological error patterns including initial and final voicing, gliding, and cluster

reduction. He also exhibited an unusual but relatively consistent substitution of a retroflex /l/ for /ɹ/.

To examine speech skills further, several measures were administered as part of the evaluation protocol, including the *Dynamic Evaluation of Motor Speech Skill* (DEMSS; Strand, McCauley, Weigand, Stoeckel, & Baas, 2013). The DEMSS is a dynamic assessment of motor speech skill to determine the degree in which a child exhibits speech motor planning and programming difficulties. It involves repetition of speech movement gestures varying in length as well as phonemic and phonotactic complexity across a variety of syllable structures. Multidimensional scoring is used to provide measures of articulatory accuracy, vowel accuracy within utterances, prosodic accuracy across utterances, and consistency of produced errors. The DEMSS is currently under development and is not yet standardized and distributed for widespread use (Strand et al., 2013); therefore, observations were not compared to a normative sample but used qualitatively. Participant 002's results on the DEMSS evidenced significantly decreased articulatory accuracy, distorted vowels and vowel substitution errors, atypical prosodic features, and inconsistent productions on repeated trials. Inconsistency of productions were also noted on the DEAP Inconsistency subtest, for which Participant 002 received a score of 80%, surpassing the test's criterion score of 40% inconsistency. These results suggest speech motor planning impairments.

Oral structure and function were examined using an oral mechanism exam (Robbins and Klee, 1987) and the MaxPT protocol (Rvachew et al., 2005; Thoonen et al., 1996, 1999). Participant 002 exhibited normal oral structures and function, although there

was a possible decreased range of movement, imprecision, and incoordination when alternating protrusion and retraction of his lips and when protruding and elevating his tongue. During the coordinated speech movement tasks, he presented with multiple phonemic substitution and distortion errors on both consonants and vowels. Vocal quality, resonance, and volume were judged to be within normal limits. On the MaxPT protocol, Participant 002 obtained a CAS score of 2 and a dysarthria score of 0. His maximum phonation duration (MPD) and maximum fricative duration (MFD) were below the normative range, his maximum monosyllabic repetition rate (MRRmono) was within normal limits, and he had difficulty producing trisyllabic sequences and produced only one correctly sequenced trial, with a rate in the lower end of the normative range. He had difficulty maintaining the sequence of target sounds across repetitions, as well as consonant additions and vowel errors.

In addition to meeting the criteria based on the MaxPT protocol and the DEAP Inconsistency subtest, Participant 002 met the clinical expert rating criterion. He received an average score of 1.88 and no scores of 0 from any of the four experts (1.5, 2, 2, 2). Thus, taken together, confidence in accuracy of CAS diagnosis is high.

As can be seen in Table 1, Participant 002 exhibited language abilities in the normal range, with low average scores for receptive and expressive language measures. He also demonstrated phonological awareness and nonverbal cognitive skills within the normal range for his age. According to the FOCUS-34, Participant 002 received an Average Focus Score of 2.0 out of 7.0, suggesting he presents with a severe limitation in his communicative participation. On the ICS, he received an Average Total Score of 3.2,

indicating he is “usually understood by his immediate family and teachers” but he is “rarely” to “sometimes” understood by other less familiar individuals such as his extended family and acquaintances.

Participant 005

At the time of the evaluation, Participant 005 was a 6-year, 0-month old male. Per case history report, he lived with his parents and one sibling. He does not present with a family history of speech sound disorders. He has a positive history of severe recurrent ear infections that resulted in temporary hearing loss and the need for myringotomy tubes; however his current hearing status was reported as within normal limits, along with his vision and swallowing. His medical history is remarkable for supraventricular tachycardia in utero at 33 weeks. Participant 005 was also diagnosed with CAS at age 3, but does not present with any previous learning or neurological diagnoses.

In terms of articulation, Participant 005 received a standard score of 45 on the GFTA-2, which is below the first percentile rank. His errors were mostly characterized by atypical substitutions (e.g., /dʒ/ for /n/, /l/ for /bw/, /gl/ for /f/) and some conformed to phonological patterns (e.g., gliding). Participant 005’s results on the DEMSS evidenced moderately decreased articulatory accuracy, distorted vowels and vowel substitution errors, and inconsistent productions on repeated trials. Inconsistency was confirmed by the DEAP, on which he received an Inconsistency score of 64%. On the DEMSS, he did not present with atypical prosodic features (0% error); however, he was noted to have significant difficulty with coarticulation during spontaneous connected speech. These

attempts were characterized by a slow rate, effortful speech, and reduced accuracy in coarticulatory transitions between words and phonemes. Combining these task results suggest Participant 005 presented with speech motor planning impairments.

Results of Participant 005's oral structure and function assessment suggest he also presented with dysarthria characterized by weakness due to reduced oropharyngeal motor skills in the absence of structural deficits. His facial and oral structures were grossly intact; however, he inconsistently exhibited a lowered mouth posture with inconsistent lingual protrusion, drooling, and pooling of saliva. He opened and closed his mandible without extraneous anterior or lateral movement and accurately rounded, protruded and retracted his lips; however, he demonstrated a reduced range of movement and incoordination when alternating lip protrusion and retraction and a decreased ability to protrude and elevate his tongue. During the coordinated speech movement tasks, he presented with multiple phonemic substitution errors, reduced vocal loudness, and hypernasality. Reduced vocal loudness and hypernasality were consistently present throughout all speech tasks, consistent with a diagnosis of dysarthria.

Participant 005 obtained a CAS score of 2 on the MaxPT. His performance on the MaxPT was below the normative range for all tasks within this protocol. He accurately produced repetitions of single syllables (MRRmono) but could not consistently produce at least 10 syllables on a single exhalation and frequently produced voicing errors on consonants. He was unable to produce trisyllabic repetitions (MRRtri). Using the standard scoring criterion, his MRRmono was 3.05 syllables/second. Combined with his reduced MPD results, Participant 005 received a Dysarthria score of 2.

Consistent with the evaluation results noted above, Participant 005 was judged to have CAS based on the four clinical expert ratings. He received an average score of 1.63 and no scores of 0 from any expert (0.5, 2, 2, 2). Thus, he met the criteria for presence of CAS, in addition to dysarthria.

As can be seen in Table 1, Participant 005 exhibited severely reduced language abilities in comparison to same-aged peers, with extremely low scores for receptive and expressive language measures. He also demonstrated moderately to severely reduced phonological awareness skills, but his nonverbal cognitive skills were within the average range for his age. According to the FOCUS-34, Participant 005 presents with a moderate to severe limitation in his communicative participation. Results of the ICS indicate he is “usually understood by his immediate family and teachers” but he is “rarely” to “sometimes” understood by other less familiar individuals such as his extended family and acquaintances.

Design

An alternating treatments design with a baseline phase, a withdrawal/maintenance phase, and unrelated untreated Control targets was used to assess learning of targets trained with different practice amounts (see Figure 1 for overview of design). An additional multiple-baselines-across-participants component was added in which the two participants received a different number of baseline points. Performance was assessed via weekly probes (described further below). This design exerts experimental control because improvements in target accuracy should only occur after the start of treatment, and only

for treated targets but not for untreated items. The total duration for the examination of practice amount was 9 weeks. There were four weeks of initial baseline measures, four weeks of treatment in which both *High* and *Low Amount* conditions were administered, and a two-week maintenance/withdrawal phase. Each participant had a minimum of four baseline measures before the start of treatment. Treatment was initiated only when speech accuracy on probes was stable, as defined by $\leq 25\%$ session-to-session changes and $\leq 10\%$ difference between the first and last baseline measures. These criteria were based on judgments by a trained research assistant. In the four-week treatment phase, treatment was provided twice per week. In each session, both conditions (*HA: High Amount, LA: Low Amount*) were administered for 25 minutes, applied to different target sets (described below). To control for systematic order effects, the order of administration of the *HA* and *LA* conditions was pseudorandomized, in that the order for the first session of the week was determined randomly (with a roll of a die) and the order was reversed for the second session of the week. Within each condition, targets were practiced in random order. A brief (5-10 minute) break was provided between conditions. The figure below depicts the design (note that TE refers to teaching episodes).

		High Amount 800 TEs, 5 targets Low Amount 800 TEs, 10 targets	
	A ₁	B ₁	A ₂
<i># of weeks:</i>	3	4	2

Figure 1. Design overview.

Target Selection

Implementation of the independent variable (practice amount) was accomplished by creating two sets of individualized speech targets (words, phrases) with a different number of targets: *HA* (5 targets) and *LA* (10 targets). In addition, a set of Control items was generated (10 items) which were never treated. Given an equal amount of time (25 minutes per session, eight sessions total) and identical treatment procedures, the five items in the *HA* set received approximately twice as much practice as the ten items in the *LA* set.

