Speech and Music Therapy in the Treatment of Childhood Apraxia of Speech
van Tellingen, Mirjam; Hurkmans, Joost; Terband, Hayo; van de Zande, Anne Marie; Maassen, Ben; Jonkers, Roel
Published in:
Journal of speech, language, and hearing research
DOI:
10.1044/2023_JSLHR-22-00619

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Version created as part of publication process; publisher's layout; not normally made publicly available

Publication date:
2023

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
Clinical Focus

Speech and Music Therapy in the Treatment of Childhood Apraxia of Speech: An Introduction and a Case Study

Mirjam van Tellingen,a,b Joost Hurkmans,a Hayo Terband,c Anne Marie van de Zande,d Ben Maassen,b and Roel Jonkersb

aRehabilitation Centre “Revalidatie Friesland,” Beetsterzwaag, the Netherlands bUniversity of Groningen, the Netherlands cUniversity of Iowa, Iowa City dRehabilitation Centre “Rijndam Revalidatie,” Rotterdam, the Netherlands

ABSTRACT

Purpose: Speech–Music Therapy for Aphasia (SMTA), a method that combines speech therapy and music therapy, is introduced as a treatment method for childhood apraxia of speech (CAS). SMTA will be evaluated in a proof-of-principle study. The first case study is presented herein.

Method: SMTA was evaluated in a study with a single-subject experimental design comparing 10 weeks of treatment with 2 months of no treatment. The research protocol included a pretest, baseline phase, treatment phase, posttest, no-treatment phase, and follow-up test. The participant was a boy with CAS aged 5;8 (years;months). Outcome measures were selected to reflect both intelligibility in daily communication as well as features of CAS and speech motor planning and programming.

Results: Results on the Intelligibility in Context Scale–Dutch (ICS-Dutch) and in the analysis of a spontaneous speech sample suggest generalization of treatment effects. Improvements were found in measures that reflect complex speech motor skills, that is, the production of consonant clusters and consistency.

Conclusions: This case study showed that speech production of the participant improved after treatment with SMTA. Although intelligibility as measured with the ICS-Dutch improved over the study period, objectifying changes at the level of intelligibility in daily communication proved to be difficult. Additional measures may be necessary to gain more insight into treatment effects at this level. Overall, the results of this first case study provide sufficient support and important leads for further evaluation of SMTA in the treatment of CAS in a proof-of-principle study.

Childhood apraxia of speech (CAS) is a speech sound disorder classified as a subtype of motor speech disorder (Shriberg et al., 2010). In CAS, a core impairment at the level of speech motor planning and programming results in errors in the production of speech sounds and prosody. Inappropriate prosody, inconsistency, and disrupted coarticulation are widely described as three key features of CAS (e.g., American Speech-Language-Hearing Association [ASHA], 2007; Shriberg et al., 1997a, 1997b, 1997c; Terband et al., 2019). Inappropriate prosody, consonant and vowel errors, and voicing and nasality errors negatively impact intelligibility in CAS (Chenausky et al., 2022; Klopfenstein, 2009; McCabe et al., 2014), which negatively affects functional communication and social participation (Hustad, 2012).

Treatments for CAS address one or more of the three core features of CAS. In dynamic temporal and tactile cueing (DTTC; Strand et al., 2006; Strand, 2020), a
method for children with more severe CAS, all three features are targeted through a focus on movements rather than phonemes, varying prosody, and high numbers of repetitions. Rapid syllable transition treatment (ReST; Ballard et al., 2010), which is used in older children with less severe CAS, addresses all three features through the focus on sounds, beats, and smoothness. Studies report positive outcomes on ratings of production accuracy for DTTC (e.g., Maas et al., 2012; Maas & Farinella, 2012; Strand et al., 2006) and on segmental accuracy and lexical stress for ReST (Ballard et al., 2012; McCabe et al., 2014). While the results of these studies for different features of CAS are promising, McCabe et al. (2014) showed that the children improved on either lexical stress production or segmental accuracy, but they remained unable to simultaneously produce both correct stress patterns and correct segments. This is known to be a major challenge in developmental speech disorders (Howard, 2007).

