# THE ACTIVE INGREDIENTS OF INTEGRAL STIMULATION TREATMENT: THE EFFICACY OF AUDITORY, VISUAL, AND AUDITORY-VISUAL CUES FOR TREATMENT OF CHILDHOOD APRAXIA OF SPEECH

---

A Thesis
Submitted to
the Temple University Graduate Board

---

In Partial Fulfillment
of the Requirements for the Degree
MASTER OF ARTS
In SPEECH, LANGUAGE, AND HEARING SCIENCES

---

by

Lauren Condoluci, B.A.
Diploma Date December 2020

Thesis Approvals:

Edwin Maas, Thesis Advisor, Department of Communication Sciences and Disorders
Francine Kohen, Department of Communication Sciences and Disorders
Susan Caspari, Department of Communication Sciences and Disorders

## ABSTRACT

The purpose of this study was to determine the relative efficacy of cuing modalities employed in Integral Stimulation (IS) treatment for childhood apraxia of speech (CAS). Previous literature has supported the use of IS for children with CAS, though there are no studies that evaluate the active ingredients of IS. This study aimed to examine the efficacy of single- and multi-modality cues in IS treatment.

The experiment was administered as a single-case, alternating treatments design consisting of three conditions (auditory-only, visual-only, and simultaneous auditory and visual). Two participants with CAS received IS treatment in every condition during each session. Probes were administered prior to starting every other session (once per week), consisting of practiced and control targets that were balanced for complexity and functionality. Perceptual accuracy of productions was rated on a 3-point scale and standardized effect sizes were calculated for each condition.

Each participant demonstrated different effects in regard to modality and treatment effects. The visual-only condition yielded the greatest effect for one participant, followed by the auditory-only cues. The other participant displayed no significant effects in any condition nor a treatment effect.

The results of this study suggest that single-modality cues may be more beneficial for some children with CAS than the clinically used simultaneous auditory-visual multi-modality cue. The significant effect of the visual-only condition in one participant indicates that visual-only cues may bypass an impaired auditory feedback system and support speech motor learning, though more research is required.

## ACKNOWLEDGEMENTS

I would like to extend my sincere gratitude to my thesis advisor, Dr. Edwin Maas, for his dedication, positivity, and unwavering support. I am incredibly fortunate to have had the opportunity to learn from his scholarship and am thankful for the empathy and encouragement he has offered me throughout my time as a graduate student. His uplifting spirit and excitement about research is truly inspiring. This thesis would not have been possible without his knowledge and support.

I also wish to express my thanks to the phenomenal members of my thesis committee, Francine Kohen and Sue Caspari, whose mentorship throughout my time as a student has been truly invaluable. You've made me a better student and clinician, and I cannot thank you enough for your contributions to this research. My deepest acknowledgement is also extended to the study's clinician, participants, and their families, whose efforts made this thesis possible.

Additionally, I would like to thank all of my graduate professors and clinical supervisors. Your wealth of knowledge, feedback, and emphasis on the importance of evidence-based practice has shaped me into the clinician I am today.

Finally, my deepest appreciation is extended to my fiancé, Matt, and my incredible parents. Thank you for always encouraging me to step out of my comfort zone and showering me with never-ending love and support. Your tireless patience, reassurance, and profound belief in my abilities has impacted me beyond words. I never would have had the courage to pursue graduate school and research without you.

## TABLE OF CONTENTS

## LIST OF TABLES

**TABLE** **PAGE**

## LIST OF FIGURES

# CHAPTER 1

# INTRODUCTION

## Background

Childhood apraxia of speech (CAS) is a pediatric speech sound disorder in which a neurologic breakdown in speech-motor planning occurs (ASHA, 2007). This impairment in articulatory control may be the result of neurological damage, an underlying disorder, or of unknown origins, impacting speech accuracy and prosody. While the American Speech-Language-Hearing Association (ASHA) recognizes a lack of conformity in diagnostic criteria, CAS is often characterized among researchers and clinicians by inconsistent articulation of vowels and consonants, laborious articulatory transitions, and atypical prosody and lexical stress (ASHA, 2007). It is estimated that approximately 1-2 children per thousand classify as having CAS (Shriberg, Aram, & Kwiatkowski, 1997).

Studies have demonstrated the lasting effects of speech sound disorder (SSD), including CAS, on school and academic performance. Raitano, Pennington, Tunick, Boada, and Shriberg (2004) demonstrated that SSD and comorbid language deficits affected outcomes on phonological awareness and letter knowledge tasks, putting children at risk for increased literacy and reading difficulties. Furthermore, Lewis, Freebairn, Hansen, Iyengar, and Taylor (2004) noted that school-aged children with CAS consistently performed more poorly than their peers with SSD and language disorders on measures of expressive and receptive language, nonword repetition, syllable sequencing, and spelling despite enrollment in speech and language services. Children with CAS experienced poorer speech sound patterns during conversational speech, and were at greater risk for language, reading, and spelling difficulties (Lewis et al., 2004).

SSD's have also been highly associated with social and emotional limitations. An extensive literature review by McCormack, McLeod, McAllister, and Harrison (2009) revealed significant social challenges for children with SSD, including limitations with teacher-child relationships and peer relationships (Overby, Carrell, & Bernthal, 2007; Glogowska, Roulstone, Peters, & Enderby, 2006). Similarly, children with CAS have been found to engage less frequently with peers and in social activities due to difficulties with intelligibility (Rusiewicz, Maize, & Ptakowski, 2018).

These studies exhibit the importance of developing and refining treatment approaches to reduce the negative impacts of CAS on individuals' overall success. To date, children with CAS require more intensive and long-term treatment than those with other SSD (ASHA, 2007). It has been estimated that children with CAS require five times more individualized treatment sessions than those with severe SSD to establish functional outcomes (Campbell, 1999). According to the American Speech-Language-Hearing Association (2016), approximately 63% of school-based speech-language pathologists (SLPs) report having at least one child with CAS on their caseload. Furthermore, school-based caseloads include on average three children with CAS (ASHA, 2016). Thus, treatment research is essential to help those with CAS maximize gains and stabilize speech sound production.

A number of motor-oriented intervention approaches have been used to treat CAS, including Rapid Syllable Transition Treatment (ReST; Ballard, Robin, McCabe, & McDonald, 2010), and Integral Stimulation-based approaches such as Dynamic Temporal and Tactile Cueing (DTTC; Strand & Debertine, 2000; Strand et al., 2006) and others (e.g., Edeal & Gildersleeve-Neumann, 2011; Maas & Farinella, 2012). These treatments incorporate principles of motor learning, which refer to the repeated practice of speech movements/gestures in optimal conditions leading to increased accuracy and transfer/generalization (see Maas et al., 2008, for review). Examples of such principles include practice schedule (greater learning with random than with blocked practice; e.g., Knock, Ballard, Robin, & Schmidt, 2000), feedback frequency (greater learning with reduced feedback frequency; Maas, Butalla, & Farinella, 2012), and practice amount (greater learning with more practice trials; Edeal & Gildersleeve-Neumann, 2011). These studies have revealed limited, yet promising, outcomes (Maas, Gildersleeve-Neumann, Jakielski, & Stoeckel 2014; Murray, McCabe, & Ballard, 2014) suggesting that these motor-based treatments and principles of motor learning may be fundamental to successful and effective CAS treatment.

Integral Stimulation (IS), in particular, has been replicated multiple times yielding successful progression towards target productions (Edeal & Gildersleeve-Neurmann, 2011; Maas et al., 2012; Maas et al., 2014; Strand & Debertine, 2000; Strand et al., 2006). IS builds upon motor learning principles through implementation of frequent, intensive treatment sessions that emphasize the production of limited and functional targets with increased gestural accuracy in speech contexts (Maas et al., 2014; Strand et al., 2006; Strand & Skinder, 1999). While the efficacy of these strategies has generated mixed findings (e.g., Edeal & Gildersleeve-Neumann, 2011; Maas & Farinella, 2012; Maas et al., 2012), the underlying ideas of motor learning suggest that IS continues to be an appropriate treatment approach for CAS (Maas et al., 2014).

A core component of IS is the use of visual and auditory models ("Watch me; listen to me; say what I say") with which feedback and cueing hierarchies rely on the accuracy and consistency of speech productions. Although there is support for the efficacy of IS overall, there is currently no evidence to indicate which of the components of IS are the (most) active ingredients. It is possible that some elements/cueing modalities of the treatment are active while other elements are potentially inactive, or perhaps even counterproductive. In order to optimize the efficacy of IS treatment for CAS, it is imperative that we understand the active ingredients and refine the approach to include only those components that enhance learning.

To optimize treatment, intervention should be informed by evidence of speech motor development as well as the underlying nature of CAS. Understanding speech motor development can be accomplished through a review of the Directions Into Velocities of Articulators (DIVA) model (Guenther, 1995; Guenther, Ghosh, & Tourville, 2006). The DIVA model outlines the neurological processes of generating movement output and integrating motor commands with auditory and somatosensory information (motor planning and programming), as well as how this speech motor skill develops.

The model describes two developmental stages: babbling and imitation (Guenther et al., 2006). In the first developmental stage (babbling), children explore speech productions through semi-random articulatory movements. This exploratory period helps to acquire the relationships between motor commands and auditory output through experimentation with motor, somatosensory, and auditory consequences (Terband, Maassen, Guenther, & Brumberg, 2009). Synaptic connections strengthen during this stage, facilitating the development of a feedback system responsible for comparing the intended phonemic output with the resulting auditory output (Terband et al., 2009; Guenther et al., 2006).

Progression into the imitation stage of learning further strengthens motor planning commands, resulting in a feed-forward system that predicts the motor commands required for a specific phonemic output. Children attempt to imitate utterances by predicting speech movements based on auditory cues. Errors in production lead to activation of the feedback system as necessary. The feedback system generates corrective motor commands in real-time based on the relationships acquired during babbling. As motor commands are fine-tuned, somatosensory information accumulates; the individual's speech system begins to master the relationships between motor, auditory, and somatosensory patterns required for intentional speech production of specific utterances. The feedforward system is tuned by learning from, and incorporating, these feedback-based corrections. Over time, the reliance on the corrective feedback system decreases and speech is accomplished through systematic feed-forward predictions resulting in desired auditory targets (Guenther et al., 2006).

Terband et al. (2009) suggest that CAS may be the result of a weakened or imprecise feed-forward system – possibly due to reduced somatosensory sensitivity or impaired auditory connections. Further, their computational simulation study showed that inaccurate feed-forward predictions were correlated with an increased reliance on feedback cues, as well as an increased demonstration of the key symptoms of CAS (Terband et al., 2009). In typical cases of early speech development, the feed-forward system strengthens its accuracy and independence based on feedback responses. A weakened feed-forward system, however, will require increased reliance on feedback cues. In the case of a somatosensory deficit, the feed-forward system may lack the tactile information required for the development of motor plans. Decreased somatosensory accuracy will result in an unreliable production of intended phonemic output, increasing the reliance on the feedback system. Similarly, weakened auditory information will require the system to depend on somatosensory feedback, thus a reliance on the feedback system will develop.

## *The Current Study*

As previously discussed, IS treatment integrates the visual and auditory cues that may facilitate the development of a robust feed-forward system. If this system is weakened in children with CAS – because of impaired somatosensory or auditory feedback processing – then IS would ideally be implemented to circumvent the deficits while simultaneously strengthening the remaining feedback systems. While IS has been successfully replicated across multiple studies, the efficacy of each specific component of IS (e.g., visual-only cues, auditory-only cues, simultaneous auditory and visual cues) has not been evaluated. This limitation in the research has prompted the current study to examine the "active ingredients" that contribute to successful CAS intervention. To date, there is no known research to determine whether some of the active ingredients of IS are more beneficial than others, and/or if any of these components are possibly counterproductive. Based on the current evidence (Strand et al., 2006; Strand et al., 1999; Maas et al., 2014; Maas et al., 2008; Terband et al., 2009), a number of hypotheses can be proposed.