Targets were words and phrases, chosen based on each participant's pre-treatment speech output level and functional relevance. The *Functional Communication Parent Questionnaire* (Wilson & Gildersleeve-Neumann, 2014) was administered in order to take into account caregiver preference on functional target selection. In the assessment phase, possible targets were elicited multiple times in order to further individualize items (e.g., fringe vocabulary) and identify items that were consistently difficult for the child.

Potential target items were divided into three sets of ten items each. Care was taken to match the sets with respect to the number of items from various interest categories (e.g., names of family members, teachers, social utterances, needs, favorite activities). In addition, the three sets were matched on the number of syllables and baseline performance accuracy as determined by a trained research assistant (all p s > 0.7, two-tailed two-sample t-tests). Each target set was created such that accuracy (based on trained research assistant's judgment) was between 1% to 50% accuracy, in order to minimize floor and ceiling effects.

After matching, the three sets were randomly assigned to the *LA*, *HA*, and Control conditions (using the ‘=rand()’ function in Microsoft Excel). Next, the ten items in the *HA* condition were divided into two matched sets of five items, while maintaining the match of each subset with the *LA* and Control conditions. One of these subsets was then randomly assigned to the *HA* condition.¹ Thus, there were three sets, as follows: *HA* set (5 items), *LA* set (10 items), and a Control set (10 items). Tables 2 and 3 below list target items for each participant; identifying information was removed (each syllable substituted with an X). Additionally, Table 4 displays the average number of syllables and baseline accuracy, represented by its standard deviation (SD) provided for each set. The SD for baseline accuracy was based on the averages of item scores and not set scores per probe.

Table 2		
<i>Participant 002's target sets</i>		
Low Amount	High Amount	Control
I want a cereal bar	breakfast	I like chocolate milk
XXX is my speech teacher	My dad's name is XX	XXXX XXX X (street address)
XXX (friend's name)	When are we going?	it's almost 6 o'clock
I like playdough	5 minutes later	X is my brother
I live in XXX	When is the movie?	Where's the bathroom?
What time is it?		Do you want to play?
Memory		I'm sorry
Talk to you later		kitchen
Beauty and the Beast		after 2 hours
What's your name?		Where are my shoes?

¹ The other five items from the *HA* set were treated in a second treatment phase, as part of the larger study of which the present study is a part.

Table 3		
<i>Participant 005's target sets</i>		
Low Amount	High Amount	Control
XX XX (name)	excuse me	Pepper (dog)
thank you	yesterday	hello
Teacher	Mickey	towel
Doing	What's your name?	tomorrow
Disney World	sister	Phanatic
XX (friend)		XX (friend)
Please		see you later
Hamburger		guacamole
Where is the bathroom?		How are you?
I'm tired		I'm thirsty

Table 4			
<i>Target sets' average number of syllables and baseline accuracy</i>			
Participant	Set	# syllables (SD)	Baseline accuracy (SD)
002	Low Amount	4.80 (1.75)	0.18 (0.21)
”	High Amount	4.60 (1.52)	0.27 (0.19)
”	Control	4.70 (1.64)	0.22 (0.22)
005	Low Amount	2.70 (1.16)	0.11 (0.23)
”	High Amount	2.60 (0.55)	0.06 (0.13)
”	Control	2.70 (0.67)	0.12 (0.20)

Procedures

Probes Procedures

Primary data for this study were based on utterances elicited on probes. Probes contained *HA* and *LA* items as well as the untreated Control items. Items on the probe list were randomized, with a different random order for each probe administration. Probes were administered during baseline, treatment, and maintenance phases. Probes for Participant 002 involved a delayed reading task to assess independent production. This

involved showing him a card with the written target, then removing the card and asking and him to produce the target. For Participant 005, probes were elicited in a delayed imitation task (delay ~2-3 seconds). During the treatment phase, probes were administered weekly at the beginning of the first treatment session in order to assess retention from the previous session. One probe was presented each week during the 2-week maintenance/withdrawal phase.

Treatment Procedures

This study implemented an integral stimulation approach to CAS treatment that implements principles of motor learning similar to previous studies (Edeal & Gildersleeve-Neumann, 2011; Maas & Farinella, 2012; Strand & Debertine, 2000). The approach begins with integral stimulation (“watch me, listen to me, say what I say”), moves to delayed imitation, and works towards independent production using multimodal cues that may be added or faded according to the treatment hierarchy. Speech rate may also be targeted progressing from slow to conversational speech rates.

All treatment was provided by a single SLP with extensive experience with treatment for CAS. A brief description of the protocol is provided below; a more detailed protocol is provided in the Appendix A. Targets were practiced in teaching episodes, which were defined as the sequence of events that included initial attempt of a target using a given elicitation method (e.g., repetition), practice (clinician feedback and cues, additional attempts), and a second attempt using the same elicitation method as for the initial attempt.

The elicitation procedures were based on the participant's needed level of support to produce the target as judged by the clinician, and consisted of one of four methods (in ascending order of difficulty): immediate imitation at reduced speech rate, immediate imitation at normal speech rate, delayed imitation at normal speech rate, and independent production (e.g., answering questions, naming pictures). In each session, each condition started with immediate imitation at normal rate. If a participant provided a correct response on the initial attempt for two consecutive teaching episodes, the next teaching episode used the next level elicitation method (e.g., delayed imitation). If a participant produced incorrect responses on the initial attempt for two consecutive teaching episodes, the elicitation method for the next teaching episode reverted to a less challenging elicitation method (e.g., immediate imitation at reduced rate), or remained at immediate imitation at reduced rate.

After the first response, the clinician provided verbal feedback and additional cues and practice (the 'teaching' part of the teaching episode). Feedback included both knowledge of performance and knowledge of results. Knowledge of results provides the participant with information about the speech movement's accuracy in relation to the target response (e.g., "Good job!"), whereas knowledge of performance involves the provision of qualitative information regarding a target response (e.g., "Your tongue was not high enough for that sound"). Verbal feedback was provided with a brief (~2 second) delay and for fewer than 100% of teaching episodes, based on previous work indicating that a reduced feedback frequency was advantageous for two of three children with CAS (Maas et al., 2012). In this study, we provided feedback on a fading schedule, decreasing

by 10% for every subsequent 10 teaching episodes within each condition (e.g., 10/10, 9/10, 8/10, etc.; as in Ballard, Maas, & Robin, 2007), in order to work from greater support to more independent productions over the course of each session. The clinician used a feedback-tracking sheet in order to determine if feedback was required for a given teaching episode. On no-feedback teaching episodes, the clinician did not provide explicit verbal feedback but waited ~2 seconds and then provided cues and practice ('teaching').

Teaching could include tactile or visual cues, simultaneous production with the clinician, verbal explanations (e.g., "try to round your lips more"), and practice on particular words or sounds, and combinations of these cues, as deemed appropriate by the clinician for a given teaching episode. In this teaching portion of the teaching episode, complete targets were attempted up to five times before closing the teaching episode with a final elicitation using the elicitation method from the beginning of the teaching episode. Feedback (knowledge of results and performance) was provided after a brief delay and on the same fading schedule as before. Subsequently, the next teaching episode began.

Within each condition, targets were practiced in random order. This randomization was implemented by creating two index cards for each target (i.e., 20 cards for *LA*, 10 cards for *HA*), shuffling the stack of cards, and taking one card from the deck for each teaching episode.

Data Analysis

The dependent variable was perceptual accuracy of target words and phrases on the probes. Perceptual accuracy was measured using a scoring system that combines segmental accuracy scores and prosodic accuracy scores for a maximum total score of 3

per item. This scoring system is adapted from the three-way scoring system used in previous studies (e.g., Maas et al., 2012; Strand et al., 2006). A full and detailed description as well as examples can be found in Appendix C. Briefly, segmental accuracy was judged on a syllable-by-syllable basis, with a score of 2 indicating that all syllables are correct, a score of 1 indicating that at least half of all syllables are correct, and a score of 0 indicating fewer than 50% of syllables correct. Segmental errors included substitutions, omissions, distortions, additions, metathesis, and unintelligible utterances, but error types were not further specified. Prosodic accuracy was scored on a binary scale, where 1 = correct and 0 = incorrect. Prosodic errors included segmented speech, omitted syllables, added syllables, syllable reversals, stress errors (e.g., equalized, incorrect, exaggerated). Self-corrections were also considered errors on the prosody scale as they affect the entire target. Thus each item could receive a maximal score of 3, with 2 points for segmental accuracy and 1 point for prosodic accuracy. Scores for items in each set were averaged and converted to percentages to obtain the average score per condition.

