The Effectiveness of Melodic Intonation Therapy with the Treatment of Childhood Apraxia of Speech

Amber Faith Worthington

Faculty Advisor: Dr. Virginia Hinton

Honors Faculty Liaison: Dr. Kristine Lundgren

Department of Communication Sciences and Disorders

The University of North Carolina - Greensboro
ABSTRACT

The purpose of this project was to create a literature review on Childhood Apraxia of Speech (CAS), Melodic Intonation Therapy (MIT) and the possible use of MIT as a treatment method for CAS. Included within this literature review is also a plan for a future study. The purpose of this study is to further examine the effectiveness of MIT for the treatment of CAS. In this study, three groups will be created, each consisting of 10 children with CAS. Each group will be exposed to a different therapy method that is used in the treatment of CAS. The three groups being; Touch Cue Method, Non-Speech Oral Motor Exercises (NSOMEs) and MIT. The phonetic inventory of each participant will be taken into account as a list of 30 stimulus words will be created that will be implemented into each therapy session. During the halfway point of the study, five new stimulus words will be introduced to test generalization skills for each treatment method. At the end of the study, five more new stimulus words will be introduced to again test generalization skills and to examine further if there is an improvement of these skills. There will be a total of 6 sessions in this study. Each session will be 60 minutes with a total of 360 minutes of exposure to each therapy method.
INTRODUCTION

Communication is a fundamental part of humanity. Often perceived as simplistic, communication is a crucial element for having a positive perspective on life. Communication can be described as the process of transmitting thoughts and ideas from the sender to the receiver. The two main aspects of communication are language and speech. Language is defined as “impaired comprehension and/or use of spoken, written, signed and/or other symbol systems” (Ostergren, 2019, p. 167).

There are five domains of language: phonology, morphology, syntax, semantics, and pragmatics. Phonology is described as the aspect of language that focuses on the ability to distinguish the phonemes within a language. Phonemes are the speech sounds that are in language (Language in Brief, n.d. para.7). Morphology comprises the rules that dictate the rules of combining speech sounds. Syntax includes the rules of grammar that are impeded into a language. Semantics encompasses the meanings of words. Lastly, pragmatics involves social communication that is necessary for devolving interpersonal bonds (“Language in Brief”, n.d. para. 4).

The other aspect of communication is speech. Speech involves the individual sounds that are part of words. Speech includes the three main elements: articulation, voice, and fluency (“What Is Speech? What is Language?”, n.d., para. 2). Articulation consists of the planning and programming of the speech production mechanisms to create speech sounds (Zuk et al., 2018, p. 583). Voice encompasses the physiological functions of the vocal folds, including respiratory functions. Lastly, fluency is the rhythm of speech, including pauses, repetitions, blocks, interjections, and prolongations (“What Is Speech? What is Language?”, n.d., para. 2).
Though communication breakdowns can be normal, it becomes concerning when breakdowns happen continuously, causing communication disorder to present itself. A communication disorder is defined as “an impairment in the ability to receive, send, process, and comprehend concepts or verbal, nonverbal and graphic symbol systems” (“Definitions of Communication Disorders and Variations”, 1993, p. 949). The focus of this project is childhood apraxia of speech (CAS) - an articulatory breakdown that prevents a child from developing normal speech production. This limits the ability to communicate effectively with others.

**CHILDHOOD APRAXIA OF SPEECH**

*Nature of the Disorder*

Childhood Apraxia of Speech (CAS) is a type of disorder that affects the “transmission” portion of communication, specifically, the neurological messages that dictate the muscles in the mouth to move in order to produce sound (“Childhood Apraxia of Speech,” n.d., para. 2). In the 2007 ASHA position, CAS, is classified as the disorder as having three core features: “inconsistent errors on consonants and vowels in repeated productions of syllables or words, lengthened and disrupted coarticulatory transitions between sounds and syllables, and inappropriate prosody, especially in the realization of lexical or phrasal stress” (American Speech-Language-Hearing Association, 2007, para.11).

One of the key features to CAS is the inability to generate a motor plan in the brain to produce appropriate sound sequences. The motor plan is the process that directs the muscles of respiration, phonation, and articulation to contract in a specific order so that intelligible speech sounds are produced. Dewey, Roy, Square-Storer, and Hayden (1988) assessed the motor planning capabilities of those with apraxia of speech. Their results indicated that children with CAS exhibited difficulties with transitions in sequences of motor tasks. In 2015, a study was
conducted to examine “spatial and temporal aspects of articulatory control in children with CAS, children with speech delay characterized by an articulation or phonological impairment, and typical development” (p.1103) particularly focusing on “movement duration, displacement, velocity, and variability” (Grigos et al., 2015, 1113). This study also recognized the influence that task complexity can have on speech output.