The first hypothesis is formed under the assumption that children with CAS may have a faulty auditory-feedback system. If natural auditory cues are unable to prompt the feed-forward mechanisms to make adjustments during speech trials, then an alternative input channel must be utilized to circumvent the deficit. A visual-only (V-only) sensory cue may allow the clinician to bypass the deficient auditory modality, facilitating the use of a new, productive means of updating the feed-forward system. In addition, contrary to simultaneous auditory-visual cueing, in visual-only cueing the clinician's model does not mask or interfere with the child's self-produced auditory signal. This type of cueing would therefore eliminate improper reinforcement of speech movements by pairing the child's incorrect movement with a correct auditory signal (the clinician's).

Another potential confound in the traditional approaches of IS is that visual cues may be a distraction to the maturation of the feed-forward system. According to the DIVA model (Guenther et al., 2006), feedforward motor commands are developed based on the incorporation of auditory feedback information from previous attempts. If auditory cues are the main source of motor learning, and the goal of speech is to produce sounds, then visual cues may be unhelpful or distracting to development. This suggests that auditory cues alone (A-only) should be sufficient to optimize motor learning, while visual cues may be irrelevant, or overwhelming to the cognitive load of the child.

Lastly, years of clinical expertise has demonstrated successful results through the implementation of simultaneous auditory-visual cues (A+V). Perhaps this success is in fact due to the multisensory nature of the approach – this approach both enhances presumably impaired channels and offers an additional channel to reduce reliance on these impaired channels in order to strengthen the feedforward commands.

In summary, the Visual-Only Hypothesis predicts optimal treatment gains in the Visual-Only cueing condition because this condition circumvents presumed impaired feedback channels and emphasizes a different, and by hypothesis, intact channel to provide feedback for updating the feedforward commands. The Auditory-Only Hypothesis predicts optimal treatment gains in the Auditory-Only cueing condition because producing an auditory signal is the primary goal of speech and this condition aims to remediate the impaired auditory feedback channel through focused stimulation. Finally, the Auditory + Visual Hypothesis predicts optimal treatment gains in the A+V condition because it combines both circumvention and strengthening of impaired channels, and reflects standard clinical practice based on clinical expertise.

# CHAPTER 2

# METHODS

## Participants

The study consisted of two monolingual English-speaking male children ages 4;4 and 4;10. Participants were recruited through the Temple University Speech Language and Hearing Center (TUSLHC), online postings through the Childhood Apraxia of Speech Association of North America (CASANA) website, flyers and advertisements throughout the Philadelphia area, as well as through referrals from licensed speech language pathologists (SLPs). Exclusionary criteria included a native language other than English, a diagnosis of a neurobehavioral disorder that may significantly impact speech, language, and communication abilities (e.g., autism spectrum disorder, Down syndrome), a parent report that suggests uncorrected vision or hearing deficits, or significant impairment of the oral mechanism determined through an oral peripheral examination (e.g., cleft palate, dysarthria).

Participants were evaluated by an experienced SLP recognized by CASANA as an expert in CAS to confirm an appropriate CAS diagnosis. A diagnosis of possible or definite CAS by this SLP was considered an inclusionary criterion. This diagnosis was based on the perceived presence of key CAS characteristics, including inconsistent vowel and/or consonant errors upon repetition of targets, challenges with transitioning between articulatory placements, and abnormal prosody. According to McCauley and Strand (2008), the most reliable form of CAS diagnosis is currently consensus between CAS experts. Therefore, participants' speech samples were also evaluated by two (P001) or three (P003) additional expert SLPs, who made determinations based on video recordings of the assessment administered by the primary SLP. These additional SLP experts used the same diagnostic criteria and awarded each child a score of 0 (no CAS), 1 (possible CAS), or 2 (CAS).

Formal and informal assessments were also included as descriptive measures to further determine the participants' speech and language abilities (see Table 1). The Goldman-Fristoe Test of Articulation, 3rd edition (GFTA-3; Goldman & Fristoe, 2016) was administered to determine articulatory accuracy of high-frequency phonemes; the Diagnostic Evaluation of Articulation and Phonology (DEAP; Dodd, Hua, Crosbie, Holm, & Ozzane, 2006) was administered to examine the presence of phonological processes and misarticulations across words; the Expressive Vocabulary Test, 2nd edition (EVT-2; Williams, 2007) was administered to review language abilities specific to expressive vocabulary and word-finding; the Peabody Picture Vocabulary Test, 4th edition (PPVT-4; Dunn & Dunn, 2007) was administered to examine receptive language abilities; finally, nonverbal portions of the Reynolds Intellectual Assessment Scales (RIAS; Reynolds & Kamphaus, 2003) were administered. Informal measures included the Dynamic Evaluation of Motor Speech Skills (DEMSS; Strand & McCauley, 2013) to determine inconsistencies between target productions and evaluate vowel accuracy, prosodic accuracy, and examine speech accuracy in words with different syllable shapes and word lengths; the Maximum Performance Protocol (MaxPT; Rvachew, Hodge, & Ohberg, 2005) to measure and differentiate symptoms related to dysarthria and apraxia during syllable repetition and prolongation tasks; an oral mechanism examination. A conversational speech sample was also elicited.

*Participant 001*

Participant One (P001) was 4-years, 4-months old at the time of the study. A case history report revealed that P001 lived at home with his two parents and younger sibling (sister; aged 3). English was noted as the primary language in the home. The participant was reported to demonstrate motor milestones within normal limits, including sitting alone at 0;4, standing alone at 0;8, and walking at 0;9. Speech and language milestones are considered delayed: first word at 2;0, and expressing a 2-or-more-word utterance by 2;6. P001's medical history was rather unremarkable. The participant was noted to experience several ear infections during toddlerhood, though was reported to have no history of hearing or vision complications. Similarly, P001 was reported to have no difficulties with swallowing, or concomitant learning disabilities, neurodevelopmental conditions, or neurological diseases. A previous diagnosis of CAS was noted.

In conversational speech, P001 presented with CAS features including abnormal prosody, vowel distortions, and inconsistent phoneme productions (both vowel and consonant phonemes) across repeated trials. Error patterns included velar fronting, prevocalic voicing, and gliding. The participant demonstrated a positive and playful demeanor throughout evaluations, benefitting from frequent verbal and visual cues to manage behavior and aid in self-regulation.

P001 scored in the 3rd percentile on the GFTA-3, presenting with error patterns including prevocalic voicing (e.g., /gʌp/ for /kʌp/), frequent use of ingressive /s/ and /f/, gliding (e.g., /pwet/ for /plet/), initial-/h/ consonant deletion (e.g., /æmə/ for /hæmɚ/), and substitution of /f/ for /θ/ (e.g., /fʌm/ for /θʌm/). Further, P001 earned a *Word Consistency Score* of 76% on the DEAP, with error patterns characterized by ingressive /s/ and /f/, prevocalic voicing, gliding, epenthesis (e.g., /bərɪtʃ/ for /brɪdʒ/), and final consonant devoicing (e.g., /wedibʌk/ for /ledibʌg/). Additionally, P001's performance on the DEMSS exhibited frequent prevocalic voicing and devoicing (e.g., /paɪ/ for /baɪ/ and /mufi/ for /muvi/), ingressive /s/, vowel distortions (e.g., /bubo/ for /bubu/), gliding, velar fronting (e.g., /fɔrdɛt/ for /fɔrgɛt/), denasalization (e.g., /bʌdi/ for /bʌni/), and inconsistent deletion of the phoneme /h/ in word-initial position.

Table 1

*Participant Information & Preliminary Evaluation Results*

| **Participant Information** | **P001** | **P003** |
|---|---|---|
| Age | 4;4 | 4;10 |
| Gender | M | M |
| CAS Score | 1, 1.5, 1 (avg = 1.17) | 2, 0, 0, 0.5 (avg = 0.625) |
| **GFTA-3** | | |
| Standard Score | 71 | 67 |
| Percentile Rank | 3rd | 1st |
| **DEAP** | | |
| Items Produced Differently (D) | 19 | 16 |
| Items Repeated 3 Times (S+D) | 25 | 24 |
| Word Inconsistency Score (D/(S+D)) | 76% | 67% |
| **EVT-2** | | |
| Standard Score | 92 | 80 |
| Percentile Rank | 30th | 9th |
| **PPVT-4** | | |
| Standard Score | 88 | 88 |
| Percentile Rank | 21st | 21st |
| **RIAS** | | |
| Nonverbal Intelligence Index Scaled Score | 89 | 104 |
| Nonverbal Intelligence Index Percentile Rank | 30th | 63rd |
| **DEMSS** | | |
| CV Word Accuracy Total | 34/36 | 32/36 |
| VC Word Accuracy Total | 34/40 | 29/40 |
| Reduplicated Syllables Word Accuracy Total | 11/16 | 12/16 |
| CVC1 Word Accuracy Total | 6/24 | 16/24 |
| CVC2 Word Accuracy Total | 17/40 | 33/40 |
| Bisyllabic1 Word Accuracy Total | N/A | 23/24 |
| Bisyllabic2 Word Accuracy Total | 16/32 | 29/32 |
| Multisyllabic Word Accuracy Total | 9/24 | 21/24 |

| Table 1<br><br>(continued) | | |
|---|---|---|
| **MaxPT** | | |
| Dysarthria Score | 0 | N/A |
| CAS Score | 2 | N/A |
| *N/A = Not Administered* | | |

*Participant 003*

At the time of the study, Participant Three (P003) was 4-years, 10-months old. P003 was reported to live at home with his two parents and three older siblings (sister, 11;0; brothers, 6;0 and 12;0). English was noted to be the primary language spoken in the home, however the participant did attend a daycare in which he was regularly exposed to Spanish. Speech and motor development were not reported for P003. Medical history was notably unremarkable with no history of hearing, vision, or swallowing difficulties. Similarly, no concomitant learning disabilities, neurodevelopmental conditions, neurological diseases, or motor disorders were reported. Typical anatomical structures were also observed. This participant did not have a previous diagnosis of CAS prior to this study.

In conversational speech, P003 exhibited fewer speech errors and CAS features than P001. Key CAS characteristics typically included vowel distortions. Abnormal prosody was intermittently observed. Error patterns included inconsistent initial and/or final consonant deletion, gliding, syllable deletion, prevocalic voicing, fronting, and stopping. P003 presented with a playful demeanor, often requiring visual and verbal cues and a strong behavior management system to aid in encouraging intentional trials throughout the evaluations and treatment sessions.

P003 earned a percentile rank in the 1st percentile on the GFTA-3. Trials were characterized by African American English (AAE) dialectal variations (e.g., unreleased /b/, /g/, and /d/ in final position; unaspirated /p/ in final position), vowel distortion of /æ/, /aɪ/, /ɔ/, and /ə/, velar fronting (e.g., /tʊti/ for /kʊki/), and weak syllable deletion (e.g., /tʃaməz/ for /pədʒaməz/). On the DEAP, P003 earned a *Word Consistency Score* of 67%. Inconsistencies were characterized by gliding (e.g., /wen/ for /ren/), substitution of /f/ for /θ/, velar fronting, final consonant deletion (e.g., /bo/ for /bot/), and deaffrication (e.g., /wat/ for /watʃ/). Further, P003's performance on the DEMSS demonstrated velar fronting, syllable deletion, final consonant deletion, vowel distortions, unreleased /p/ in initial and final positions, substitution of /tʃ/ for /s/, and coalescence (e.g., /badɛt/ for /fɔrgɛt/).