Probes were scored from audio recordings by a rater blinded to treatment status and time of probe collection, to control for rater bias. Probe recordings were edited by a different research assistant to remove clues as to time of probe collection (e.g., references to holidays) and treatment status (e.g., if the clinician or child commented on a particular item having been treated). Probes were scored in randomized order to control for perceptual drift, and thus, data analyses were performed after the post-treatment probes given during the 2-week maintenance/withdrawal phase. A second blinded independent

rater scored 6/37 (16.2%) of the probes to determine inter-rater reliability. Average percentage of scores within 1 scale point was 93% (SD = 4%, range 88-98%).

Data were analyzed by visual inspection supplemented with standardized effect sizes (d statistic; Beeson & Robey, 2006), in order to compare and quantify changes relative to baseline for each set. Effect size (d) is computed as [(average accuracy post-treatment – average accuracy pre-treatment / standard deviation pre-treatment)].

Following Maas et al. (2012), improvement was operationally defined as $d > 1$ (i.e., post-treatment performance exceeds the baseline standard deviation). The percent change from baseline to the post-treatment phase was also calculated; however, effect sizes were used as the primary method of comparison. Based on the motor learning literature and previous research with children with CAS, a larger d was predicted for the *HA* set than for the *LA* set, because *HA* items received approximately twice as much practice as *LA* items. Both *HA* and *LA* sets were also compared to the Control set to determine whether any improvements on treated items exceed improvements on untreated items (i.e., whether gains on treated items can be ascribed to the treatment rather than extraneous effects such as maturation). Both *HA* and *LA* sets were expected to show larger effect sizes than the Control set.

CHAPTER 4

RESULTS

Figures 2 and 3 display the mean set scores (converted to percentages) for each participant. Table 5 depicts the effect size results for each participant in both the *HA*, *LA*, and Control conditions.

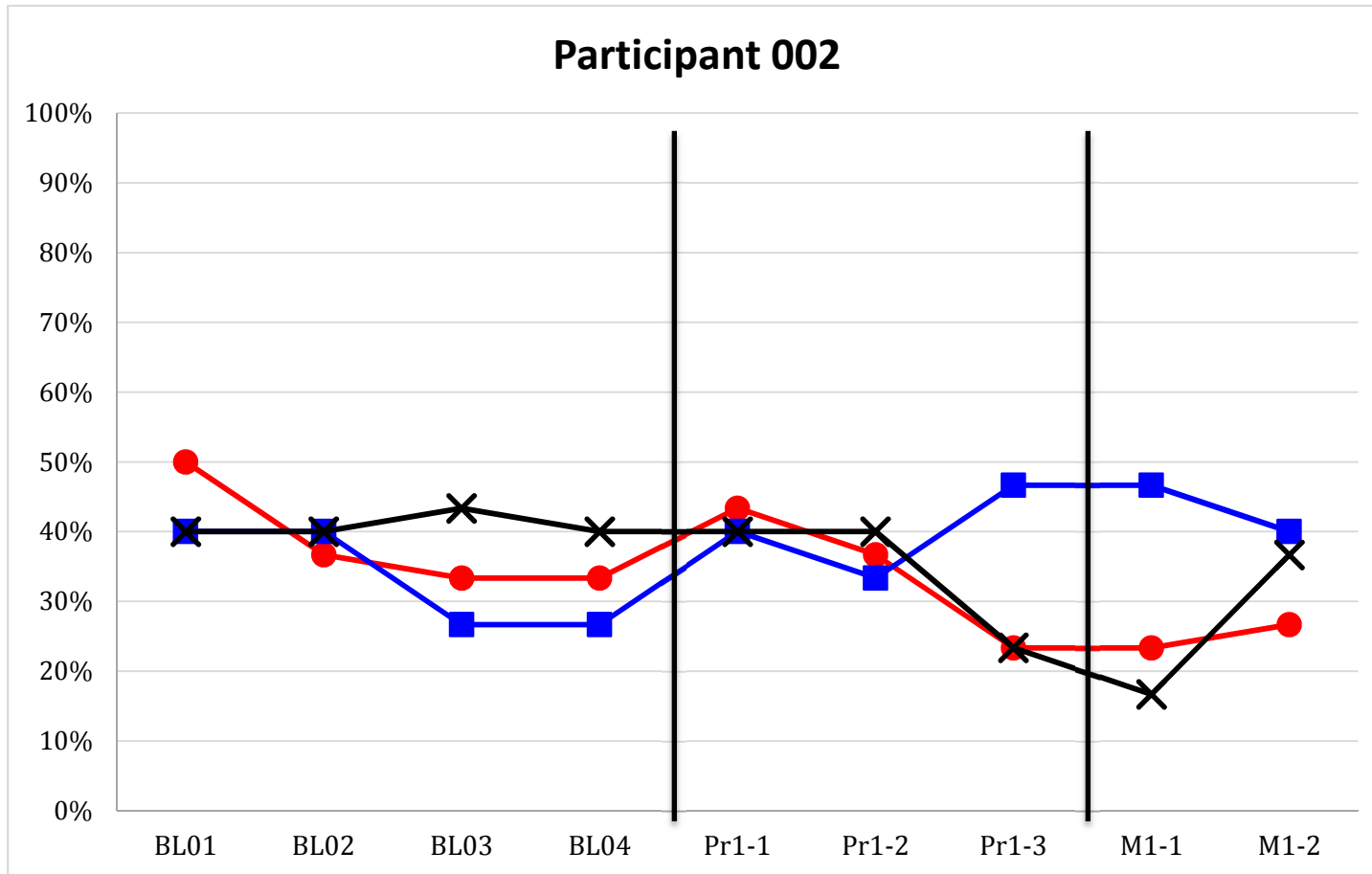


Figure 2. Graphical representation of Participant 002's mean set scores.

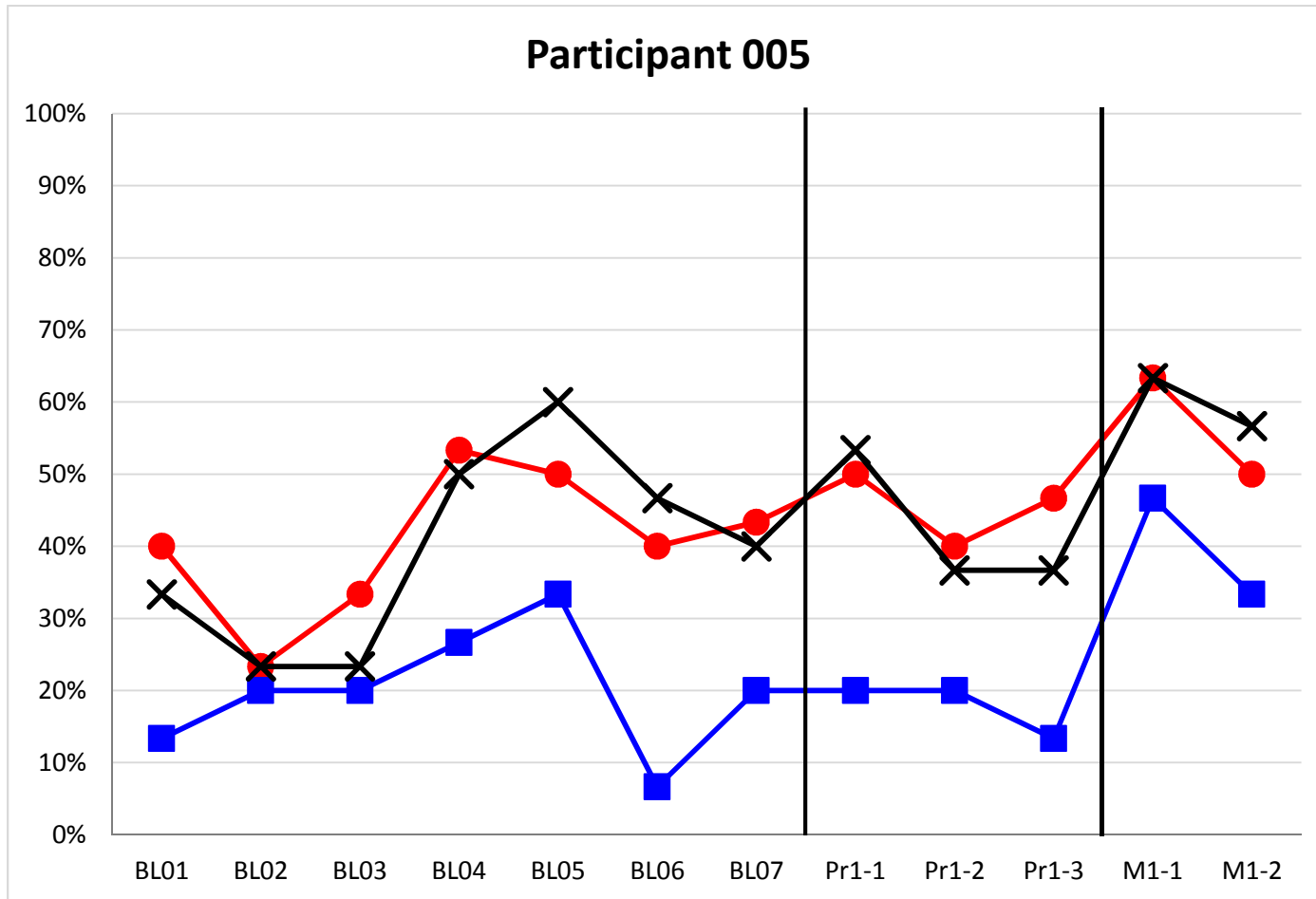


Figure 3. Graphical representation of Participant 005's mean set scores

- Low Amount
- High Amount
- ×— Control

Table 5						
<i>Effect sizes and percent changes between conditions for both participants</i>						
Participant	Low Amount d	Low Amount %change	High Amount d	High Amount %change	Control d	Control %change
002	-1.68	-13%	1.30	+10%	-8.50	-14%
005	1.61	+16%	2.32	+20%	1.48	+20%