**Reported comorbidities**

In addition to difficulties with motor planning for speech, children with apraxia may exhibit cognitive differences compared to children without the disorder. In 2015, a study was conducted which demonstrated that children with CAS were impaired on “complex sensorimotor and sequential memory functions” as well as “had concomitant deficits on sequential functions” (Nijland et al., 2015, p. 559). These findings further accept the idea that the motor-based feature of CAS does stem from sequential functioning issues.

In addition to these features, children with CAS may also experience speech inconsistencies. Iuzzini-Seigel, Hogan, and Green sought to determine if speech inconsistency is a core feature of CAS or if it can be attributed to comorbidity of a language impairment. For this study two groups were studied, those who have only CAS and those who have CAS with a language impairment. The findings concluded that speech inconsistency is a core feature of CAS since the level of inconsistency was similar for both groups (Iuzzini-Seigel et al., 2017, p.1206). Though these inconsistent errors on phonemes imply a deficiency in articulation control, there is still the question of which specific domains of movement are specifically deficient in CAS.

In a previous study, Sadagopan and Smith (2008) examined the variability of articulation movement as utterance length increased. They found that in children and adolescents, articulation movement variability increased when the phrase “buy Bobby a puppy” was
incorporated into a larger utterance than when produced in isolation. However, this study did not include participants with CAS in its research pool. Moss and Gigos (2012) expanded the concept in a study that included the participants with CAS, articulation/phonological impairment as well as typical development. They found that word length may affect articulator movement with a higher rate for children with compared to typically developing children or children with articulation disorders (Moss & Grigos, 2012, p. 132). To further examine articulatory control, the motion capture system, Vicon 460, was used to track lip, jaw, and head movement on the participants (Grigos et al., 2015, p.1108). The study differentiated the three groups of children by movement duration and variability, however, “regarding duration, both groups of children with speech impairment produced longer jaw movement durations than children in the typical development group” (Grigos et al., 2015, p. 1113). This research indicates that the speech production difficulties that are a characteristic feature of CAS may arise from speech motor deficits that make accurate speech challenging to maintain.

While previous studies have demonstrated that speech motor deficits are present in children with CAS, however, the question remains if the “core phenotype is limited to transcoding (planning/programming) deficits or if speakers with CAS also have deficits in auditory perceptual encoding (representational) and/or memory (storage and retrieval of representations) processes” (Shriberg et al., 2012, p. 445). Transcoding is the process of planning the motor processes that must be present in order to produce fluid and accurate speech. Auditory perceptual encoding involves processes that alter auditory input into phonemic representation and memory is retrieving those representations. A study was conducted to determine what the core domains of this disorder are. Four groups were included: typical language development; speech delay but later typical language; speech delay but later language impairment; and CAS.
The researchers measured the three domains by using the Syllable Repetition Task (STR) (Shriberg et al., 2012, p. 445). STR is a language task that tests the speaker’s ability to repeat nonsense words (Shriberg et al., 2009, p. 1190). The findings from this study are consistent with the belief that CAS has deficits in all three domains: encoding, memory, and transcoding processes. The participants with CAS had significantly lower “memory, and transcoding scores than participants in each of the three comparison groups, and significantly lower encoding scores than participants” who had typical speech and those with a language delay that now had typical speech (Shriberg et al., 2012, p. 474). Thus the children with CAS had the core features of deficiency in the domains of auditory-perceptual encoding, memory, and transcoding. A second conclusion is that these findings are “consistent with literature findings indicating primarily quantitative, rather than conclusive differences in the speech, prosody, and voice patterns of participants with CAS compared to participants with other moderate to severe speech sound disorders” (Shriberg et al., 2012, p. 477). Many researchers have difficulty classifying CAS. This research further supports “for the etiological classification framework for speech sound disorders, including CAS…compared to other organizational proposals for congenital, environmental, and acquired speech disorders based on taxonomic similarities in speech patterns” (Shriberg et al., 2012, p. 477).