## Design

The study was administered as a single-case, alternating treatment design consisting of three conditions (V-only, A-only, A+V). The treatment study was designed to last 26 weeks, consisting of an assessment phase, baseline phase, two treatment phases, two maintenance phases, and two follow-up phases (see Table 2). All three conditions were treated in each session throughout the treatment phases. Untreated words or phrases (probes) were implemented throughout baseline, treatment, and maintenance phases as experimental controls. Follow-up phases were planned to be four weeks apart, with Follow-Up 2 being eight weeks post the final phase of maintenance ($Maint_2$).

To account for systematic order effects, target sets (with their assigned condition – see Target Selection; e.g., A-only, V-only, and A+V) were randomized at the beginning of each session by the participant. The children were given three dice – each of which were assigned to a condition – and the participants' roll determined the order in which conditions were administered. Additionally, to control for number of trials, each therapy condition was timed while administering treatment. This served as a management tool to prevent targets within each condition from being targeted in significantly more trials than others. Throughout both phases of treatment ($Tx_1$ and $Tx_2$), target sets remained the same in each condition (see Target Selection).

Table 2

*Treatment Timeline*

| Phase | Assess. | | | | Baseline | | | | $Tx_1$ | | | | Maint.$_1$ | | $Tx_2$ | | | | Maint.$_2$ | | Follow-Up$_1$ | Follow-Up$_2$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Week | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 22 | 26 |
| Probe | | | | | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | | X |

*Target Selection*

Each treatment condition was assigned a target set that was treated during each session across both phases. A total of five targets sets were curated: V-only target set, A-only target set, A+V target set, and two Control target sets. Targets were generated based on their functional relevance to the participants in order to ensure motivation of trials and functional use outside of treatment. Parent report indicated words and names that were of interest to the participants (see Appendix A).

Target sets were examined and balanced to address three criteria. First, sets were balanced for the number of "visible" phonemes (e.g., bilabials, labiodentals, interdentals, rounded vowels and consonants, and diphthongs) to ensure that no target set contained significantly higher numbers of phonemes that could provide visual cues during treatment – specifically during the V-only condition (see Table 3 and 4). Additionally, target sets were balanced to control for the number of syllables within each set. This increased the likelihood that words of comparable difficulty were targeted in each set. Target sets were also controlled for functional relevance and interest to maintain similar levels of motivation across treatment conditions. Balanced target sets were then randomly assigned to treatment conditions.

Table 3

*Participant 001 Balanced Target Sets*

| | # syllables (avg.; range) | # consonants (avg.; range) | # of "visible" consonants* (avg.; range) | # of "visible" vowels** (avg.; range) |
|---|---|---|---|---|
| A+V | 3; 2-4 | 4.2; 2-6 | 1.2; 0-3 | 1; 0-2 |
| A-only | 3; 2-4 | 4.4; 2-7 | 1.4; 0-3 | 1; 0-2 |
| V-only | 3; 2-5 | 5.8; 4-8 | 1.4; 1-2 | 1.2; 0-2 |
| Control Set 1 | 2.8; 2-4 | 5.4; 3-9 | 1.2; 0-3 | 1.6; 1-2 |
| Control Set 2 | 3; 2-5 | 4.6; 3-7 | 1; 0-2 | 0.8;0-1 |

** Includes bilabial, labiodental, and interdental consonants and consonants with lip-rounding (e.g., /ʃ/, /tʃ/, /ʒ/, /dʒ/, /ɹ/)*
*** Includes rounded vowels and diphthongs*

Table 4

*Participant 003 Balanced Target Sets*

| | # syllables (avg.; range) | # consonants (avg.; range) | # of "visible" consonants* (avg.; range) | # of "visible" vowels** (avg.; range) |
|---|---|---|---|---|
| A+V | 2.8; 2-4 | 4.4; 1-6 | 1.8; 0-4 | 0.6; 0-2 |
| A-only | 2.8; 2-4 | 5; 3-7 | 2.2; 1-3 | 1; 0-2 |
| V-only | 2.8; 2-4 | 5; 4-6 | 1.6; 0-3 | 0.8; 0-1 |
| Control Set 1 | 3.4; 2-7 | 6.6; 4-9 | 2; 1-3 | 1; 0-3 |
| Control Set 2 | 3; 1-5 | 5.4; 3-8 | 2.2; 1-3 | 0.8; 0-1 |

** Includes bilabial, labiodental consonants, and consonants with lip-rounding (e.g., /ʃ/, /tʃ/, /ʒ/, /dʒ/, /ɹ/)*
*** Includes rounded vowels and diphthongs*

## Procedures

### *Treatment Procedures*

Participants attended approximately two treatment sessions per week, with schedule modifications, as needed. Treatment sessions were administered as one-hour sessions. Each treatment condition was treated for 15-minutes blocks with five-minute breaks between conditions. Treatment conditions were randomized by the participant at the start of the session by using three dice. Each die was assigned a condition by the participant; the participant matched a colored die to a list of targets corresponding to a condition. The number displayed after a roll would determine the order in which conditions were administered (i.e., lowest number rolled was administered first). When ties occurred, the tied dice were rolled until the numbers differed. Targets were treated randomly each session by shuffling the deck of index cards to avoid systematic order effects.

Targets in each treatment condition were elicited using the hallmark prompts of IS treatment (e.g., "Watch me," "listen to me," and "say what I say"). In the V-Only condition, the clinician prompted, "Watch me carefully, we are going to say *target* three times together. Ready, let's say *target*. Go!" The SLP then mouthed/mimed the articulatory movements of the target without voicing to solely provide visual cues. In the A-only condition, the SLP prompted, "Listen to me carefully, we are going to say *target* three times together. Ready, let's say *target*. Go!" The clinician then produced the target while covering their mouth with their hand or a clipboard, only providing auditory cues. Finally, in the A+V condition, the SLP prompted, "Watch me and listen to me carefully, we are going to say *target* three times together. Ready, let's say *target*. Go!" The clinician then produced the target without covering their mouth to provide auditory and visual cues.

A treatment episode was defined as a sequence of events in which the SLP elicited a target and provided support for the production if needed. Treatment episodes for each target began with elicitation of the target either through Integral Stimulation (with an immediate or delayed production attempt by the child) or in response to a question (for independent production attempts). If the participant's trial was correct, the clinician elicited a second attempt – either via Integral Stimulation or in response to a question – to reinforce motor learning and accuracy. If the participant's trial was incorrect, the SLP would provide feedback as appropriate (feedback was faded across treatment episodes within each condition) and direct the child to attempt three trials of the target in slowed simultaneous production. During these three simultaneous attempts, the conditions differed: simultaneous production was elicited in correspondence to the target's condition (e.g., the clinician and participant would produce a target in the V-Only condition given only visual cues). Following simultaneous production, the SLP elicited a final attempt of the target using the same elicitation method as the beginning of the treatment episode. See Appendix B for an outline of the treatment protocol.

Following both productions (i.e., first and final attempt), the SLP provided delayed verbal feedback on a fading feedback schedule (five out of the first five treatment episodes, four out of the next five treatment episodes, etc.). Target-specific feedback (e.g., "I heard the air blow out when you said /s/. Good job!") and knowledge of results (e.g., "Great work!") was faded throughout treatment episodes (Ballard et al., 2007) though was not systematically controlled. For example, during the first treatment episode in a condition, the clinician would provide feedback for all five trials of the target. In the treatment episode for the final target of the condition, the clinician would provide feedback for only one trial of the target.

### *Probe Procedures*

Participants' results were measured by determining the accuracy of targets without the provision of feedback from clinicians during probes. Probes were administered once per week during baseline, treatment, and maintenance phases, as well as during the final follow-up phase. During treatment phases, probes were administered at the beginning of the session to capture generalization from previous sessions and prevent temporary effects.

Probe trials consisted of repetition of targets from all target sets (e.g., A-Only, V-Only, A+V, Control 1, and Control 2) in direct imitation. Targets were elicited in random order to prevent the participant from anticipating the next target. The clinician produced a target and then prompted the participant to repeat the item. Participants attempted each item in direct imitation one time (25 probes total).

## Data Analysis

The dependent variable was perceptual accuracy of probes. All data were analyzed by two trained analysts (undergraduate students in Communication Sciences and Disorders) who were not involved in administering treatment or probes, and who were blinded to treatment condition and time point of the probe (e.g., baseline, follow-up, etc.). The analysts scored each production of an item (treated target or untreated control item) on a 3-point rating scale, where 2 = correct, 1 = minor error, and 0 = major error or multiple errors. Major and minor errors were operationally defined (see Appendix C) and included judgment of both segmental and suprasegmental aspects. This system has been used previously to document treatment efficacy in CAS and has been shown to be reliable and sensitive to change (e.g., Maas et al., 2012; Strand & Debertine, 2000; Strand et al., 2006).

Reliability between analysts on a point-to-point basis across all probes was 64% for P001 and 54% for P003. The vast majority of disagreements were within a single point (e.g., Rater1 score = 0, Rater2 score = 1). Reliability within one scale point was 97% for P001 and 95% for P003. To ensure adequate interrater reliability and maintain sensitivity, scores of the two raters were combined into a single score as follows: if raters' scores were equal, the score was deemed reliable and remained the same (e.g., if both raters scored a 0, the probe received a 0; if both raters scored a 2, the probe received a 2; etc.); if scores differed, the probe received a 1 (e.g., if Rater1 scored 0 and Rater2 scored 1, probe received a 1; if Rater1 scored 2 and Rater2 scored 0, probed received a 1; etc). In other words, when the raters agreed on a score, that score was maintained, and when the raters disagreed, the score was 1, indicating neither clear-cut correct or incorrect responses.

Data were averaged by condition for each timepoint and plotted for visual analysis. In addition, standardized effect sizes for each condition were computed according to the formula of Busk and Serlin (1992): effect size d = ($mean_{post} - mean_{pre}$) / standard $deviation_{pre}$. Following Maas and Farinella (2012), a significant effect was defined as $d > 1$ (see also Behrman, 2014; McAllister et al., 2016), meaning that the change in mean accuracy exceeds the baseline standard deviation.

According to the Auditory + Visual Hypothesis, it was predicted that a greater effect size would be observed in the A+V condition than in the V-Only and A-Only conditions. Similarly, it was predicted that the V-Only condition would yield a greater effect size than the A+V and A-Only conditions, according to the Visual-Only Hypothesis. Further, the Auditory-Only Hypothesis predicts a larger effect size in the A-Only condition than the A+V and V-Only conditions. Finally, it was expected that all treated probes, regardless of condition (i.e., A-Only, V-Only, and A+V), would yield a larger effect size than untreated items (i.e., $Control_1$ and $Control_2$).

# CHAPTER 3

# RESULTS

Table 5 and Table 6 display P001's and P003's accuracy of probes across Baseline, Maintenance$_2$, and Follow-Up phases, respectively. Accuracy of probes across all treatment phases are reflected in Figure 1 (P001) and Figure 2 (P003).