Both DTTC and ReST can be described as articulatory–kinematic approaches, using interventions such as visual and tactile cues as well as feedback on the knowledge of performance. These methods also use some rate/rhythm control strategies, such as reduced speech rate and specific drill of lexical stress. Other methods, such as melodic intonation therapy (MIT; Albert et al., 1973; Helfrich-Miller, 1984) and Speech–Music Therapy for Aphasia (SMTA; De Bruijn et al., 2005; Hurkmans et al., 2018), are primarily described as rate/rhythm control–type approaches. The interventions in these methods are aimed at speech rate, stress, and intonation. MIT and SMTA were originally developed for adults with nonfluor aphasia and apraxia of speech (AoS; Hurkmans et al., 2015; Merrett et al., 2014). Both AoS and CAS are described as disorders in the planning and programming of speech movements (ASHA, 2007; Hurkmans, 2016) and share various characteristics, such as inconsistent errors in the realization of phonemes, syllable segregation, vowel distortions, groping, and effect of articulatory complexity (Iuzzini-Seigel & Murray, 2017; Ziegler, 2008). Therefore, rate/rhythm control approaches might be effective in the treatment of CAS.

There is a limited number of studies on the use of rate/rhythm control approaches and music or musical elements in the treatment of CAS (van Tellingen et al., 2022). Four of eight studies in the systematic review by van Tellingen et al. (2022) evaluated the use of MIT (Helfrich-Miller, 1994; Krauss & Galloway, 1982; Lagasse, 2012; Martikainen & Korpilahdi, 2011). The results of these studies vary and need to be interpreted with caution because the methodological quality of these studies was rated insufficient (van Tellingen et al., 2022).

In this study, SMTA is evaluated in the treatment of CAS in the first case of a series in a single-subject experimental design. The background and protocol for SMTA are introduced in the next section.

**SMTA**

SMTA is a combination of speech therapy and music therapy in which a speech-language pathologist (SLP) and a music therapist (MT) provide the treatment simultaneously. It is used in clinical practice with children from 3 years of age onwards with motor speech disorders, including (suspected) CAS. This method uses musical parameters that support the prosody of speech on word, phrase, and sentence levels and facilitate the sequencing and timing of speech movements. The musical compositions are tailored to individual needs, as the MT composes melodies to support the functionally relevant speech targets.

There are two lines of treatment, namely, a speech therapy line and a music therapy line, that are conducted simultaneously. Speech therapy includes three levels: (a) syllables, (b) words, and (c) sentences. These levels allow for a focus on movements, rather than individual speech sounds. As an exception, vowels may be practiced in isolation at the first level. Target items on the word and sentence levels are designed both to fit the speech targets based on speech assessment and to be functionally relevant (and therefore motivating) for the individual child in daily communication. For example, when a child has difficulty producing consonant clusters, and their brother’s name is “Steven,” this could make this name an excellent target item. Items may be both personal, such as names of family members, and more formulaic, such as “thank you.” Music therapy follows a structured procedure that starts with singing, followed by rhythmical chanting and speaking, which is derived from MIT. In SMTA, the final step of speaking is divided into smaller steps, including simultaneous speaking as well as alternating and semi-spontaneous speech. MIT uses rhythm and melody to simplify and exaggerate prosody, limiting melody in an alternation of a limited number of pitches (Sparks, 1981). The exercises in SMTA are designed to musically support natural speech, using the musical parameters melody, rhythm, meter, tempo, and dynamics. For each target item, a new melody is composed to support the prosodic features of the spoken utterance. This allows for selection of targets that are specifically tailored to the communicative needs and interests, speech sound inventory, and speech motor processes of the child. During an exercise, musical parameters may be used to adjust the exercise as needed. Variations in tempo, for example, may increase or decrease the difficulty of the exercise (De Bruijn et al., 2005; Hurkmans et al., 2018). SMTA has been shown to be an effective treatment method for Dutch adults with AoS and
aphasia in a proof-of-principle study with five speakers with aphasia and AoS (Hurkmans et al., 2015). In that study, intelligibility of verbal communication in daily life improved, as well as articulation.

SMTA is originally based on various similarities between language and music, such as shared hierarchical structures (Hurkmans, 2016; Patel, 2003; Peretz & Zatorre, 2003) and shared neural processing (e.g., Brown et al., 2006; see Hurkmans, 2016, for further discussion). SMTA has been shown to improve speech production at the level of motor planning and programming (Hurkmans et al., 2015). To provide the rationale for SMTA in the treatment of CAS, three theoretical frameworks on music and speech and the potential working mechanisms of music in the treatment of speech will be discussed: (a) similarities between and overlap in the processing of speech and music, (b) overlap in prosodic features in music and speech, and (c) mechanisms of music with regard to motivation and mood.