**Treatment Options**

One treatment option that is used when working with cases of CAS is the Touch Cue Method. This is a method is based “based on the assumption that the child has difficulty in firmly establishing and integrating voluntary oral motor movements necessary for purposes of speech sound production” (Bashir et al., 1984, p. 128). By using touch points on the face, particularly around the mouth, the therapist is firmly establishing the different motor movements needed for
speech. For this technique, three stages are used. The first stage “uses a series of drills involving nonsense syllables to teach the topographic cues, improve articulatory sequencing, and develop accurate self-monitoring of production (Bashir et al., 1984, p. 129). This first step is used to increase the child’s awareness of the differences between the target sounds and the produced sounds. The second stage “drills that incorporate the previously learned sequential movements into mono and polysyllabic words. Monosyllabic words having the CVC structure are used” (Bashir et al., 1984, p. 133). These words will include both real and nonsense words. Part of this stage is to focus on the use of correct prosody when using polysyllabic words. The overall goal for this stage is aid in the sequencing between different points of articulation. The last stage “applies previously learned sequencing and self-monitoring skills, first to controlled multiword utterances, and then to spontaneous speech” (Bashir et al., 1984, p. 135). The main part of this stage is the transition of control from the clinician to the client. This is to work on the self-monitoring skills that the client must use in spontaneous speech. Eventually, the therapy method will be terminated once the client has met two conditions. The first being “speech is intelligible all of the time” and the second being “the client can monitor and, subsequently, self-correct speech most of the time” (Bashir et al., 1984, p. 136).

One type of therapy method that has been used in the treatment of CAS are non-speech oral motor exercises (NSOMEs). In 2008, a nationwide survey was conducted to research the reoccurrence of this treatment method in therapy. With 27% of speech language pathologists (SLPs) completing this survey, it was discovered 85% of SLPs used NSOMEs “to address speech sound production problems” (Loft & Watson, 2008, p. 394). NSOMEs are oral exercises that are meant to build muscle mass in the mouth. This includes but is not limited to “blowing, tongue wagging and pushups, cheek puffing, the alternating movement of pucker-smile, “big
“smile,” and tongue-to-nose-to-chin” (Loft, 2008, p. 253). These exercises are frequently used in therapy rooms; around 85% of speech pathologists in the United States (Loft & Watson, 2008, p. 394). One of the most attractive things about these exercises is that they are advertised to diminish any muscle weakness in the oral cavity. Though it is true that these exercises do build up muscle mass, it is unnecessary for effective speech. In fact “…only a fraction of maximum tongue force is used in speech production, and such strength tasks are not representative of the tongue's role during typical speaking. As a result, caution should be taken when directly associating tongue strength to speech…” (Wenke et al., 2006, p. 2). However, it is necessary to recognize that NSOMEs do not have evidence-based practice to support their use. Firstly, these exercises should not be used in cases of childhood apraxia since it is a neurological disorder. In these cases, the oral muscles do not experience any atrophy or weakness. Therefore, it would be illogical to use muscle-based exercises. Another issue is that the movement of the tongue outside of the mouth is a core feature of NSOMEs. The issue is that for normal speech, the tongue remains within the oral cavity when producing speech sounds. Therefore, it seems implausible to spend therapy time with the tongue outside of the mouth. Instead, techniques such as phonetic placement as well as using visual and tactile cues should be used to produce correct speech. Due to the fact that these treatments are not evidence-based yet they are highly used in treatment plans, other treatment methods must be explored.

**Long Term Outcomes**

Similar to the lack of research for therapy options, there have been limited longitudinal studies for the prognosis of those with CAS. However, since CAS is classified as a persistent speech sound disorder, there are risks for academic, vocational and social difficulties. In 2017, a study was conducted to test the potential of reading difficulties students with language
impairments as well as persistent speech sound disorders, including CAS. After screening the participants reading comprehension skills, it was shown that children with functional speech disorders, language impairments as well as CAS were at a higher risk for reading difficulties (Zipoli & Merritt, 2017, p. 101). Since students with CAS could have the potential for reading based difficulties, this continues to emphasize the need for an appropriate therapy technique when treating these cases.