Participant 001

Table 5

*Participant 001 Percent Accuracy of Probes*

| Condition | Baseline Average (SD) | Maint.$_2$ Average | Follow-Up Average | d Maint.$_2$ | d Follow-Up | d All |
|---|---|---|---|---|---|---|
| **A-Only** | **36.7% (5.8%)** | **50.0%** | **45.0%** | **2.31** | **1.44** | **1.88** |
| **V-Only** | **26.7% (5.8%)** | **50.0%** | **40.0%** | **4.04** | **2.31** | **3.18** |
| **A+V** | **26.7% (5.8%)** | **35.0%** | **30.0%** | **1.44** | **0.58** | **1.01** |
| Control 1 | 13.3% (15.3%) | 45.0% | 35.0% | 2.07 | 1.42 | 1.75 |
| Control 2 | 40.0% (10.0%) | 45.0% | 50.0% | 0.50 | 1.00 | 0.75 |
| **Control** | **26.7% (12.6%)** | **45.0%** | **42.5%** | **1.46** | **1.26** | **1.36** |

P001 demonstrated relatively stable baseline accuracy for probes in the A-Only and A+V conditions. V-Only probes displayed a sharp decline in accuracy by the third baseline session, with Control probes demonstrating variability throughout all baseline sessions. All conditions except the A-only condition showed some evidence of improvement toward the end of the first treatment phase, most immediately for the V-only condition. After the first treatment phase, P001 showed above-baseline performance for all conditions except the AV-condition. Treatment phase 2 started with a drop in performance for all conditions, with increasing performance toward the end of the

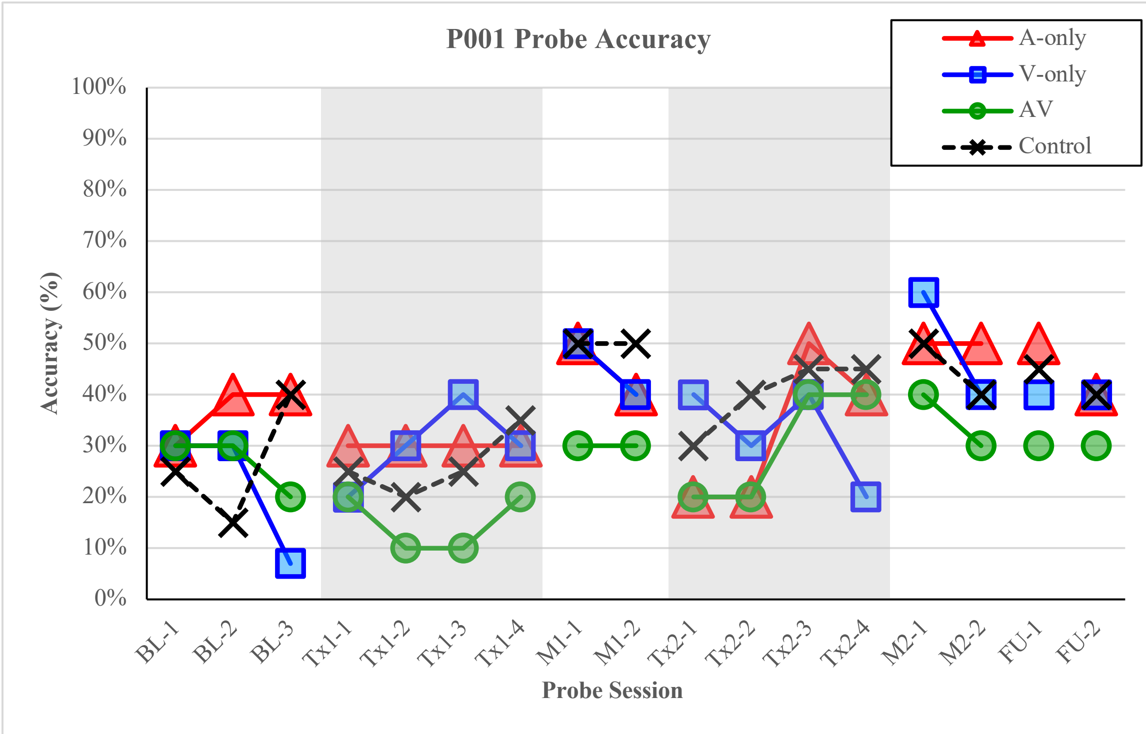



Figure 1. Participant 001 Probe Accuracy.

treatment phase for all conditions except V-only, which showed a drop at the last treatment probe. Following treatment phase two, P001 displayed an increase in probe accuracy across the V-Only, A-Only, and Control conditions. After a slight drop in performance at the second maintenance probe, accuracy mostly stabilized at above-baseline levels for all conditions except the AV condition.

Effect sizes following the second Maintenance phase indicate effects (d > 1) for all conditions, including the combined control sets (d = 1.46; 18.3% gain). However, only the V-only condition showed an effect size notably greater than the Control condition (d = 4.04; 23.3% gain). The A-only condition effect size (d = 2.31; 13.3% gain) was greater than the combined control sets, although the difference with one of the control sets (d = 2.07; 31.7% gain) was negligible. This pattern was maintained at follow-up, where the V-only condition showed the largest effect size (d = 2.31; 13.3% gain), and all other conditions except the AV condition showed effects > 1.

In sum, P001 showed improvements in all conditions following treatment phase 2. The V-only condition was the only one that consistently exceeded the effect size in the Control condition at both Maintenance 2 and at follow-up.

Participant 003

Table 6

*Participant 003 Percent Accuracy of Probes*

| Condition | Baseline Average (SD) | Maint.2 Average | Follow-Up Average | d Maint.2 | d Follow-Up | d All |
|---|---|---|---|---|---|---|
| **A-Only** | **30.0% (8.2%)** | **10.0%** | **30.0%** | **-2.45** | **0.00** | **-1.63** |
| **V-Only** | **27.5% (5.0%)** | **25.0%** | **30.0%** | **-0.50** | **0.50** | **-0.17** |
| **A+V** | **40.0% (8.2%)** | **45.0%** | **40.0%** | **0.61** | **0.00** | **0.41** |
| Control 1 | 32.5% (9.6%) | 20.0% | 30.0% | -1.31 | -0.26 | -0.96 |
| Control 2 | 31.9% (9.0%) | 45.0% | 50.0% | 1.46 | 2.02 | 1.65 |
| **Control** | **32.2% (8.4%)** | **32.5%** | **40.0%** | **0.03** | **0.93** | **0.33** |

P003 demonstrated some variability during baseline probes across all conditions – most notably in the V-Only and A+V conditions – but no obvious improving trends. Following the first treatment phase, P003 displayed a level of accuracy comparable to baseline, with a slight increase in accuracy only in the A+V condition. Notable variability of accuracy is evident across sessions in all conditions, however. Treatment phase 2 began with stable accuracy for all conditions, with a sharp increase in accuracy in all conditions – most significantly in the A+V, A-Only, and Control conditions. By the end of treatment phase 2, all conditions maintained probe accuracy above baseline with the exception of the V-Only condition. Averaged across both maintenance probes (M2-1 and

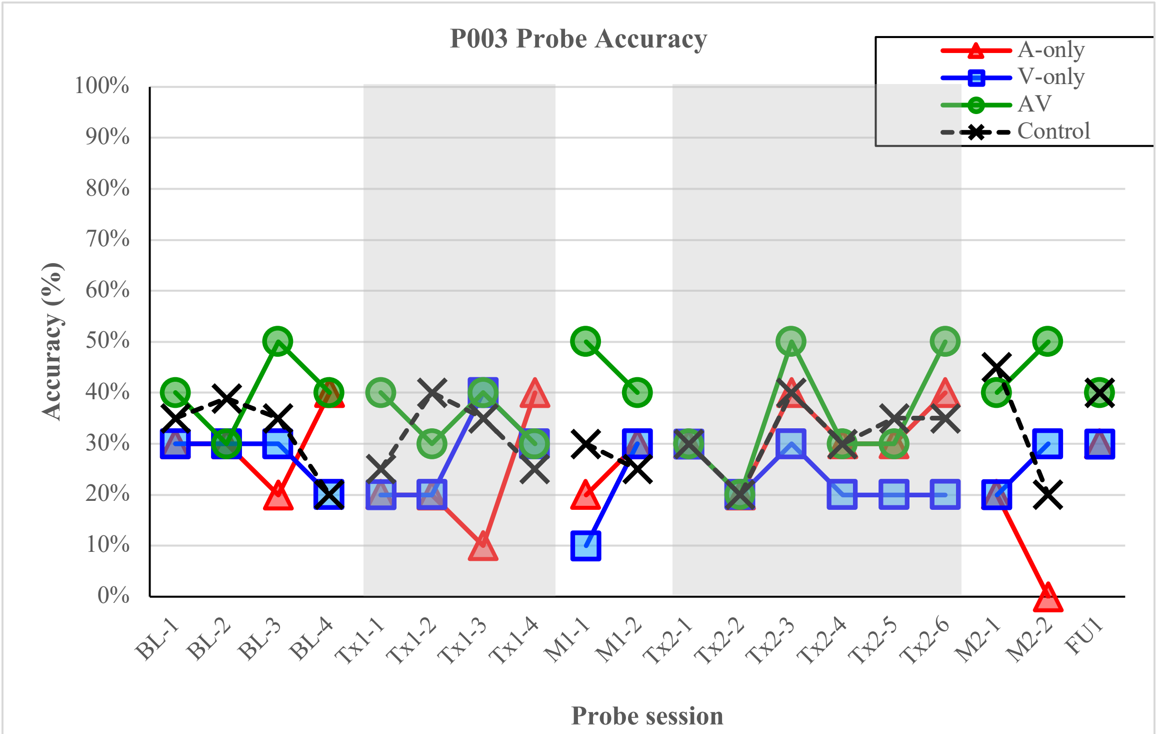



Figure 2. Participant 003 Probe Accuracy.

M2-2), only the AV condition showed a modest increase in accuracy, with little to no change for the control set or the V-only set, and a decrease for the A-only set. Once again, significant variability in probe accuracy was noted during the second maintenance phase, with accuracy returning back to baseline accuracy by follow-up for all conditions except the control set.

Effect sizes following maintenance phase 2 corroborated these visual patterns: no significant improvements for any condition, and a decrease in accuracy for the A-only condition (d = -2.45; -20.0% decline). Overall, the control condition also showed no effect, although the Control 2 set displayed the most gain (d=1.46; 13.1% gain). At follow-up, none of the conditions showed an effect (all d < 1), again with the exception of Control 2 which showed a significant increase (d=2.02; 18.1% gain).

Overall, P003 did not display progress in any condition, and in fact showed a significant regression from baseline in the A-Only condition (d=-1.63; 13.3% decrease). The second Control condition demonstrated significant gain (d=1.65; 14.8% increase), though when averaged with Control 1, Control effect size was insubstantial.

# CHAPTER 4

# DISCUSSION

The purpose of this study was to evaluate the efficacy of the different cueing modalities used in IS therapy for treatment of CAS. This study considered the possibility that some cues may be more beneficial, or counterproductive, during treatment. In addition, the study aimed to evaluate the efficacy of IS treatment by comparing treated and untreated utterances. Given that the results did not display a uniform pattern between participants, each participant will be discussed individually within each section below, to highlight the effects specific to each child.

## Modality Effects

Findings with respect to modality effects differed between children, in that only one child showed a clear effect. P001 demonstrated a clear advantage of the V-Only condition compared to other conditions. At follow-up, P001 displayed an effect size of d = 2.31 in the V-Only condition. The effect was even stronger immediately following treatment. The Visual-Only Hypothesis predicted that the V-Only condition would yield the greatest effect size due to its potential to circumvent an insufficient auditory feedback system. The visually-specific effects displayed in P001's performance suggest that, at least for some children with CAS, visual-only cues may be most beneficial to maximize the efficacy of IS treatment. The modality with the smallest effect size was the A+V condition, suggesting that combined cues may pose a cognitive challenge (in some children) as this condition requires simultaneous processing of two different input modalities. This notion was further supported by the finding that the A-Only condition also resulted in a robust change immediately following treatment, although this effect dissipated somewhat at follow-up. Thus, single-modality cues appeared to be most effective for P001, but especially the V-Only condition, which showed the strongest retention.