Participant 002's Results

Visual inspection of Participant 002's mean set scores revealed that all three sets were comparable in accuracy during baseline, with stable or slightly declining accuracy before treatment. Both treated sets (*HA*, *LA*) showed a slight increase upon initiation of treatment. For the *HA* set, but not for the *LA* set, this upward trend was maintained throughout the remainder of the study. Both the *LA* and Control sets decreased in accuracy compared to baseline (*LA*: $d = -1.68$ %change = -13%; *Control*: $d = -8.50$, %change = -14%). The decrease is similar in terms of absolute amount of change, although the negative effect size for the Control set is larger than that in the *LA* condition. This is due to the fact that the Control set exhibited less variability during the baseline phase. In contrast to the *LA* and Control conditions, the *HA* condition did improve by +10%, resulting in an effect size of 1.30.

Participant 005's Results

Visual inspection of Participant 005's mean set scores reveals that accuracy initially rose during baseline but then plateaued and stabilized. During the treatment phase there is no clear evidence of improvement in any condition, but immediately following treatment all three sets show a notable increase in accuracy. The *HA* condition showed a greater change (*HA*: $d = 2.32$, %change = +20%) in comparison to the *LA* and Control conditions (*LA*: $d = 1.61$, %change = +16%; *Control*: $d = 1.48$, %change = +20%), although mean set scores upon follow-up are lower than those for the *LA* and Control sets. The *HA* condition resulted in a larger effect size due to its lower mean and standard deviation during baseline. There was no clear difference between the Control condition and the *LA* condition.

CHAPTER 5

DISCUSSION

This study examined the role of practice amount in an integral stimulation based treatment in two children with CAS. As previously noted, it was predicted that a greater amount of practice would lead to greater improvement (i.e. larger effect size for *HA* than for *LA*), consistent with the nonspeech motor learning literature and previous findings in the speech motor learning literature regarding children with CAS. Similarly, it was predicted that treated items (both *HA* and *LA*) would show greater improvement than untreated items.

Consistent with the first prediction, both participants showed greater improvement for the *HA* condition than the *LA* and Control conditions, as indexed by larger effect sizes. For Participant 002, this difference rested on a relatively small absolute gain for *HA* items (+10%) in the context of decreasing accuracy for the *LA* and Control items (-13% and -14%, respectively). Although the reasons for the declining accuracy for *LA* and Control sets are unclear, it does appear that a relatively large amount of practice (*HA*) was able to overcome this negative trend. For Participant 005, all sets did show improvement, but improvement was greatest in the *HA* condition as well. Thus, across both participants, a greater amount of practice resulted in greater improvement in effect size and percent of change. These results are broadly consistent with the principles of motor learning literature (speech and nonspeech), suggesting greater practice leads to greater learning (e.g., Park & Shea, 2003, 2005; Shea & Kohl, 1991; Kwakkel et al.,

2004; Cherney et al., 2008; Varley et al., 2016; Edeal & Gildersleeve-Neumann, 2011; Namasivayam et al., 2015). As this study did not confound practice amount and distribution (as in Namasivayam et al., 2015), gains observed under the *HA* condition cannot be ascribed to confounding effects from a more intensive practice schedule.

It should be noted however, that the changes and effects are relatively modest and did not result in (near) 100% accuracy. Furthermore, despite being treated, targets in the *LA* condition did not show clear improvement that exceeded improvement for untreated Control targets, contrary to the second prediction. Although the effect size for *LA* items was larger than that for Control items for Participant 002, this is attributable largely to lower variability during baseline for the Control items and does not reflect a substantial difference in absolute gain. Taken together, it appears that the amount of practice provided in the *LA* condition was insufficient to induce clear gains over and above no treatment. This observation suggests that there may be a minimal amount of practice that is required before improvements can be expected under this approach.

Various reasons may account for the relatively limited gains in this study, as compared to previous studies using similar integral stimulation treatment approaches for CAS (e.g., Edeal & Gildersleeve-Neumann, 2011; Strand et al., 2006). One reason may in fact relate to practice amount, in that the overall amount of practice provided in the present study was less than in previous studies. For example, children in Strand et al. (2006) received two 30-minute sessions per day, five days per week, for six consecutive weeks (i.e., 30 hours), with relatively small and consistent target sets (~5 targets). Edeal and Gildersleeve-Neumann (2011) included two children, one of whom received three

weekly 40-minute treatment sessions over ten weeks (~20 hours; ~10 hours per condition) and one of whom received two weekly 50-minute sessions over five weeks (< 10 hours; < 5 hours per condition). Edeal and Gildersleeve-Neumann (2011) noted that the former child made greater gains than the latter, consistent with the notion that practice amount influences learning in children with CAS. By contrast, children in the present study received only eight sessions (two weekly 50-minute sessions over four weeks), for a total of ~6.5 hours (~3 hours per condition). It is possible that greater gains, and greater condition differences, would have emerged with more practice.

Second, this study employed a rigorous study design, with a number of methodological factors that could account for the lack of stronger and clearer effects, including probe procedures, analysis procedures, and target selection criteria. With respect to probe procedures, this study used delayed production tasks. In treatment, Participants 002 and 005 did not reach the delayed elicitation method for teaching episodes until the end, and even then not consistently. As such, they did not have much opportunity to practice delayed imitation/reading tasks. Therefore, these tasks may have been too novel and complex, and may not constitute an accurate reflection of their learning. Probes were also obtained at the beginning of subsequently practiced sessions, rather than the end of practiced sessions as in some previous studies (e.g., Edeal & Gildersleeve-Neumann, 2011; Strand & Debertine, 2000; Strand et al., 2006). Probes administered at the end of a treatment session may reflect performance enhancements that are temporary in nature (Maas et al. 2008) and thus may inflate estimates of the amount of learning (retention). Further, probes in Strand et al. (2006) included each trained item

ten times, possibly allowing for within-probe performance improvements. In the present study, each item occurred only once on each probe, giving children only one opportunity to produce the target accurately.

With respect to analysis procedures, probe data in the present study were analyzed by raters blinded to treatment status and time of probe, which reduces potential bias and effects of perceptual drift. In some previous studies, probes were scored by the treating clinicians as the study progressed (e.g., Edeal & Gildersleeve-Neumann, 2011; Strand et al., 2006). Further, Edeal and Gildersleeve-Neumann (2011) only scored accuracy of treated sounds in words, rather than scoring accuracy of the entire word or phrase (as in the present study). Achieving accuracy on a single sound is arguably a lower bar than achieving accuracy on all sounds and suprasegmental aspects such as prosody.

Finally, targets for the children in this study were more complex (including word and phrase levels) compared to previous work (e.g., Edeal & Gildersleeve-Neumann, 2011; Strand et al., 2006). This was largely because targets were selected on the basis of functional relevance to enhance motivation and not based on prior phonetic inventory or speech sound developmental skill. As a result, it is possible that targets may have been too far outside of the participants' ability/skill level and/or zone of proximal development (Vygotsky, 1978; 1987), which is theorized to support children in successfully completing tasks that otherwise would not be possible without support from an individual with greater competence in their environment. This may have caused frustration rather than enhancing motivation.

Clinical Implications

Greater learning was evident for higher amounts of practice in both participants, although effects were modest. Further, it may be the case that for some children a small amount of integral stimulation treatment is no different than receiving no treatment whatsoever. Considering the service delivery model of typical clinical settings (e.g., outpatient, school-based, etc.), a majority of children with CAS likely receive low amounts of treatment. Knowing that amount matters (as shown here and elsewhere, e.g., Edeal & Gildersleeve-Neumann, 2011), an argument can be made for implementing very high amounts of practice in order to show significant improvement and learning. Moving forward, SLPs may take this notion into consideration by examining how to best optimize number of trials per session. For example, by choosing a smaller set of items, SLPs can increase the amount of practice on those items. Given that this reduces the number of items that can be targeted, it would be important to select targets that are functionally relevant and may be phonetically achievable, in order to optimize motivation. Once these items become established in a child's inventory, new items may be added and practiced further using a similar structure. This type of session design may yield more mastery and success rather than focusing on a wider set of items that potentially yield limited to no progress.