**Introduction of Melodic Intonation Therapy**

A technique that is gaining popularity for the treatment of CAS is Melodic Intonation Therapy (MIT). MIT was initially designed to treat severe nonfluent aphasic patients (Norton et al., 2009, p. 431). This treatment uses the natural prosody of language “to improve expressive language by capitalizing on preserved function (singing) and engaging language-capable regions in the undamaged right hemisphere” (Norton et al., 2009, p. 431). Many studies have been conducted to test the effectiveness of MIT for cases of Aphasia. In 2013, a study was done to research the effectiveness of MIT when treating expressive language. This study followed the case of a 65-year-old participant (KM) who experienced a “unilateral, right hemisphere, intracerebral, hemorrhagic stroke during the embolization of an arteriovenous malformation” (Morrow-Odom and Swann, 2013, p. 1325). KM received a total of 32 sessions over seven weeks. During those sessions, an hour and a half were dedicated to treatment with MIT and 20 minutes for cognitive activities. The words that were focused on were life-sustaining words as well as personal words such as the names of her family members (Morrow-Odom and Swann, 2013, p. 1329-1331). After the treatment, her scores showed significant improvement. Though she still had global aphasia, her aphasia severity score increased from 72 to 77 with a standard error of plus or minus one. Her alternative communication score also showed significant
improvement, with her score rising from 69 to 81 with a standard error of plus or minus 0.8
(Morrow-Odom and Swann, 2013, p. 1333). These significant increases of scores show that there
is a potential of this treatment method working outside of the original confines of MIT.

The treatment typically involves the clinician asking the client to sing a target phrase
while tapping of the left hand, eventually having the support of the clinician decrease until the
patient sings the phrase alone. The last step would involve the transition from singing to
speaking the phrase (Van der Meulen et al., 2012, p. S46). Though the treatment seems simple
by nature, it is incredibly complex. MIT uses four aspects; intonation, tapping, inner rehearsal,
and auditory-motor feedback training; to produce intelligible speech. Intonation is the core of
MIT. Overall, MIT tries to increase reliance on the right hemisphere of the brain. Studies have
proven that the right hemisphere of the brain predominately deal in the processing of song
prosody and therefore a “slower rate of articulation and continuous voicing that increases
connectedness between syllables and words in singing, may reduce dependence on the left
hemisphere” (Norton et al., 2009, p.433). The tapping of the left hand also helps reduce the
dependence on the left hemisphere of the brain. The tapping of the left hand engages a right-
hemisphere sensorimotor network that controls both hand and mouth movements (Gentilucci,
2008, p. 944). By engaging this part of the brain, it is theorized that fine motor movement
necessary for speech would be more precise. Inner rehearsal is especially necessary for apraxic
speech. If a client can practice “silently intoning the target phrase may reinitiate a cascade of
activation from a higher level in the cognitive-linguistic architecture thus giving the speaker
another attempt to sequence the motor commands correctly” (Norton et al., 2009, p.433). The
last aspect of MIT is auditory-motor feedback training. For clients who are aphasic or apraxic, it
may be increasingly difficult to process auditory feedback efficiently to self-correct. However, if
words are sung, “phonemes are isolated and thus, can be heard distinctly while still connected to the word. In addition, sustained vowel sounds provide time to ‘think ahead’ about the next sound, make internal comparisons to the target, and self-correct when sounds produced begin to go awry” (Norton et al., 2009, p.433). Overall, the process of MIT is aimed at increasing reliance on the right hemisphere of the brain while at the same time giving the client ample opportunity to self-correct.

Since MIT allows for the client to have an increased amount of time for self-correction, the incorporation of MIT for the intervention of CAS has been brought into question. In 2011, a study was conducted to test the effectiveness of MIT and Touch Cue Method (TCM). TCM is a therapy technique that involves “the sequencing of speech sounds is supported by giving touch cues on the child’s face and neck” (Martikainen & Korpilahti, 2011, p.11). The study used MIT for the first 6-weeks of treatment and TCM for the second 6-weeks of treatment. The study found that “speech sound errors decreased and sequencing abilities increased significantly after the MIT period, and the positive progression continued during the TCM period” (Martikainen & Korpilahti, 2011, p.9). This study did not test the generalization skills of MIT since treatment was followed immediately by treatment of TCM. The conclusion for this study is that MIT and TCM is an appropriate treatment option for CAS intervention. However, research on MIT is still needed to know the effectiveness of this treatment method individually.