In contrast, P003 did not demonstrate any condition effects immediately following treatment nor at follow-up. His performance varied with no clear connection to treatment phases, and no effects were evident either visually nor in effect size measures. The reasons for the lack of effects is unclear, but it should be noted that no treatment effects (treated vs. untreated items) were evident in his data either (see also below).

Taken together, the present findings suggest that some children with CAS may benefit from single-modality simultaneous production cues, in particular visual-only cues, although not all children may show such effects. These findings are important as they suggest that the common practice of using simultaneous auditory and visual productions are less effective than single-modality, and especially visual-only, cues. Clearly, further research is needed to replicate these effects and determine any participant-specific factors that predict which modality may be optimal for a given child. The present results at least suggest that such modality effects may be present and that certain cue modalities produce greater learning for some children.

## Integral Stimulation Effects

With respect to the question of overall treatment efficacy, results again were mixed, with P001 showing a clear treatment effect and P003 showing no effect. For

P001, immediately following treatment, the A-only and the V-only conditions showed effect sizes above and beyond the control condition, although the difference for the A-only condition was relatively small when compared to one of the control sets. The advantage of the V-only condition over the untreated control condition was maintained at follow-up, supporting the efficacy of IS treatment.

It is interesting to note that the A+V condition, which is perhaps the more typically used cueing condition, did not produce gains that were reliably different from the untreated control condition. This contrasts with previous studies that supported the efficacy of IS treatment with A+V simultaneous production (e.g., Maas et al., 2019; Strand et al., 2006). The lack of a treatment effect for the A+V condition in our study may reflect an effect of specific target items in this set, in that perhaps the items randomly assigned to this set were more challenging or less motivating, despite our best efforts at matching sets. Alternatively, it is possible that this combined cue is less effective than single-modality cues, and that previous studies have found effects in spite of use of this cue. For example, an effect may have emerged with more treatment, whereas in the present study each set of items received treatment only in one-third of the session, since there were three conditions to be compared. It is also possible that previous studies, when incorporating A+V simultaneous production, quickly faded the auditory model, effectively making these previous studies more in line with our V-only condition. In our study protocol, fading of auditory cues was not allowed in order to be able to compare the conditions.

It is unclear why no effects emerged for P003. One possibility is that the treatment was insufficiently intense. Recent work has shown that more practice is better than less practice (e.g., Edeal & Gildersleeve-Neumann, 2011; Maas et al., 2019) and that massed practice is more effective than distributed practice for most children with CAS (Maas et al., 2019; Thomas et al., 2014). The present study included three different target sets, all of which were practiced in each session. Therefore, each target set only received 15 minutes of practice per session, compared to about 25 minutes per session in Maas et al. (2019). Perhaps treatment effects may have emerged if treatment had continued and/or had been offered in a more massed procedure.

Another possibility that may have contributed to limited treatment effects is that P003 was not sufficiently motivated. Anecdotal observations support this interpretation, as P003 often required significant redirection and encouragement (both by clinician and parents) to engage in practice. Motivation is an important prerequisite for learning and for engaging in structured learning tasks, such as those in IS treatment. Previous research with IS treatment has also reported limited treatment gains in some children with CAS likely due to lack of motivation and/or compliance (e.g., Maas et al., 2012; Strand et al., 2006).

A final possibility for limited treatment effects in P003 is his diagnosis. While P003 demonstrated symptoms of CAS, including vowel distortions, intermittent abnormal prosody, and atypical speech error patterns, the expert clinicians considered P003's presentation and diagnosis to be less definitive than P001's. It is possible that P003 did not have CAS, had a milder form of CAS, or even a different subtype of CAS, thus resulting in different treatment effects.

While P003 did not demonstrate clear effects, it is also notable that the untreated control items demonstrated some improvement throughout the study. Perhaps P003 developed an aversion to treated targets and was therefore more motivated by untreated control items during probes. Additionally, it is also possible that the control targets were more intrinsically motivating, or easier/less complex, than treated items despite our best efforts at balancing the targets.

Overall, our findings are consistent with prior literature on IS treatment in that a treatment effect was observed for some but not all children. Clearly, further research is needed to replicate and extend these findings.

## Future Directions

Based on the results of this study, it is possible that different children with CAS may respond differently to particular cue modalities. Participant 001 benefitted most from the V-Only condition. Future studies should replicate and evaluate the effects of different cue modalities in larger and more diverse groups of children with CAS. Perhaps some children benefit most from single-modality cues (e.g., V-Only, A-Only) and others profit from multi-modality cues (e.g., A+V). Similarly, future studies should consider identifying the learning style of the participants (e.g., visual learner, auditory learner) as this may predict which cues are most optimal. Future research should also investigate the efficacy of additional cue modalities, including gestural cues, tactile cues, rhythmic cues, etc., as these types of cues are often utilized in clinical practice. Lastly, it is possible that fading the auditory modality during the simultaneous condition may also provide additional benefits, which future studies should also consider.

## Conclusions

The results of the present single-case experimental design study with two children with CAS show that single-modality simultaneous production cues, and in particular visual-only cues, may be optimal in IS treatment for CAS. One child showed a clear advantage of visual cues over auditory-only and auditory + visual cues, consistent with the hypothesis that visual information helps bypass a potentially impaired auditory feedback processing channel. This child also demonstrated a clear treatment effect following treatment, with larger effect sizes for treated items (in particular visual-only items, but also auditory-only items) than for untreated items, supporting the efficacy of the overall approach. However, these effects of cue modality or treatment were not evident for the other participant with CAS, for unknown reasons. Nevertheless, these findings are consistent with the literature on IS treatment for CAS, in that the overall approach was shown to produce treatment effects for at least some children with CAS.

## REFERENCES CITED

American Speech-Language-Hearing Association. (2007). *Childhood apraxia of speech* [Technical Report]. Available from www.asha.org/policy.

American Speech-Language-Hearing Association. (2016). 2016 Schools Survey report: SLP caseload characteristics. Available at www.asha.org/research/memberdata/schoolssurvey/

Austermann Hula, S. N., 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.

Ballard, K. J., Robin, D. A., McCabe, P., & McDonald, J. (2010). A treatment for dysprosody in childhood apraxia of speech. *Journal of Speech, Language, and Hearing Research, 53,* 1227-1245.

Behrman, A. (2014). Segmental and prosodic approaches to accent management. *American Journal of Speech-Language Pathology, 23,* 546-561.

Busk, P. L., & Serlin, R. C. (1992). Meta-analysis for single case research. In T. R. Kratochwill & J. R. Levin (Eds.), *Single-case research design and analysis: New directions for psychology and education* (pp. 187-212). Hillsdale, NJ: Erlbaum.

Campbell, T. F. (1999). Functional treatment outcomes in young children with motor speech disorders. In A. J. Caruso & E. A. Strand (Eds.), *Clinical Management of Motor Speech Disorders in Children* (pp. 385-396). New York: Thieme.

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.

Glogowska, M., Roulstone, S., Peters, T. J., & Enderby, P. (2006). Early speech- and language-impaired children: Linguistic, literacy, and social outcomes. *Developmental Medicine & Child Neurology, 48,* 489-494.

Goldman, R., & Fristoe, M. (2000). *The Goldman-Fristoe Test of Articulation (2nd Edition)*. Austin, TX: Pro-Ed.

Guenther, F. H. (1995). Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. *Psychological Review, 102*(3), 594-621.

Guenther, F. H., Ghosh, S. S., & Tourville, J. A. (2006). Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain and Language, 96*(3), 280-301.

Lewis, B. A., Freebairn, L. A., Hansen, A. J., Iyengar, S. K., & Taylor, H. G. (2004). School0age follow-up of children with childhood apraxia of speech. *Language, Speech, and Hearing Services in Schools, 35*(2), 122-140.

Knock, T. R., Ballard, K. J., Robin, D. A., Schmidt, R. A. (2000). Influence of order of stimulus presentation on speech motor learning: A principled approach to treatment for apraxia of speech. *Aphasiology, 14(5-6),* 653-668.

Maas, E., & Farinella, K. A. (2012). Random vs. blocked practice in childhood apraxia of speech. *Journal of Speech, Language, and Hearing Research, 55*(2), 561-578.

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., Gildersleeve-Neumann, C. E., Jakielski, K. J., & Stoeckel, R. (2014). Motor-based intervention protocols in treatment of childhood apraxia of speech (CAS). *Current Developmental Disorders Reports, 1,* 197-206.

Maas, E., Gildersleeve-Neumann, C. E., Jakielski, K. J., Kovacs, N., Stoeckel, R., Vradelis, H., & Welsh, M. (2019). Bang for your buck: A single-case experimental design study of practice amount and distribution in treatment for childhood apraxia of speech. *Journal of Speech, Language, and Hearing Research, 62(9),* 3160-3182.

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, and Hearing Research, 17*, 277-298.

McAllister Byun, T., Swartz, M. T., Halpin, P. F., Szeredi, D., Maas, E. (2016). Direction of attentional focus in biofeedback for treatment for /r/ misarticulation. *International Journal of Language & Communication Disorders, 51(4),* 384-401.

McCauley, R. J., & Strand, E. A. (2008). A review of standardized tests of nonverbal oral and speech motor performance in children. *American Journal of Speech-Language Pathology, 17,* 18-91.

McCormack, J. McLeod, S., McAllister, L., & Harrison, L. J. (2009). A systematic review of the association between childhood speech impairment and participation across the lifespan. *International Journal of Speech-Language Pathology, 11,* 155-170.

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*(3), 486-504.

Overby, M., Carrell, T., & Bernthal, J. (2007). Teachers' perception of students with speech sound disorders: A quantitative and qualitative analysis. *Language, Speech, and Hearing Services in Schools*, *38*(4), 327-341.

Raitano, N.A., Pennington, B. F., Tunick, R.A., Boada, R., and Shriberg, L. D. (2004). Pre-literacy skills of subgroups of children with speech sound disorders. *Journal of Child Psychology and Psychiatry*, *45*(4), 821-835.

Reynolds, C. R., & Kamphaus, R. W. (2003). *Reynolds Intellectual Assessment Scales (RIAS)*. Lutz, FL: Psychological Assessment Resources, Inc.

Rusiewicz, H. L., Maize, K., & Ptakowski, T. (2018). Parental experiences and perceptions related to childhood apraxia of speech: Focus on functional implications. *International Journal of Speech-Language Pathology, 20*(5), 569-580.

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.

Shriberg, L. D., Aram, D., & Kwiatkowski, J. (1997). Developmental apraxia of speech: Descriptive and theoretical perspectives. *Journal of Speech, Language, and Hearing Research, 40,* 273-285.

Skinder, A., Strand, E. A., & Mignerey, M. (1999). Perceptual and acoustic analysis of lexical and sentential stress in children with developmental apraxia of speech. *Journal of Medical Speech-Language Pathology, 7*(2), 133-144.

Strand, E. A., & Skinder, A. (1999). Treatment of developmental apraxia of speech: Integral stimulation methods. In A. J. Caruso & E. A. Strand (Eds.), *Clinical Management of Motor Speech Disorders in Children* (pp. 109-148). New York, NY: Thieme.

Strand, E. A., & Debertine, P. (2000). The efficacy of integral stimulation intervention with developmental apraxia of speech. *Journal of Medical Speech-Language Pathology, 8,* 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*(4), 297-307.

Terband, H., Maassen, B., Guenther, F. H., & Brumberg, J. (2009). Computational neural modeling of speech motor control in childhood apraxia of speech (CAS). *Journal of Speech, Language, and Hearing Research, 52*(6), 1595-1609.

Thomas, D. C., McCabe, P., & Ballard, K. G. (2014). Rapid Syllable Transitions (ReST) treatment for childhood apraxia of speech: The effect of lower dose-frequency. *Journal of Communication Disorders, 51,* 29-42.