Considering the service delivery model of typical clinical settings (e.g., outpatient, school-based, etc.), demands of third-party payers, high caseloads, and funding constraints, future research is warranted to determine the optimal practice

amount in consideration with other principles of motor learning required to best serve this population in real-world conditions.

CHAPTER 6

CONCLUSION

The purpose of this study was to examine how practice amount, a key principle of motor learning, affects maintenance of speech motor learning. This study also served as a replication examining the effectiveness of integral stimulation approach to CAS treatment using a single-case experimental design with two participants. Findings from the present study should be interpreted with caution given the small sample size and modest effects, but show some evidence suggesting that a higher amount of practice yields greater learning in children with CAS. For both participants in this study, treatment using a low amount of practice did not seem to yield improvements in comparison to receiving no treatment. This important implication suggests that integral stimulation treatment, at least as implemented and studied here, may only be truly effective if provided in the context of high amounts of practice. To best support optimal treatment outcomes for children with CAS, further high-quality research (including more participants and well-controlled study designs) are needed to examine how the principles of motor learning should be realistically applied and can be optimally generalized to real-world settings.

REFERENCES CITED

- American Speech-Language-Hearing Association. (2007). Childhood Apraxia of Speech [Technical Report]. Available at www.asha.org/policy.
- American Speech-Language-Hearing Association. (2012). 2012 Schools Survey report: SLP caseload characteristics. Available at www.asha.org/research/memberdata/schoolssurvey/
- Allen, M. M. (2013). Intervention efficacy and intensity for children with speech sound disorder. *Journal of Speech, Language, and Hearing Research, 56*, 865-877.
- Baas, B. S., Strand, E. A., Elmer, L. M., & Barbaresi, W. J. (2008). Treatment of severe childhood apraxia of speech in a 12-year-old male with CHARGE association. *Journal of Medical Speech-Language Pathology, 16*(4), 181-191.
- Ballard, K. J., Maas, E., & Robin, D. A. (2007). Treating control of voicing in apraxia of speech with variable practice. *Aphasiology, 21*(12), 1195-1217.
- Beeson, P. M., & Robey, R. R. (2006). Evaluating single-subject treatment research: Lessons learned from the aphasia literature. *Neuropsychology Review, 16*, 161-169.
- Caruso, A. J., & Strand, E. A. (Eds.). (1999). *Clinical management of motor speech disorders in children* (pp. 21-24). New York: Thieme.
- Cherney, L. R., Patterson, J. P., Raymer, A., Frymark, T., & Schooling, T. (2008). Evidence-based systematic review: Effects of intensity of treatment and constraint-induced language therapy for individuals with stroke-induced aphasia. *Journal of Speech, Language, and Hearing Research, 51*(5), 1282-1299.

- Davis, B., & Velleman, S. L. (2000). Differential diagnosis and treatment of developmental apraxia of speech in infants and toddlers. *Infant-Toddler Intervention: The Transdisciplinary Journal*, 10, 177–192.
- Delaney, A. L., & Kent, R. D. (2004, November). *Developmental profiles of children diagnosed with apraxia of speech*. Poster session presented at the annual convention of the American-Speech-Language-Hearing Association, Philadelphia.
- Dodd, B., Hua, Z., Crosbie, S., Holm, A., & Ozanne, A. (2006). *Diagnostic Evaluation of Articulation and Phonology*. San Antonio, TX: PsychCorp.
- Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test (4th Edition)*. AGS/Pearson.
- Edeal, D. M., & Gildersleeve-Neumann, C. E. (2011). The importance of production frequency in therapy for childhood apraxia of speech. *American Journal of Speech-Language Pathology*, 20, 95-110.
- Goldman, R., & Fristoe., M. (2000). *Goldman-Fristoe 2 Test of Articulation*. Minneapolis, MN: Pearson Assessments.
- Gillon, G.T., & Moriarty, B.C. (2007) Childhood apraxia of speech: children at risk for persistent reading and spelling disorder. *Seminars in Speech and Language*, 28(1), 48–57.
- Guenther, F. H., Ghosh, S. S., & Tourville, J. A. (2006). Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain and Language*, 96, 280–301.

- Hula, S. N. A., Robin, D. A., Maas, E., Ballard, K. J., & Schmidt, R. A. (2008). Effects of feedback frequency and timing on acquisition, retention, and transfer of speech skills in acquired apraxia of speech. *Journal of Speech, Language, and Hearing Research, 51*(5), 1088-1113.
- Kwakkel, G., van Peppen, R., Wagenaar, R. C., Dauphinee, S. W., Richards, C., Ashburn, A., ... & Langhorne, P. (2004). Effects of augmented exercise therapy time after stroke a meta-analysis. *Stroke, 35*(11), 2529-2539.
- Law, J., Boyle, J., Harris, F., Harkness, A., & Nye, C. (2000). Prevalence and natural history of primary speech and language delay: Findings from a systematic review of the literature. *International Journal of Language and Communication Disorders, 35*(2), 165-188.
- Lee, T. D., Magill, R. A., & Weeks, D. J. (1985). Influence of practice schedule on testing schema theory predictions in adults. *Journal of Motor Behavior, 17*(3), 283-299.
- Lewis, B.A., Freebairn, L.A., Hansen, A.J., Iyengar, S.K., & Taylor, H.G. (2004). School-age follow-up of children with childhood apraxia of speech. *Language, Speech, and Hearing Services in Schools, 35*(2), 122-140.
- Maas, E., Butalla, C.E., & Farinella, K.A. (2012). Feedback frequency in treatment for childhood apraxia of speech. *American Journal of Speech-Language Pathology, 21*, 239-257.

- Maas, E., & Farinella, K. A. (2012). Random versus blocked practice in treatment for childhood apraxia of speech. *Journal of Speech, Language, and Hearing Research, 55*, 561-578.
- Maas, E., Robin, D. A., Austermann Hula, S. N., Freedman, S. E., Wulf, G., Ballard, K. J., & Schmidt, R. A. (2008). Principles of motor learning in treatment of motor speech disorders. *American Journal of Speech-Language Pathology, 17*, 277-298.
- McLeod, S., Harrison, L. J., & McCormack, J. (2012). The intelligibility in context scale: Validity and reliability of a subjective rating measure. *Journal of Speech, Language, and Hearing Research, 55*(2), 648-656.
- Murray, E., McCabe, P., & Ballard, K. J. (2014). A systematic review of treatment outcomes for children with Childhood Apraxia of Speech. *American Journal of Speech-Language Pathology, 23*, 486-504.
- Namasivayam, A. K., Pukonen, M., Goshulak, D., Hard, J., Rudzicz, F., Rietveld, T., Maassen, B., Kroll, R., & van Lieshout, P. (2015). Treatment intensity and childhood apraxia of speech. *International Journal of Language and Communication Disorders, 50*, 529-546.
- Park, J. H., & Shea, C. H. (2003). Effect of practice on effector independence. *Journal of Motor Behavior, 35*, 33-40.
- Park, J. H., & Shea, C. H. (2005). Sequence learning: Response structure and effector transfer. *Quarterly Journal of Experimental Psychology, 58A*, 387-419.
- Reynolds, C. R., & Kamphaus, R. W. (2003). *Reynolds Intellectual Assessment Scales (RIAS)*. Lutz, FL: Psychological Assessment Resources, Inc.

- Robbins, J., & Klee, T. (1987). Clinical assessment of oropharyngeal motor development in young children. *Journal of Speech and Hearing Disorders, 52*, 271-277.
- Robin, D. A., Maas, E., Sandberg, Y., & Schmidt, R. A. (2007). Motor control and learning and childhood apraxia of speech. In P. Hall, L. Jordan, & D. Robin (Eds.), *Developmental Apraxia of Speech: Theory and Clinical Practice (2nd Edition)* (pp. 67-86). Austin, TX: Pro-Ed.
- Rosenbek, J. C., Lemme, M. L., Ahern, M. B., Harris, E. H., & Wertz, R. T. (1973). A treatment for apraxia of speech in adults. *Journal of Speech and Hearing Disorders, 38*(4), 462-472.
- Rvachew, S., Hodge, M., & Ohberg, A. (2005). Obtaining and interpreting maximum performance tasks from children: A tutorial. *Journal of Speech-Language Pathology and Audiology, 29*, 146-157.
- Schmidt, R. A. (1975). A schema theory of discrete motor skill learning. *Psychological Review, 82*, 225-260.
- Schmidt, R. A. (2003). Motor schema theory after 27 years: Reflections and implications for a new theory. *Research Quarterly for Exercise and Sport, 74*, 366-375.
- Schmidt, R. A., & Lee, T. D. (2005). *Motor Control and Learning: A Behavioral Emphasis (4th edition)*. Champaign, IL: Human Kinetics.
- Semel, E., Wiig, E. H., & Secord, W. A. (2003). *Clinical Evaluation of Language Fundamentals-Fourth edition*. Toronto, Canada: The Psychological Corporation.
- Shea, C. H., & Kohl, R. M. (1991). Composition of practice: Influence on the retention of motor skills. *Research Quarterly for Exercise and Sport, 62*, 187-195.