RESEARCH PLAN

Purpose

Based on the limited implementation related to MIT as an effective treatment for CAS, future information is needed before it can be designated as an evidence-based practice for this disorder. A wide range of participants will be included in this study, after completing a variety of
tests to conclude eligibility. In this study, therapy sessions would be conducted to test the effectiveness of MIT on children with CAS by using different types of stimulus, hopefully concluding in positive clinical implications. To determine if MIT effectively treats cases of CAS, three groups will be created, each using a different therapy method; MIT, Non-speech Oral Motor Exercises (NSOME), and Touch Cue Method.

METHOD

Participants

Four criteria must first be met. The first criteria would be based on passing an audiometric screening. Next, corrected visual acuity is necessary to interact with stimuli appropriately. Native American English is also required to reduce variability resulting from the production of non-English words. Lastly, the participants must show characteristics of CAS that are defined by ASHA in their technical report on the disorder (“Childhood Apraxia of Speech: Technical Report”, 2007 para. 20). Assessments would also need to be conducted to qualify for participation. The first assessment administered would be a hearing screening ranging from 25 dB for 1000-4000 Hz. This would be to ensure that all participants had normal levels of hearing to properly hear the instructions and stimuli that are given in this study. An oral mechanism screening should also be administered to ensure that physical abnormalities were present that could interfere with pronunciation. A variety of standardized tests may also have to be given to the potential participants to determine inclusion criterion. The first test given will be the Primary Test of Nonverbal Intelligence (PTONI). This test is to be administered to ensure that the participants have the cognitive ability to complete the necessary tasks asked of them for the study. The next standardized test will be The Test of Auditory Comprehension of Language (TACL) to assess the receptive language skills necessary to complete the tasks asked of them.
Running Header: The Effectiveness of Melodic Intonation Therapy with the Treatment of Childhood Apraxia of Speech

during the study. *The Clinical Assessment of Articulation and Phonology* (CAAP) will be used to measure the speech pronunciation skills for single words. CAAP can also be used to assess the consonant inventory of the participants. This will be necessary for determining appropriate stimulus words/phrases for each participant. Finally, *The Kaufman Speech Praxis Test* for Children will be needed to determine the vowel inventory of participants as well as the complexity of syllables produced and accuracy for the production of nonsense words. The accuracy of nonsense words is necessary to understand the variability of CAS and its effect on the output of actual words. *Diadochokinesis* (DDK) should also be performed to establish syllable repetitions. After this screening is conducted, I hope to have a minimum of 30 participants; therefore, being able to split them into the three groups of 10.

**PROCEDURE**

**Stimulus**

All groups will use the same stimulus determination method. To determine the stimulus words to be used in the study, the phonemic inventory of the participants will be taken into consideration to ensure that each participant is given stimuli containing consonants that they can produce. From these personal phonemic inventories, 80-100 stimuli, words/phrases will be created. A random number generator will choose 40 of the words/phrases. Ten of these words will be used as probes at the end of sessions to test the generalization of skills to unknown words/phrases. Ultimately, the 30 stimuli words/phrases will be taught using MIT, NSOME, or Touch Cue Method during the therapy sessions. These 30 stimuli will be used in every session; the only new words that will be introduced will be the ten probe words/phrases. These probes will be equally distributed in the study. The first use of the probe phrases will be at the midway point of the study. The second use will be at the end of the study. This is to examine the possible
improvement of generalization skills. The session should be 60 minutes weekly with 360 minutes of treatment in total, around six sessions.

**Results**

The results will be measured by two standards: articulation accuracy and generalization accuracy. Articulation accuracy will be used to determine the accuracy of stimuli words/phrases. Generalization accuracy is to be used for probe words/phrases. The mean performance accuracy = accurate productions/by the total number of outputs. Mean performance accuracy accounts for both of these standards. These accuracy ratings will be taken at the end of each session, hopefully showing an increase in the mean performance accuracy ratings. These results will be compared between the three groups to determine the effectiveness of each therapy method, hopefully showing that MIT does have the potential to be a credible therapy method.

**Clinical Implications**

The clinical implications of this study would be hopefully providing another therapy tool to use on the treatment of CAS. However, if the study shows that this treatment method is not valid for CAS, then it would add further research to this complex disorder. No matter the outcome of the investigation, the field of speech pathology will positively benefit from it. However, if this study was able to show the potential for MIT being a valid therapy method, eventually training courses could be offered to be able to specialize in this therapy technique. This is a practice that is common among speech therapists. Hopefully in the future, before training courses are offered, further studies could also be conducted to see how MIT affects the brain in cases of CAS since the majority of these studies focus on those with Aphasia.
References