Williams, K. T. (2007). *Expressive Vocabulary Test (2nd Edition)*. Pearson.

## APPENDIX A

## PARTICIPANT 001 TARGET SETS

### *Set 1 – Audio-Visual (A+V) Condition*

*Consonant Data*

| | Bilabials | Labio-dentals | Inter-dentals | Rounded | Visible | Alveolars | Other | Nonvisible | Total |
|---|---|---|---|---|---|---|---|---|---|
| **E (sister)** | 1 | 0 | 0 | 0 | 1 | 1 | 0 | **1** | **2** |
| **Juice please** | 1 | 0 | 0 | 1 | 2 | 3 | 0 | **3** | **5** |
| **Let's play Zingo!** | 1 | 0 | 0 | 0 | 1 | 5 | 1 | **6** | **7** |
| **Noni's house** | 0 | 0 | 0 | 0 | 0 | 4 | 1 | **5** | **5** |
| **Mr. J** | 1 | 0 | 0 | 2 | 3 | 1 | 0 | **1** | **4** |
| ***Total*** | 4 | 0 | 0 | 3 | **7** | 14 | 2 | **16** | **23** |
| ***Mean*** | | | | | **1.4** | | | **3.2** | **4.6** |

*Vowel Data*

| | Rounded | Diphthong | Visible | Unrounded* | Nonvisible | Total Vowels | Total Syllables |
|---|---|---|---|---|---|---|---|
| **E (sister)** | 0 | 0 | 0 | 3 | 3 | 3 | **3** |
| **Juice please** | 1 | 0 | 1 | 1 | 1 | 2 | **2** |
| **Let's play Zingo!** | 1 | 0 | 1 | 3 | 3 | 4 | **4** |
| **Noni's house** | 1 | 1 | 2 | 1 | 1 | 3 | **3** |
| **Mr. J** | 1 | 0 | 1 | 2 | 2 | 3 | **3** |
| ***Total*** | **4** | **1** | **5** | **10** | **10** | **15** | **15** |
| ***Mean*** | | | **1** | | **2** | **3** | |
| *Unrounded includes spread and neutral vowels | | | | | | | |

*Set 2 – Control$_1$ Condition*

*Consonant Data*

| | Bilabials | Labio-dentals | Inter-dentals | Rounded | Visible | Alveolars | Other | Nonvisible | Total |
|---|---|---|---|---|---|---|---|---|---|
| **See you next time!** | 1 | 0 | 0 | 0 | 1 | 5 | 2 | **7** | **8** |
| **Playground** | 1 | 0 | 0 | 1 | 2 | 3 | 1 | **4** | **6** |
| **My sneakers** | 1 | 0 | 0 | 0 | 1 | 3 | 1 | **4** | **5** |
| **Cheerios** | 0 | 0 | 0 | 2 | 2 | 1 | 1 | **2** | **4** |
| **Mrs. XX (L)** | 1 | 0 | 0 | 1 | 2 | 4 | 0 | **4** | **6** |
| ***Total*** | 4 | 0 | 0 | 5 | **8** | 16 | 5 | **21** | **29** |
| ***Mean*** | | | | | **1.6** | | | **4.2** | **5.8** |

*Vowel Data*

| | Rounded | Diphthong | Visible | Unrounded* | Nonvisible | Total Vowels | Total Syllables |
|---|---|---|---|---|---|---|---|
| **See you next time!** | 1 | 1 | 2 | 2 | 2 | 4 | **4** |
| **Playground** | 0 | 1 | 1 | 1 | 1 | 2 | **2** |
| **My sneakers** | 1 | 1 | 2 | 1 | 1 | 3 | **3** |
| **Cheerios** | 1 | 0 | 1 | 2 | 2 | 3 | **3** |
| **Mrs. XX (L)** | 1 | 0 | 1 | 2 | 2 | 3 | **4** |
| ***Total*** | **4** | **3** | **7** | **8** | **8** | **15** | **16** |
| ***Mean*** | | | **1.4** | | **1.6** | **3** | |
| *Unrounded includes spread and neutral vowels | | | | | | | |

*Set 3 – Audio-Only (A-Only) Condition*

*Consonant Data*

| | Bilabials | Labio-dentals | Inter-dentals | Rounded | Visible | Alveolars | Other | Nonvisible | Total |
|---|---|---|---|---|---|---|---|---|---|
| **I'm so mad!** | 2 | 0 | 0 | 0 | 2 | 2 | 0 | **2** | **4** |
| **Tacos** | 0 | 0 | 0 | 0 | 0 | 2 | 1 | **3** | **3** |
| **What's your name?** | 2 | 0 | 0 | 0 | 2 | 3 | 1 | **4** | **6** |
| **XXXX (A)** | 0 | 0 | 0 | 0 | 0 | 1 | 1 | **2** | **2** |
| **That's silly** | 0 | 0 | 1 | 0 | 1 | 4 | 0 | **4** | **5** |
| ***Total*** | 4 | 0 | 1 | 0 | **5** | 12 | 3 | **15** | **20** |
| ***Mean*** | | | | | **1** | | | **3** | **4** |

*Vowel Data*

| | Rounded | Diphthong | Visible | Unrounded* | Nonvisible | Total Vowels | Total Syllables |
|---|---|---|---|---|---|---|---|
| **I'm so mad!** | 1 | 1 | 2 | 1 | 1 | 3 | **3** |
| **Tacos** | 1 | 0 | 1 | 1 | 1 | 2 | **2** |
| **What's your name?** | 2 | 0 | 2 | 2 | 2 | 4 | **3** |
| **XXXX (A)** | 0 | 0 | 0 | 4 | 4 | 4 | **4** |
| **That's silly** | 0 | 0 | 0 | 3 | 3 | 3 | **3** |
| ***Total*** | **4** | **1** | **5** | **11** | **11** | **16** | **15** |
| ***Mean*** | | | **1** | | **2.2** | **3.2** | |
| *Unrounded includes spread and neutral vowels | | | | | | | |

*Set 4 – Visual-Only (V-Only) Condition*

*Consonant Data*

| | Bilabials | Labio-dentals | Inter-dentals | Rounded | Visible | Alveolars | Other | Nonvisible | Total |
|---|---|---|---|---|---|---|---|---|---|
| **Dunkin' Donuts** | 0 | 0 | 0 | 0 | 0 | 7 | 1 | **8** | **8** |
| **No thanks!** | 0 | 0 | 1 | 0 | 1 | 3 | 1 | **4** | **5** |
| **Waffles** | 1 | 1 | 0 | 0 | 2 | 2 | 0 | **2** | **4** |
| **Science museum** | 2 | 0 | 0 | 0 | 2 | 4 | 2 | **6** | **8** |
| **My school** | 1 | 0 | 0 | 0 | 1 | 2 | 1 | **3** | **4** |
| ***Total*** | 4 | 1 | 1 | 0 | **6** | 18 | 5 | **23** | **29** |
| ***Mean*** | | | | | **1.2** | | | **4.6** | **5.8** |

*Vowel Data*

| | Rounded | Diphthong | Visible | Unrounded* | Nonvisible | Total Vowels | Total Syllables |
|---|---|---|---|---|---|---|---|
| **Dunkin' Donuts** | 1 | 0 | 1 | 3 | 3 | 4 | **4** |
| **No thanks!** | 1 | 0 | 1 | 1 | 1 | 2 | **2** |
| **Waffles** | 0 | 0 | 0 | 1 | 1 | 1 | **2** |
| **Science museum** | 1 | 1 | 2 | 3 | 3 | 5 | **5** |
| **My school** | 1 | 1 | 2 | 0 | 0 | 2 | **2** |
| ***Total*** | **4** | **2** | **6** | **8** | **8** | **14** | **15** |
| ***Mean*** | | | **1.2** | | **1.6** | **2.8** | |
| *Unrounded includes spread and neutral vowels | | | | | | | |

*Set 5 – $Control_2$ Condition*

*Consonant Data*

| | Bilabials | Labio-dentals | Inter-dentals | Rounded | Visible | Alveolars | Other | Nonvisible | Total |
|---|---|---|---|---|---|---|---|---|---|
| **M (friend)** | 1 | 0 | 0 | 0 | 1 | 1 | 2 | **3** | **4** |
| **EW (name)** | 1 | 0 | 0 | 1 | 2 | 4 | 0 | **4** | **6** |
| **Legos** | 0 | 0 | 0 | 0 | 0 | 2 | 1 | **3** | **3** |
| **Uncle B** | 1 | 0 | 0 | 1 | 2 | 3 | 1 | **4** | **6** |
| **Hippos** | 1 | 0 | 0 | 0 | 1 | 1 | 1 | **2** | **3** |
| ***Total*** | 4 | 0 | 0 | 2 | **6** | 11 | 5 | **16** | **22** |
| ***Mean*** | | | | | **1.2** | | | **3.2** | **4.4** |

*Vowel Data*

| | Rounded | Diphthong | Visible | Unrounded* | Nonvisible | Total Vowels | Total Syllables |
|---|---|---|---|---|---|---|---|
| **M (friend)** | 0 | 0 | 0 | 2 | 2 | 2 | **2** |
| **EW (name)** | 0 | 1 | 1 | 3 | 3 | 4 | **5** |
| **Legos** | 1 | 0 | 1 | 1 | 1 | 2 | **2** |
| **Uncle B** | 0 | 1 | 1 | 1 | 1 | 2 | **4** |
| **Hippos** | 1 | 0 | 1 | 1 | 1 | 2 | **2** |
| ***Total*** | **2** | **2** | **4** | **8** | **8** | **12** | **15** |
| ***Mean*** | | | **0.8** | | **1.6** | **2.4** | |
| *Unrounded includes spread and neutral vowels | | | | | | | |

Participant 003 Target Sets

*Set 1 – Control$_1$ Condition*

*Consonant Data*

| | Bilabials | Labio-dentals | Inter-dentals | Rounded | Visible | Alveolars | Other | Nonvisible | Total |
|---|---|---|---|---|---|---|---|---|---|
| **Mr. XX (E)** | 1 | 0 | 0 | 1 | 2 | 3 | 1 | **4** | **6** |
| **What's up** | 2 | 0 | 0 | 0 | 2 | 2 | 0 | **2** | **4** |
| **I wanna be the door holder** | 2 | 0 | 1 | 0 | 3 | 4 | 1 | **5** | **8** |
| **Slam dunk** | 1 | 0 | 0 | 0 | 1 | 3 | 2 | **5** | **6** |
| **String beans** | 1 | 0 | 0 | 1 | 2 | 4 | 1 | **5** | **7** |
| ***Total*** | 7 | 0 | 1 | 2 | **10** | 16 | 5 | **21** | **31** |
| ***Mean*** | | | | | **2** | | | **4.2** | **6.2** |

*Vowel Data*

| | Rounded | Diphthong | Visible | Unrounded* | Nonvisible | Total Vowels | Total Syllables |
|---|---|---|---|---|---|---|---|
| **Mr. XX (E)** | 2 | 0 | 2 | 2 | 2 | 4 | **4** |
| **What's up** | 0 | 0 | 0 | 2 | 2 | 2 | **2** |
| **I wanna be the door holder** | 2 | 1 | 3 | 4 | 4 | 7 | **7** |
| **Slam dunk** | 0 | 0 | 0 | 2 | 2 | 2 | **2** |
| **String beans** | 0 | 0 | 0 | 2 | 2 | 2 | **2** |
| ***Total*** | **4** | **1** | **5** | **12** | **12** | **17** | **17** |
| ***Mean*** | | | **1** | | **2.4** | **3.4** | |
| *Unrounded includes spread and neutral vowels | | | | | | | |