- Shriberg, L. D., Aram, D. M., & Kwiatkowski, J. (1997). Developmental apraxia of speech: I. Descriptive and theoretical perspectives. *Journal of Speech, Language, and Hearing Research, 40*, 273-285.
- Shriberg, L. D., & Strand, E. A. (February, 2014). *A diagnostic marker to discriminate childhood apraxia of speech from speech delay*. Paper presented at the Seventeenth Biennial Conference on Motor Speech: Motor Speech Disorders & Speech Motor Control, Sarasota, FL.
- Smith, N. L. (1981). The certainty of judgments in health evaluations. *Evaluation and Program Planning, 4*(3), 273–278.
- Strand, E. (1995). Treatment of motor speech disorders in children. *Seminars in Speech and Language, 16*, 126–139.
- Strand, E., & Debertine, P. (2000). The efficacy of integral stimulation intervention with developmental apraxia of speech. *Journal of Medical Speech-Language Pathology, 8*(4), 295-300.
- Strand, E. A., McCauley, R. J., Weigand, S. D., Stoeckel, R. E., & Baas, B. S. (2013). A motor speech assessment for children with severe speech disorders: Reliability and validity evidence. *Journal of Speech, Language, and Hearing Research, 56*(2), 505-520.
- Strand, E. A., Stoeckel, R., & Baas, B. (2006). Treatment of severe childhood apraxia of speech: A treatment efficacy study. *Journal of Medical Speech-Language Pathology, 14*, 297- 307.

- Thomas-Stonell, N., Washington, K., Oddson, B., Robertson, B. & Rosenbaum, P. (2013). Measuring communicative participation using the FOCUS: Focus on the Outcomes of Communication Under Six. *Child: Care, Health, and Development*, 39, 474–480.
- Thoonen, G., Maassen, B., Gabreëls, F., & Schreuder, R. (1999). Validity of maximum performance tasks to diagnose motor speech disorders in children. *Clinical Linguistics & Phonetics*, 13, 1-23.
- Thoonen, G., Maassen, B., Wit, J., Gabreëls, F., & Schreuder, R. (1996). The integrated use of maximum performance tasks in differential diagnostic evaluations among children with motor speech disorders. *Clinical Linguistics and Phonetics*, 10, 311-336.
- Varley, R., Cowell, P. E., Dyson, L., Inglis, L., Roper, A., & Whiteside, S. P. (2016). Self-Administered Computer Therapy for Apraxia of Speech Two-Period Randomized Control Trial With Crossover. *Stroke*, STROKEAHA-115.
- Vygotsky, L.S. (1978). *Mind in society*. (Cole, A., John-Steiner, V., Scriber, S., & Souberman, E. , Eds.). Cambridge, MA: Harvard University Press. (Original work published 1930).
- Vygotsky, L.S. (1987). *Thought and language*. (A. Kozulin, Ed.). Cambridge, MA: MIT Press. (Original work published 1934).
- Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). *Comprehensive Test of Phonological Processing*. Austin, TX: Pro-Ed.

- Williams, K. (2007). *Expressive Vocabulary Test (2nd Edition)*. Circle Pines, MN: American Guidance Services.
- Wilson, A., & Gildersleeve-Neumann, C. E. (2014). *Functional Communication Parent Questionnaire*. Unpublished instrument, Portland State University.
- World Health Organization. (2001). International classification of functioning, disability, and health (icf). Geneva, Switzerland: World Health Organization.
- Wright, D. L., Black, C. B., Immink, M. A., Brueckner, S., & Magnuson, C. (2004). Long-term motor programming improvements occur via concatenating movement sequences during random but not blocked practice. *Journal of Motor Behavior*, 36, 39-50.
- Wulf, G., & Schmidt, R. A. (1997). Variability of practice and implicit motor learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23(4), 987.

APPENDIX A

TREATMENT PROTOCOL

Teaching episodes will be presented in random order for both conditions. There will be 5 targets for the massed condition and 10 targets for the distributed condition in each treatment phase. A card deck will contain 5 stimulus cards for each target word per condition. Random order will be implemented by shuffling the cards before the treatment session and reshuffling the deck as needed during the session once every card has been used. Steps for each teaching episode are as follows:

1. Start teaching episode by taking stimulus card from deck.
2. Elicit first attempt using one of the following elicitation methods (see below for criterion to switch between elicitation methods):
 - Immediate imitation, slow rate (“watch me, listen carefully, and say ____.”)
 - Immediate imitation, normal rate (“watch me, listen carefully, and say ____.”)
 - Delayed imitation, normal rate (“watch me, listen carefully, and say ____.” Point to child after 3 seconds to cue delayed imitation)
 - Spontaneous production (Show picture associated with target and say “What is this?”)
3. Child responds (*count as attempt #1*)

If production is *correct*:

4. Wait 2-3 seconds
 - a. provide feedback, if indicated for this trial (e.g., “good one!” “that was right”)
 - b. don’t provide feedback, if not indicated for this trial. Instead provide general encouragement (e.g., “Keep doing the best you can”)
5. Elicit another attempt (e.g., “can you say that again?”)
6. Child responds (count as attempt #2)
7. Wait 2-3 seconds
 - a. provide feedback, if indicated for this trial.
 - b. don’t provide feedback, if not indicated for this trial; general encouragement.
8. Place a check mark on the feedback tracking sheet (see below)
9. Go to Step 1 with next target to start next teaching episode (e.g., “Let’s do another one”).

If production is *incorrect*:

4. Wait 2-3 seconds
 - a. provide feedback, if indicated for this trial (e.g., “good try, but not quite”, “close!”, “your mouth was too open”)
 - b. don’t provide feedback, if not indicated for this trial. Instead provide general encouragement (e.g., “Keep doing the best you can”)
5. Provide appropriate support (e.g., slow simultaneous production, tactile/gestural cue)
6. Elicit another attempt using same elicitation method as for attempt #1
7. Child responds (count as attempt #2).
8. Wait 2-3 seconds
 - a. provide feedback, if indicated for this trial.
 - b. don’t provide feedback, if not indicated for this trial; general encouragement.
9. Place a check mark on the feedback tracking sheet (see below)
10. Go to Step 1 with next target to start next teaching episode (e.g., “Let’s do another one”).

Continue until a total of 100 trials have been completed or one half of the session time has elapsed. After a brief break, begin the second treatment condition and continue until a total of 100 trials have been completed or the session is ended.

Feedback is given with decreasing frequency during each condition. Feedback is provided for 100% (10) of the first 10 teaching episodes, 90% (9) of the next 10 teaching episodes, and so on until feedback is given on 10% (1) of the last 10 teaching episodes. To facilitate keeping track of feedback schedules, the SLP will have a feedback tracking sheet with 100 slots with the fading feedback schedule marked (cf. Ballard et al., 2007). By placing a check mark in a slot for each teaching episode, the SLP can easily see whether the next trial requires feedback or not.

Start with immediate imitation at normal rate. The criterion to increase difficulty level of elicitation condition is 2/2 consecutive correct on the first attempt for a given target within a session. That is, if the child produces target A correctly after the first elicitation on two consecutive teaching episodes of target A (separated by other targets or not), then the third teaching episode of target A will be elicited with the next level elicitation method. Similarly, if the child produces a target incorrectly on 2/2 consecutive teaching episodes of target A with an elicitation method, the next teaching episode of target A will revert to the previous difficulty level of elicitation method (or stay in immediate imitation at slow rate).