One of the original fundamental ideas for SMTA is the overlap in neural processing for language and music, which can be expanded to speech. Fujii and Wan (2014) showed that an overlap in neural processing of rhythm in music and speech, combined with synchronization and entrainment to a pulse, explains how rhythm supports the recovery of speech production. Overlap between music and speech can also be found in prosody, which is a characteristic for music and speech (e.g., Boutsen, 2003). Both include highly related features of sound such as melody and pitch, rhythm and duration, and dynamics and intensity (Hurkmans, 2016). Pitch, duration, and intensity are described as the features that combine to express stress in many languages (Terband et al., 2019). Through the overlap in features, music can be used to support speech prosody (Hurkmans et al., 2015). The third theoretical framework concerns the positive effects of music on mood and motivation in speech-language interventions, which have been summarized by Merrett et al. (2014) as one of the possible working mechanisms of MIT. In short, music is believed to have a positive effect on mood and motivation, which may contribute to the effect of interventions that utilize musical elements. Together, these frameworks provide insight into the potential mechanisms that contribute to the effects of SMTA.

SMTA encompasses all principles of motor learning that are recommended in the treatment of CAS (Maas et al., 2014; Murray et al., 2014; Strand, 2020). The use of music in SMTA allows for a high number of trials per session, which is crucial for motor learning (Maas et al., 2014; Strand, 2020), as singing (including a minimum of 20 trials per exercise) is regarded as more pleasant than realizing a high number of trials in a drill-type exercise. Usually up to five different exercises are conducted during a treatment session, alternating with small musical activities, such as singing a song or playing an instrument. Children are given autonomy within the session as they are invited to select the targets they want to practice and choose the musical activities they wish to engage in. This autonomy and the highly relevant target items can both contribute to increased motivation (Strand, 2020; Wulf et al., 2018). During an exercise, verbal feedback is kept to a minimum, to avoid disruptions of the flow of the exercise. Nonverbal knowledge of performance feedback is provided by focusing the child’s attention on the provided oral example in which the speech therapist may emphasize a specific movement. Feedback in the form of knowledge of results is also provided nonverbally and focuses on accurate realizations of the target word. Placing the child’s attention on accurate realizations raises their expectancies of their own ability. Before and at the end of an exercise, the functional use of the target is emphasized, to direct the attention to an external focus on results. This focus on feedback at the level of knowledge of results is recommended in the treatment of children with CAS to stimulate learning and retention (Strand, 2020) and has consistently been shown to enhance learning regardless of task, age, skill level, or (dis)ability (Wulf et al., 2018).

Treatment Protocol

SMTA is provided by trained SLPs and MTs. When a child is referred for SMTA, the SLP formulates target items with the child and/ or their parents or caregivers, depending on the age of the child. These target items are at the word and/or sentence level and are selected to be both functionally relevant for the child and relevant for the speech targets that are formulated based on speech assessment. Target items at the syllable level may be added for consonants or clusters that are still difficult for the child. Ideally, these targets at the syllable level will subsequently be used in targets at the word and/or sentence level.

The MT composes new melodies that support the natural melody, rhythm, and prosody of the target items. To this end, the MT uses melody, meter, rhythm, tempo, and dynamics to compose a melody that is close to the spoken prosody of the target item. This implies that musical features such as complex melodic structures, large intervals, and syncopation should be avoided (for Western languages). The musical parameters can also be used to influence the difficulty of the exercise, for example, for meter, a 6/8 beat elicits fluency more than a 4/4 beat. The composed melody consists of repetitions of the target item. The number of repetitions varies with the length of the target item, for example, four repetitions for a sentence or eight repetitions for a word. All melodies are new.
and specifically composed for the target item, as famous or previously used melodies will elicit the words and sentences that go with these melodies.