*Set 2 – Visual Only (V-Only) Condition*

*Consonant Data*

| | Bilabials | Labio-dentals | Inter-dentals | Rounded | Visible | Alveolars | Other | Nonvisible | Total |
|---|---|---|---|---|---|---|---|---|---|
| **X Road** | 0 | 0 | 0 | 3 | 3 | 2 | 0 | **2** | **5** |
| **Pastor XX** | 1 | 0 | 0 | 0 | 1 | 3 | 1 | **4** | **5** |
| **Juice Box** | 1 | 0 | 0 | 1 | 2 | 2 | 1 | **3** | **5** |
| **Yes please** | 1 | 0 | 0 | 0 | 1 | 3 | 1 | **4** | **5** |
| **Hallelujah** | 0 | 0 | 0 | 0 | 0 | 2 | 2 | **4** | **4** |
| ***Total*** | 3 | 0 | 0 | 4 | **7** | 12 | 5 | **17** | **24** |
| ***Mean*** | | | | | **1.4** | | | **3.4** | **4.8** |

*Vowel Data*

| | Rounded | Diphthong | Visible | Unrounded* | Nonvisible | Total Vowels | Total Syllables |
|---|---|---|---|---|---|---|---|
| **X Road** | 1 | 0 | 1 | 1 | 1 | 2 | **2** |
| **Pastor XX** | 1 | 0 | 1 | 3 | 3 | 4 | **4** |
| **Juice box** | 1 | 0 | 1 | 1 | 1 | 2 | **2** |
| **Yes please** | 0 | 0 | 0 | 2 | 2 | 2 | **2** |
| **Hallelujah** | 1 | 0 | 1 | 3 | 3 | 4 | **4** |
| ***Total*** | **4** | **0** | **4** | **10** | **10** | **14** | **14** |
| ***Mean*** | | | **0.8** | | **2** | **2.8** | |
| *Unrounded includes spread and neutral vowels | | | | | | | |

*Set 3 – Audio-Visual (A+V) Condition*

*Consonant Data*

| | Bilabials | Labio-dentals | Inter-dentals | Rounded | Visible | Alveolars | Other | Nonvisible | Total |
|---|---|---|---|---|---|---|---|---|---|
| **XX (J)** | 0 | 0 | 0 | 1 | 1 | 3 | 0 | **3** | **4** |
| **Gimme my change** | 2 | 0 | 0 | 2 | 4 | 1 | 1 | **2** | **6** |
| **Basketball** | 2 | 0 | 0 | 0 | 2 | 3 | 1 | **4** | **6** |
| **XX (L)** | 0 | 0 | 0 | 0 | 0 | 1 | 0 | **1** | **1** |
| **Where's my tie?** | 2 | 0 | 0 | 0 | 2 | 2 | 0 | **2** | **4** |
| ***Total*** | 6 | 0 | 0 | 3 | **9** | 10 | 2 | **12** | **21** |
| ***Mean*** | | | | | **1.8** | | | **2.4** | **4.2** |

*Vowel Data*

| | Rounded | Diphthong | Visible | Unrounded* | Nonvisible | Total Vowels | Total Syllables |
|---|---|---|---|---|---|---|---|
| **XX (J)** | 0 | 0 | 0 | 2 | 2 | 2 | **2** |
| **Gimme my change** | 0 | 1 | 1 | 3 | 3 | 4 | **4** |
| **Basketball** | 0 | 0 | 0 | 3 | 3 | 3 | **3** |
| **XX (L)** | 0 | 0 | 0 | 2 | 2 | 2 | **2** |
| **Where's my tie?** | 1 | 2 | 3 | 1 | 1 | 4 | **4** |
| ***Total*** | **1** | **3** | **4** | **11** | **11** | **15** | **15** |
| ***Mean*** | | | **0.8** | | **2.2** | **3** | |
| *Unrounded includes spread and neutral vowels | | | | | | | |

*Set 4 – $Control_2$ Condition*

*Consonant Data*

| | Bilabials | Labio-dentals | Inter-dentals | Rounded | Visible | Alveolars | Other | Nonvisible | Total |
|---|---|---|---|---|---|---|---|---|---|
| **No thanks** | 0 | 0 | 1 | 0 | 1 | 3 | 1 | **4** | **5** |
| **Take a shot** | 0 | 0 | 0 | 1 | 1 | 2 | 1 | **3** | **4** |
| **Where's my suit jacket?** | 2 | 0 | 0 | 1 | 3 | 4 | 1 | **5** | **8** |
| **Church** | 0 | 0 | 0 | 2 | 2 | 0 | 0 | **0** | **2** |
| **Mrs. XX (W)** | 2 | 0 | 0 | 1 | 3 | 3 | 0 | **3** | **6** |
| ***Total*** | 4 | 0 | 1 | 5 | **10** | 12 | 3 | **15** | **25** |
| ***Mean*** | | | | | **2** | | | **3** | **5** |

*Vowel Data*

| | Rounded | Diphthong | Visible | Unrounded* | Nonvisible | Total Vowels | Total Syllables |
|---|---|---|---|---|---|---|---|
| **No thanks** | 1 | 0 | 1 | 1 | 1 | 2 | **2** |
| **Take a shot** | 0 | 0 | 0 | 3 | 3 | 3 | **3** |
| **Where's my suit jacket** | 2 | 1 | 3 | 3 | 3 | 6 | **5** |
| **Church** | 0 | 0 | 1 | 0 | 0 | 1 | **1** |
| **Mrs. XX (W)** | 0 | 0 | 0 | 4 | 4 | 4 | **4** |
| ***Total*** | **3** | **1** | **4** | **11** | **11** | **15** | **15** |
| ***Mean*** | | | **0.8** | | **2.2** | **3** | |
| *Unrounded includes spread and neutral vowels | | | | | | | |

*Set 5 – Audio-Only (A-Only) Condition*

*Consonant Data*

| | Bilabials | Labio-dentals | Inter-dentals | Rounded | Visible | Alveolars | Other | Nonvisible | Total |
|---|---|---|---|---|---|---|---|---|---|
| **Watch T.V.** | 1 | 1 | 0 | 1 | 3 | 1 | 0 | **1** | **4** |
| **Grilled cheese** | 0 | 0 | 0 | 2 | 2 | 3 | 1 | **4** | **6** |
| **XX (Jo)** | 0 | 0 | 0 | 1 | 1 | 2 | 0 | **2** | **3** |
| **Praise the Lord** | 1 | 0 | 1 | 1 | 3 | 3 | 0 | **3** | **6** |
| **I like drumming** | 1 | 0 | 0 | 1 | 2 | 2 | 0 | **2** | **4** |
| ***Total*** | 3 | 1 | 1 | 6 | **11** | 11 | 1 | **12** | **23** |
| ***Mean*** | | | | | **2.2** | | | **2.4** | **4.6** |

*Vowel Data*

| | Rounded | Diphthong | Visible | Unrounded* | Nonvisible | Total Vowels | Total Syllables |
|---|---|---|---|---|---|---|---|
| **Watch T.V.** | 0 | 0 | 0 | 3 | 3 | 3 | **3** |
| **Grilled cheese** | 0 | 0 | 0 | 2 | 2 | 2 | **2** |
| **XX (Jo)** | 1 | 0 | 1 | 1 | 1 | 2 | **2** |
| **Praise the Lord** | 2 | 0 | 2 | 2 | 2 | 4 | **3** |
| **I like drumming** | 0 | 2 | 2 | 2 | 2 | 4 | **4** |
| ***Total*** | **3** | **2** | **5** | **10** | **10** | **15** | **14** |
| ***Mean*** | | | **1** | | **2** | **3** | |
| *Unrounded includes spread and neutral vowels | | | | | | | |

## APPENDIX B

TREATMENT PROTOCOL

Steps for each treatment episode are outlined in the Protocol Flowchart below.

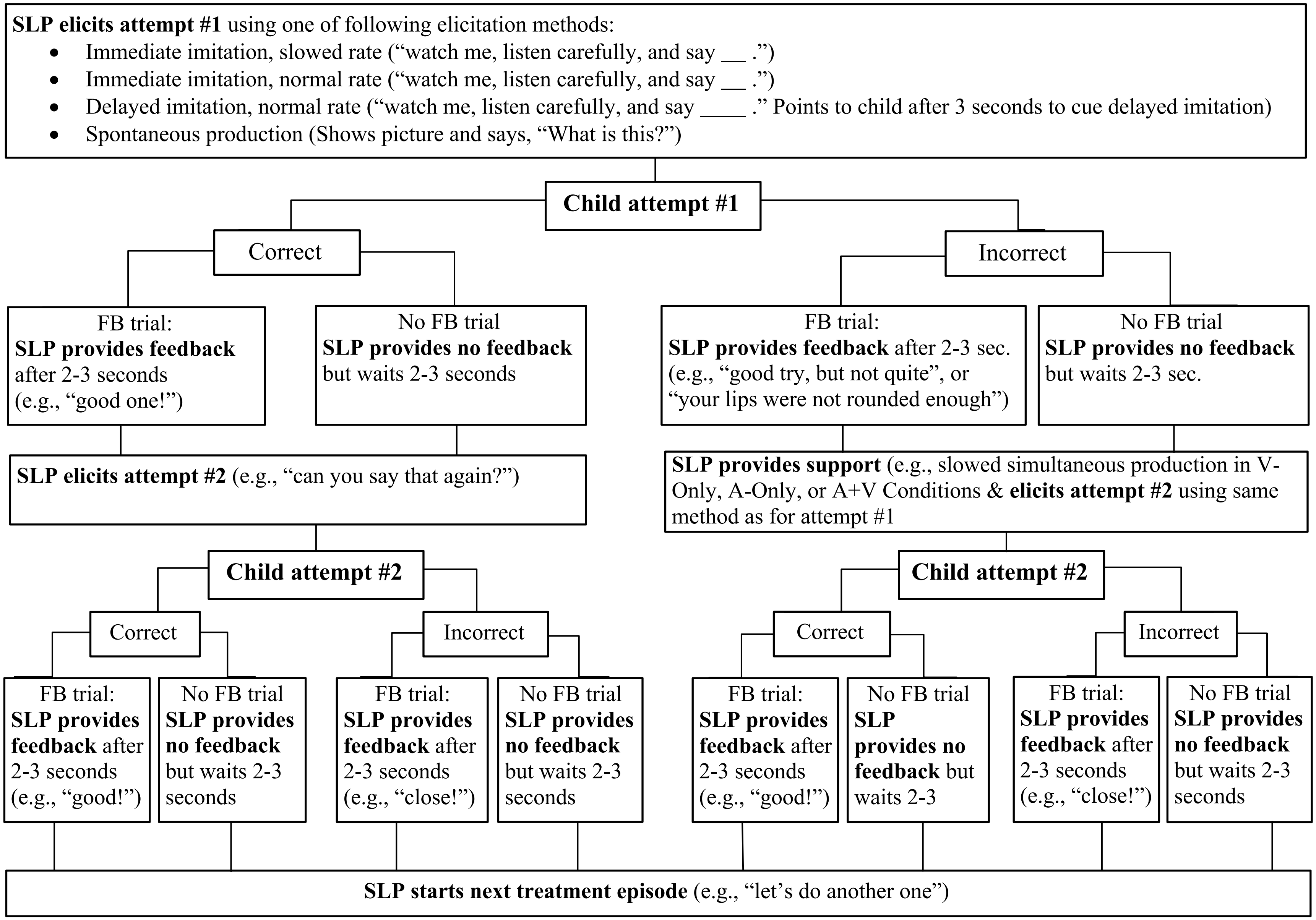
In each treatment condition, there are five targets (15 total).