APPENDIX B

PROTOCOL FLOWCHART

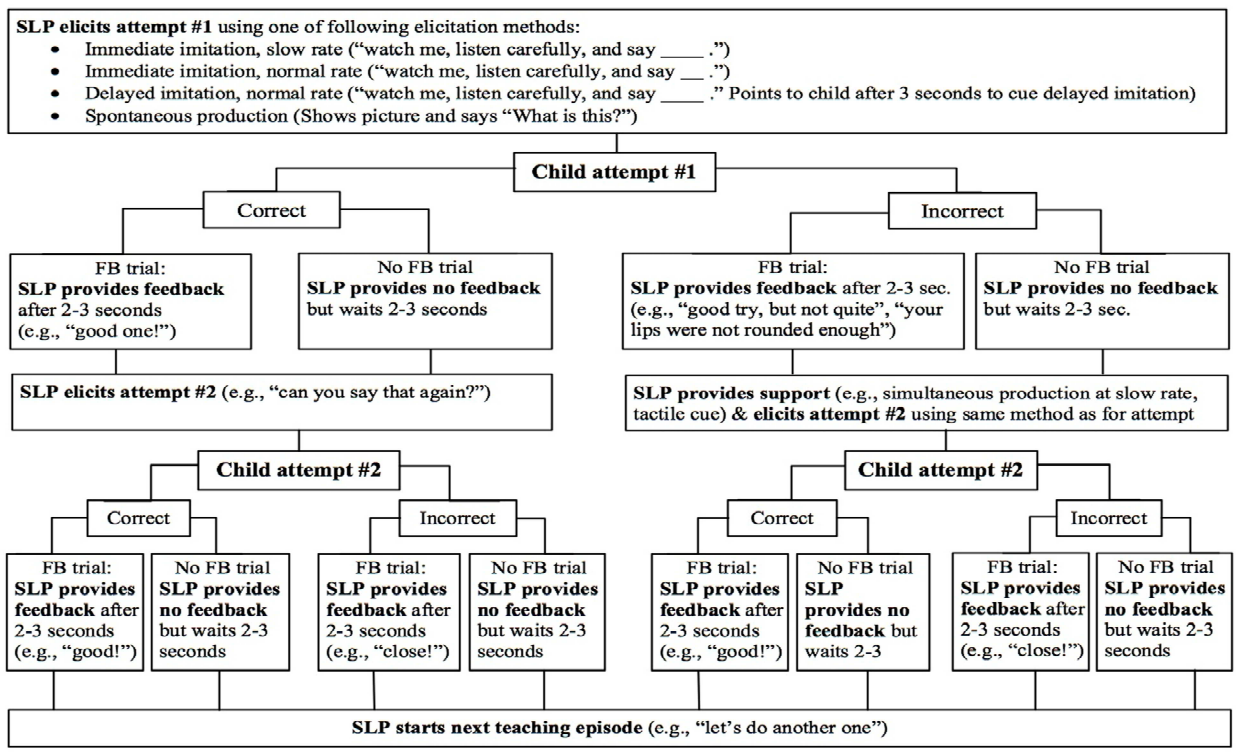


Figure 4. Protocol Flowchart.

APPENDIX C

PROBE SCORING INSTRUCTIONS

OVERVIEW

The task is to score the accuracy of productions of treatment targets, untreated transfer items, and untreated Control items. These accuracy scores constitute the primary data to determine whether the treatment has an effect and whether one condition leads to greater improvement than another. This scoring must be completed by blinded analysts, and can therefore only begin once all data are available for a child (i.e. after the final follow-up probe), so that probes can be randomized for time point. This means that as an analyst, you should score the probes on your list in the order in which they are given, and it means you should avoid learning which targets were practiced in treatment.

THE SCORING SYSTEM

Background & Overview

The scoring system has to meet several goals:

(a) It has to be valid. This involves two things: (1) it has to be able to differentiate minor from major errors, weighted to some extent by target complexity, and (2) it has to be sensitive to change – i.e. It has to be able to detect changes (e.g., due to treatment), including for targets of different complexity levels. For example, a minor error on a 5-syllable phrase is less serious than a minor error on a 1-syllable word.

(b) It has to be reliable. Different raters should be able to agree, and the same rater should be consistent with her/himself. This also relates to replicability of research findings. The detailed instructions below are intended to ensure reliability. (c) It has to be (at least somewhat) clinically relevant.

The scoring system described below meets these criteria. This scoring system is adapted from Strand et al. (2006), and was also used in Maas et al. (2012) and Maas and Farinella (2012), but adapted to account for different target lengths/complexities (operationalized here in terms of number of syllables). NOTE that this involves judging accuracy (not intelligibility) of the child's production.

The system includes segmental and prosody scoring components, both of which involve a binary scale (correct/incorrect). Briefly, for Segmental scoring, raters judge the accuracy of each syllable in a target, and for Prosody scoring, raters judge the prosodic accuracy of the entire target word or phrase. Segmental accuracy of the entire target word or phrase is then scaled to a 3-point scale (correct, minor errors, major errors) according to the criteria outlined in Table 1 below; this is accomplished automatically in Excel using a formula so raters do not need to apply those criteria. Combining the Segmental and Prosodic scores, each target has a maximum possible score of 3 (2 points for segmental accuracy + 1 point for prosodic accuracy). Each scoring component is explained in further detail below, followed by some general guidelines and examples.

Table 6		
<i>Segmental score criteria</i>		
Score	Description	Criteria
2	Correct	No segmental errors in any syllable in any word.
1	Minor Error	≥ 50% but < 100% of syllables correct
0	Major Error	< 50% of syllables correct

Segmental Scoring

For Segmental Scoring, a score is derived based on the proportion of syllables with (one or more) segmental errors (defined in Table 2 below). The criteria for the Segmental Score are given in Table 1 above; these criteria are designed to be applicable to both single-word targets and multi-word targets (but note that monosyllabic targets can only receive a 0 or 2). A further rationale for aspects of this system is provided below.

Raters do not need to apply the criteria in Table 1; these scores are calculated automatically in the Excel sheet. Instead, raters simply indicate for each syllable whether it was correct or not (in terms of segmental accuracy). Do not distinguish between major and minor segmental errors, nor classify error types. The judgment is only whether the syllable contains one or more of the listed errors or not. If a syllable is correct, enter 1 in the relevant cell in the spreadsheet; if a syllable has any of the listed errors, enter 0 (see example in Figure 1). Based on these judgments, a formula in the Excel file then automatically applies the score criteria given in Table 1, as well as a percentage of syllables correct (see cell U2 in Figure 1).

Table 7		
<i>Segmental error types (judged for each syllable) and explanations</i>		
Error Type	Description	Notes
Substitution	Consonant or vowel that sounds like an example of a different phoneme.	May be distorted but sounds like a phoneme that is not the target.
Omission	Missing sound(s)	Vowel and whole-syllable omissions also affect the Prosody Score.

Table 7		
<i>Continued</i>		
Error Type	Description	Notes
Distortion	A consonant or vowel that is recognizable as the target sound but is not clear or well-articulated.	May sound like a sound between the target and some other sound, but not always (e.g., a lateral lisp). This includes excessive aspiration of plosives, derhoticization of rhotic vowels/diphthongs and consonants, and excessive lengthening of consonants, Vowel lengthening is not scored for the Segmental score but affects the Prosody score if it affects stress.
Addition	A syllable contains a consonant addition.	Scored under Prosody Score.
Metathesis	Two consonants are presented in an incorrect order within a syllable	Metathesis between syllables is counted as 2 syllables in error (e.g., /kato/ for 'taco' is 2 syllables with errors). Metathesis within a syllable is counted as only 1 syllable in error (e.g., /supənæm/ for 'Superman' is 1 syllable with errors). Whole-syllable or whole-word metatheses are scored as 0 in Prosody Scoring. For Segmental Scoring, treat each syllable as if appeared in the correct position
Unintelligible	The syllable is unable to be perceptually matched to any target word form, it is unrecognizable.	

Rationales

The rationale for not distinguishing between major and minor errors (or error types) is as follows:

(a) This distinction tends to have poor reliability, even if raters agree on presence of error: Two raters may agree there is an error, but one may consider it major (substitution) and the other minor (distortion). This is true even for otherwise very good and reliable raters.

(b) This distinction is onerous for raters. To achieve acceptable reliability for this distinction, the instructions and rating system become unwieldy, leading to excessively long rating sessions and associated frustration of raters (leading to possible lapses of attention). This further means reduced ecological validity / clinical feasibility.

The rationale for counting the number of syllables with segmental errors rather than counting the total number of segmental errors (per syllable or total) is as follows:

(a) Theoretically, segments may not constitute the primary unit of speech motor planning. Instead, many models assume a syllable-sized unit (e.g., Guenther et al., 2006)

(b) Related to (a), intra-syllabic coarticulation is especially strong, and thus it is expected that an error on one segment may also affect production of nearby segments.

(c) Related to (b), it is often difficult to determine whether an error affects only one segment or multiple segments within a syllable. In keeping with the desire to keep the system manageable (and reduce time and frustration on our raters' part), we simply count the number of syllables with (one or more) segmental errors.