Treatment episodes are presented in random order within all conditions. Random order of treatment episodes is implemented by shuffling the cards before the treatment session and reshuffling the deck as needed during the session once every card has been used.

Feedback is given with decreasing frequency during each condition. Feedback is provided for 100% (5) of the first five treatment episodes, 80% (4) of the next five treatment episodes, and so on until feedback is given on 20% (1) of the last five treatment episodes. To facilitate keeping track of feedback schedules, the SLP has a feedback tracking sheet with 100 slots with the fading feedback schedule marked (Ballard et al., 2007).

Start each condition in each session with immediate imitation at normal rate. The criterion to increase difficulty level of elicitation condition is 2/2 consecutive correct *Attempts #1*. That is, if the child produces a correct response on *Attempt #1* on two successive treatment episodes (whether these are the same target or not), then the third treatment episode begins with the next level elicitation method. Similarly, if the child produces 2/2 consecutive incorrect *Attempts #1* with an elicitation method, the next treatment episode reverts to the previous difficulty level of elicitation method (or stay in immediate imitation at slow rate).

Continue until session time has elapsed (determined by an egg-timer set to 15-minutes per condition). After a 5-minute break, begin the second treatment condition and continue until session time is up.


SLP elicits attempt #1 using one of following elicitation methods:
• Immediate imitation, slowed 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 ____ ." Points to child after 3 seconds to cue delayed imitation)
• Spontaneous production (Shows picture and says, "What is this?")
Child attempt #1
Correct
Incorrect
FB trial: SLP provides feedback after 2-3 seconds (e.g., "good one!")
No FB trial SLP provides no feedback but waits 2-3 seconds
FB trial: SLP provides feedback after 2-3 sec. (e.g., "good try, but not quite", or "your lips were not rounded enough")
No FB trial SLP provides no feedback but waits 2-3 sec.
SLP elicits attempt #2 (e.g., "can you say that again?")
SLP provides support (e.g., slowed simultaneous production in V-Only, A-Only, or A+V Conditions & elicits attempt #2 using same method as for attempt #1
Child attempt #2
Correct
Incorrect
FB trial: SLP provides feedback after 2-3 seconds (e.g., "good!")
No FB trial SLP provides no feedback but waits 2-3 seconds
FB trial: SLP provides feedback after 2-3 seconds (e.g., "close!")
No FB trial SLP provides no feedback but waits 2-3 seconds
Child attempt #2
Correct
Incorrect
FB trial: SLP provides feedback after 2-3 seconds (e.g., "good!")
No FB trial SLP provides no feedback but waits 2-3
FB trial: SLP provides feedback after 2-3 seconds (e.g., "close!")
No FB trial SLP provides no feedback but waits 2-3 seconds
SLP starts next treatment episode (e.g., "let's do another one")

## APPENDIX C

### PROBE DATA ANALYSIS INSTRUCTIONS

*The Scoring System*

*Background & Overview*

For our purposes, the scoring system has to meet several goals:

(a) It has to be <u>valid</u>. For our purposes that involves two things: (1) it has to be able to differentiate minor from major errors, and (2) it has to be sensitive to change – i.e., it has to be able to detect changes (e.g., due to treatment).

(b) It has to be <u>reliable</u>. 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. We will check this, of course.

(c) It has to be (at least somewhat) <u>clinically relevant</u>. We are using perceptual judgments rather than acoustic measures, for example. This also means that it has to be manageable. We don't want to do narrow transcription but rather use a bit more of a holistic judgment of accuracy.

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). NOTE that we judge ***accuracy*** (not *intelligibility*) of the child's production.

*THE 3-way SCORING SYSTEM*

This scoring system has 3 levels:

2 = correct response
1 = response has one minor error
0 = incorrect response, involving more than one minor error

**<u>Score of 2 (correct):</u>**
Production must have <u>no errors</u>, must be <u>fluent</u>, and must <u>sound natural or normal</u> to you.

1. Errors include omissions, additions, substitutions, distortions (including excessive lengthening), incorrect stress patterns, sound or syllable segmentation ("choppy" speech), unintelligible utterances, etc.
2. Self-corrections are also considered errors because they indicate that the child was not able to produce a correct response on the first attempt. So even if a child self-corrects to a "correct" response, this would still not be scored a 2, but instead would be scored a 1 (assuming no other errors, normally fluent response, etc.).

**<u>Score of 1 (minor error):</u>**
Production must contain no more than one "minor error." In the Strand et al. (2006) paper, a minor error was defined as "one distinctive feature off consonant production, or close approximation of movement gesture, or only one mild vowel distortion" (p. 301). In other words,

a score of 1 represents a close approximation of the target. In our application of this system, we have modified and operationalized this a bit by also including judgment of suprasegmental accuracy and including self-corrections.

1. *Segmental*: Vowel and consonant distortions (including excessive lengthening) or consonant substitutions involving a difference of no more than a single phonetic feature (place of articulation, manner of articulation, or voicing). Distortions are defined as sounds that are recognizable as the target sound but that are not clear, well-articulated examples of that target sound. Distortions may sound like a sound between the target and some other sound (e.g., a vowel halfway between /æ/ and /ɛ/, a consonant halfway between /s/ and /ʃ/) but sometimes they do not approximate another target (e.g., a lateral lisp).
2. *Suprasegmental*: Equalized stress patterns, exaggerated stress patterns, sound or syllable segmentation ("choppy" speech). Suprasegmental errors only apply to words or phrases with more than one syllable.
3. Self-corrections are also considered errors, because they indicate that the child was not able to produce a correct response on the first attempt. So, if a child self-corrects to a "minor error" response this would still not be scored a 1, but instead would be scored a 0.

**Score of 0 (incorrect):**

Production contains more than one minor error. This includes any combination of two or more minor errors (as defined above), or one or more major error. Major errors can be segmental or suprasegmental:

1. *Segmental*: Any omission or addition of a consonant or vowel (or syllable), any substitution of a consonant that differs from the target by more than one phonetic feature, and vowel substitutions (sounds like a good example of a different vowel). Also includes sound reversals/transpositions (e.g., /kæt/ → /tæk/), because that is essentially two errors.
2. *Suprasegmental*: Incorrect stress pattern, i.e. stress on the wrong syllable (not merely equalized or exaggerated stress). Applies only to words or phrases consisting of more than one syllable.

| SCORE = 0 | Examples | SCORE = 1 | Examples |
|---|---|---|---|
| More than one sound in error (distorted, added, substituted, omitted) | kæt → gæd<br>ˈbʌtɚflaɪ → ˈbætɚfaɪ<br>kæt → tæk | | |
| Omission of sound or syllable | kæt → kæ<br>ˈbʌtɚflaɪ → ˈbʌflaɪ<br>ˈbʌtɚflaɪ → ˈbʌtɚfaɪ | | |
| Addition of sound or syllable | kæt → kætəl<br>ˈbʌtɚflaɪ → ˈbʌtɚflaɪp | Consonant distortion: Sounds like bad example of target consonant, or between target and other consonant (e.g., excessive aspiration; lateral lisp, excessive lengthening) | kæt → kætʰ<br>sʌn → s̪ʌn<br>sʌn → s̪ʌn s/θ?<br>ˈbʌtɚflaɪ → ˈbʌtɚf:laɪ<br>**NOT:**<br>sʌn → θʌn<br>(substitution error = 0) |
| Vowel "substitution": Sounds like good example of different vowel. | kæt → kɪt<br>ˈbʌtɚflaɪ → ˈbætɚflaɪ | Vowel "distortion": Sounds like bad example of target vowel, or between target and other vowel.<br>Also includes excessive lengthening. | kæt → kæt æ/ɛ?<br>ˈbʌtɚflaɪ → ˈbʌtɚflaɪ ʌ/ɪ?<br>kæt → kæ:t<br>**NOT:**<br>kæt → kæs æ/ɛ?<br>ˈbʌtɚflaɪ → ˈbʌtɚfaɪ ʌ/ɪ? |
| Major consonant error: More than 1 feature off target | kæt → kæf, bæt, gæd<br>ˈbʌtɚflaɪ → ˈbʌzɚflaɪ | Minor consonant error: No more than 1 feature off target for 1 consonant (may be perceived as substitution). | kæt → gæt, kæp<br>ˈbʌtɚflaɪ → ˈdʌtɚflaɪ, ˈmʌtɚflaɪ<br>**NOT:**<br>kæt → gæs, tæp<br>(2 minor errors = 0)<br>ˈbʌtɚflaɪ → ˈtʌtɚflaɪ<br>(2 features off target = 0) |
| Major prosodic error: Incorrect stress pattern (not merely equal stress) | ˈbʌtɚflaɪ → bətɚˈflaɪ<br>ˈbʌtɚflaɪ → ˈbʌtɚˈflaɪ<br>ˈbeɪbi → beɪˈbi | Minor prosodic error: Equal stress, exaggerated stress, syllable segregation, slow rate. | ˈbʌtɚflaɪ → ˈbʌˈtɝˈflaɪ<br>ˈbʌtɚflaɪ → ˈbʌ(.)tɚ(.)flaɪ<br>**NOT:**<br>ˈbʌtɚflaɪ → bʌ(.)tɚ(.)ˈflaɪ<br>(2 prododic errors = 0)<br>ˈbʌtɚflaɪ → bætɚˈflaɪ<br>(V error + prosodic error = 0) |
| Sound error plus prosodic error (e.g., syllable segregation, equal stress pattern) | ˈbʌtɚflaɪ → ˈbæˈtɝˈflaɪ<br>ˈbeɪbi → ˈbeɪ(.)ˈbi | Restart after partial attempt (no errors in first *complete* attempt) | kæt → t … kæt<br>kæt → k … kæt |
| Any other errors that do not qualify for a 1 (e.g., unintelligible; two minor prosodic errors) | | Self-correction (no errors in self-correction) | kæt → kæf… kæt |

*General Scoring Rules & Procedures*

1. The judgment is about ***accuracy*** (not *intelligibility*) of the child's production and encompasses both segmental and suprasegmental accuracy of the entire word or phrase.
2. Generally, **judge the first attempt** by a child. Self-corrections lose a point.
    - ***NOTE***: If there is overlap between speakers (clinician and child) so that you can't judge the child's attempt, then score the second attempt but subtract one point. The reason is that children are supposed to wait for the clinician to finish talking, and overlapping responses indicate failure to wait. So, if the second response is correct (would be a 2), then give only a 1. If the second response is a minor error (i.e. 1), then give a 0. Do not go below 0 (0 is the lowest possible score). *Make a note when this happens*.
3. **Dialectal variations or normal speech phenomena are not considered errors**. For example, some vowel reduction in unstressed syllables is normal, tapping of intersyllabic /t/ is normal, vowel nasalization before a nasal coda is normal, unreleased stop at the end of a word is normal, etc. In other words, you need to **use your knowledge of normal, non-disordered speech** and its contextual variations to make these judgments.
4. Targets are elicited via a repetition task, therefore ***judgment is relative to the model*** presented by the clinician. In other words, if the clinician uses atypical or exaggerated prosody and the child imitates that, this is not an error on the child's part. If the clinician adds or omits a word, then judge the child's response only in relation to the model. Be sure to *make a note when this happens*.
5. **Make notes** about observations that may be relevant. No need to go crazy with notes, but please make notes about the following:
    a. 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.
    b. 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. We nevertheless score these as speech errors, but we want to be able to separate these types of errors if needed, so please add a comment (LANG?).
    c. Clinician model differs from intended target (e.g., clinician says *'Miss B.'* for *'Mrs. B.'*).
    d. Anything that might affect validity of the scores. For example, background noise, poor recording quality, silly voice/yawning/giggling/crying, etc.
    e. No need to <u>provide a transcription</u> when scoring, in the interest of efficiency/time management.