The rationale for relating the number of syllables with segmental errors to number of syllables per target utterance is that this approach enables gradient scores that take into account target length. Although this reduces sensitivity for monosyllabic targets (can only be 2 or 0), there are very few monosyllabic targets for any participant, so this is unlikely to have a large impact.

The rationale for setting the criterion for a score of 1 at $\geq 50\%$ is to maintain gradience for two-syllable words (i.e., if one of the syllables is correct, this gets a score of 1 rather than 0).

The rationale for counting missing vowels in both the Segmental and the Prosody Score is as follows:

(a) These errors do affect both and can be considered more serious than the other errors

(b) It minimizes confusion of scoring contingencies (e.g., if a syllable is missing, then how does one judge prosodic accuracy? If missing syllables are scored only in Prosody Scoring, then how many syllables are considered in error for the Segmental Score?)

Prosody Scoring

For Prosody Scoring, we only score correct (1) vs. incorrect (0), based on presence of the errors and abnormalities listed in Table 3 below. Raters judge stress patterns, correct number of syllables, and fluency (syllable repeats, syllable segmentation). In addition, we include self-corrections here as well. A further rationale for this system is provided below.

A couple of preliminary points:

- For monosyllabic targets, there can be no stress pattern errors or segmentation.

For these targets, the Prosody Score is based on presence/absence of self-corrections and added syllables.

- Voice quality is not addressed in the Prosody Score. If a child has a consistently breathy or harsh voice, this likely is either reflective of a comorbid dysarthria or idiosyncratic speaking style, neither of which are addressed in treatment. NOTE: The same applies to consistently altered resonance (e.g., consistently hyponasal) and vocal loudness (e.g., reduced loudness).

A judgment of correct prosody requires a production to be continuous (not choppy) with adequate stress patterning – on the first attempt (basically, a production free of the errors listed in Table 3). Score 1 if prosody is correct, and 0 if any of the errors listed below are present. A formula in the Excel file then automatically computes the total score.

Rationale

The rationale for this approach is as follows:

(a) Suprasegmental scoring with the 0-1-2 system turned out to have poor reliability, largely due to difficulty in distinguishing major from minor errors as defined in a previously attempted system. However, in most cases raters agreed on presence vs. absence of error.

(b) Rating of speaking rate is particularly difficult and subject to individual variation (both on the speaker's part and on the listener's part).

(c) Binary scoring is expected to be more efficient and less frustrating to raters

Table 8		
<i>Prosodic error types and explanations</i>		
Error Type	Description	Notes
Segmentation	Pauses between sounds/syllables (“choppy” speech).	Score 0 regardless of whether all or only some syllables are segmented.

Table 8		
<i>Continued</i>		
Error Type	Description	Notes
Omitted syllable(s)	One or more syllables are missing.	Missing syllables affect prosody, BUT normal contractions that result in an omitted vowel/syllable should not be penalized because those are not considered true omissions. Score both syllables correct unless there are errors.
Added syllable(s)	One or more syllables are added.	Vowel additions result in added syllables, which affect prosody. This also includes partial or whole-syllable repetitions (disfluencies). For Segmental Scoring: If the added syllable is in the middle of the word, give the child the benefit of the doubt by scoring the syllable that is segmentally accurate as the target, regardless of position
Syllable reversals	Two or more syllables are switched.	For Segmental Scoring, score each syllable as though it appeared in correct position.
Reduced/equalized stress	Reduced differentiation between stressed and unstressed syllables	This can be the result of vowel lengthening or shortening, alterations of loudness, and/or changes in pitch.
Exaggerated stress	Correct stress pattern but exaggerated	This can be the result of vowel lengthening or shortening, alterations of loudness, and/or changes in pitch.

Table 8		
<i>Continued</i>		
Error Type	Description	Notes
Incorrect stress	Reversed stress pattern-not simply equalized or exaggerated (stressed becoming unstressed, and vice versa).	This can be the result of vowel lengthening or shortening, alterations of loudness, and/or changes in pitch.
Self-corrections	Child initiates another unprompted attempt.	If the clinician elicits another attempt, it is likely due to inattention or background noise, etc. In this case, score the newly-elicited attempt.

Note. If any of these are present, Prosody Score = 0.

General Scoring Rules & Procedures

1. Generally, judge the first attempt by a child, unless the child produces a spontaneous self-correction (i.e. not prompted by the clinician). Spontaneous self-corrections are scored incorrect for Prosody Scoring but are scored the same as first attempts with respect to segmental accuracy (you only score segmental accuracy for self-corrections because the prosody score is already 0). □
2. If targets are elicited via a repetition task, then judgment is relative to the model presented by the clinician. In other words, if the clinician uses abnormal prosody and the child imitates that, this is not an error on the child's part. Be sure to make a note if this happens. □
3. Dialectal variations or normal speech phenomena are not considered errors. Examples:
 -
 - o Some vowel reduction in unstressed syllables is normal.
 - o Tapping of intersyllabic /t/ or /d/ is normal. □
 - o Vowel nasalization before a nasal coda is normal.
 - o Unreleased stop at the end of a word is normal, as is unreleasedness of the first stop in a two-stop sequence (e.g., 'act').
 - o Contractions are common in typical speech and so should not be penalized (e.g., 'It's' instead of 'It is' is OK), assuming no other errors: in other words, score both syllables as separate target syllables. Conversely, contractions being produced in full should not be penalized either, assuming no other errors (e.g. 'do not' instead of 'don't' is OK). In other words, score both syllables as one target syllable (both syllables as produced must then be correct).

o For the purpose of this scoring system, consider syllable-final rhotics to be part of a rhotic diphthong or rhotic monophthong rather than as consonants or as added vowels. That is, the word ‘where’ would be considered 1 syllable (/weə/) in which the /ɛə/ sequence is considered a diphthong (1 syllable). Words like ‘sister’ are considered to end in a rhotic monophthong (/sɪstə/). Deduct points for syllable additions (score as 0 on Prosody Scoring; see below) only if a child produces this with two clearly distinguishable syllables to the point where it no longer sounds normal (e.g., /we.ə/ for ‘where’). Keep in mind that a changing quality is inherent because they are considered diphthongs (which may sound like an added “half-syllable”, as for vowels preceding syllable-final lateral /l/ (e.g. ‘feel’ /fiəl/ or regular, non-rhotic diphthongs such as /ɔɪ/ in ‘toy’).

o There are many other examples of typical alterations of sounds that happen as normal speech phenomena, and these should not be considered errors. Use knowledge of normal, non-disordered speech and its contextual variations to make these judgments.

4. For Segmental Scoring, use the Sonority Sequencing Principle and Onset Maximization Principle to determine syllable boundaries (and thus whether one or two syllables contain errors).

o Sonority Sequencing Principle (SSP): Syllables prefer a rise in sonority toward the nucleus and a flat or declining sonority slope following the nucleus.

o Onset Maximization Principle (OMP): Assign as many consonants to the onset of a syllable as possible (i.e. without violating the SSP).

o NOTE: When added or deleted syllables result in resyllabification, score these segments as part of the target syllable nevertheless.

5. Make notes about observations that may be relevant.

o If you are unsure about a score, please make a note to describe the error and/or what was confusing. These notes will help us resolve any discrepancies that may arise.

o Errors that could be morphological or syntactic errors. For example, if a child says ‘I can play with the iPad’ for the target ‘Can I play with the iPad?’, the syllable reversal could be a language error. Similarly, the syllable omission in ‘What you want to drink?’ for ‘What do you want to drink?’ could be a language error. Nevertheless score these as speech errors (0 on Prosody Scoring; see below) and add a comment (e.g., LANG?) to separate these types of errors if needed.

o Clinician model differs from intended target (e.g., clinician says ‘Miss B.’ for ‘Mrs. B.’).

o Anything that might affect validity of the scores. For example, background noise, poor recording quality, silly voice/yawning/giggling/crying, etc.

#	Production	Syll. 1	Segment Score	Prosody Score	Total Score	Comments
1	pliz	1	2	1	3	Correct
2	(plid-) pliz	1	2	0	2	Segment Score = 2 because no errors Prosody score = 0 because self-correction
3	plid	0	0	1	1	Segment Score = 0 because < 50% of syllables correct (substitution error) Prosody Score = 1 because no self-correction and no added syllables
4	pidə	0	0	0	0	Segment Score = 0 because < 50% of syllables correct (omission & substitution error) Prosody Score = 0 because added syllable
5	plidə	0	0	0	0	Segment Score = 0 because < 50% of syllables correct (substitution error) Prosody Score = 0 because added syllable NOTE: due to added syllable, target /z/ is now part of added syllable. Score /z/ → /d/ as segmental error for syllable 1.
6	plizə	1	2	0	2	Segment Score = 2 because no errors (resyllabification of /z/ is not a segmental error for syll. 1) Prosody Score = 0 because added syllable