During a therapy session of 30 min, up to five target items will each be practiced in a fixed structure (see Tables 1 and 2). The child is usually seated opposite of the SLP, for visual assistance, for example, the oral example that the SLP provides. The MT is seated beside the child and SLP, creating a triangle-shaped setup. The introduction of the target item by the SLP includes naming the target clearly. The introduction may be supported by an object, photograph, or picture of the target item. This is followed by a demonstration of the target item by the MT. Directly thereafter, the child and the SLP join in with the MT and sing the melody, usually 2 times, but more repetitions can be added if deemed necessary. The next step is rhythmic chanting. In this phase, the melody fades out and the musical support is reduced to rhythmic assistance. The child and SLP can join in with the rhythmic support by tapping with the hand or foot, but this is not required. The choice to do this should be made based on what is helpful and not distracting for the child in their efforts to produce the target item. After rhythmic chanting, the rhythmic support is removed, and the SLP and the child simultaneously produce the target item repeatedly. Then, the SLP introduces turn-taking (direct imitation) with a hand gesture, signaling when the SLP will speak and when the child is invited to speak. Finally, the SLP poses a question that will elicit the target item. In this step, any visual support used in the introductory phase may be used to repeatedly elicit the target item. During the exercises and between the phases, verbal feedback is to be kept to a minimum. Feedback may be nonverbal, with facial expression or small gestures, but interruptions of the flow of the exercise should be avoided.

Changes may be necessary during the exercise, such as a change in tempo or an extra repetition of singing. Ideally, the SLP and MT develop a cooperation so that these changes can be made during the exercise by nonverbal cues to one another, without disrupting the flow of the exercise. Most children enjoy a break between exercises with singing a favorite song or playing an instrument for a short amount of time.

When a child produces a target item correctly at semispontaneous (elicited) speaking during therapy, recordings of the complete exercise can be made for practice at home. While practice at home with these recordings lacks the opportunity to provide interventions during the exercise, it does create opportunities for increasing treatment dose. Practicing at home might also contribute to the transfer of the target items to spontaneous speech in daily communication outside the treatment setting as realization of a target item in the pragmatically intended context provides a greater experience of success.

Clinical experiences with SMTA in the treatment of CAS are positive, but up to now, there were no efficacy studies of SMTA in the treatment of CAS. This study represents the first single-subject–design study into the effectiveness of SMTA in the treatment of CAS. The main research question was whether intelligibility in daily communication improves after treatment with SMTA. Secondary research questions focused on the effect of SMTA on the production of consonants, vowels, and clusters in spontaneous speech; picture naming; nonword imitation; and measures of speed, accuracy, consistency, and fluency in a diadochokinesis (DDK) task.

Case Study

Method

SMTA was evaluated in a study with a single-subject experimental design comparing 10 weeks of

<table>
<thead>
<tr>
<th>Phase</th>
<th>Activity</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>Starting song/short conversation</td>
<td>3 min</td>
</tr>
<tr>
<td>First exercise</td>
<td>Following protocol in Table 2</td>
<td>4 min</td>
</tr>
<tr>
<td>Second exercise</td>
<td>Following protocol in Table 2</td>
<td>4 min</td>
</tr>
<tr>
<td>Short break</td>
<td>Sing a song or play an instrument</td>
<td>2 min</td>
</tr>
<tr>
<td>Third exercise</td>
<td>Following protocol in Table 2</td>
<td>4 min</td>
</tr>
<tr>
<td>Fourth exercise</td>
<td>Following protocol in Table 2</td>
<td>4 min</td>
</tr>
<tr>
<td>Short break</td>
<td>Sing a song or play an instrument</td>
<td>2 min</td>
</tr>
<tr>
<td>Fifth exercise</td>
<td>Following protocol in Table 2</td>
<td>4 min</td>
</tr>
<tr>
<td>Conclusion</td>
<td>Make recordings/sing a closing song/other closing activity</td>
<td>3 min</td>
</tr>
</tbody>
</table>

Note. This structure may vary across children. The time set here per exercise allows for repetition of the entire exercise or practicing a part (word or syllable) of the target item separately before the complete target item.
Participants

Child + SLP + MT
Child + SLP
Child + SLP
4
Child + SLP + MT
4
8
Child + SLP
3 repetitions of target item

Number of repetitions

Table 2. Structure of an Speech–Music Therapy for Aphasia exercise.

<table>
<thead>
<tr>
<th>Step</th>
<th>Participants</th>
<th>Number of repetitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction of the target word/sentence</td>
<td>SLP (+ child)(^a)</td>
<td>1–2 repetitions of target item</td>
</tr>
<tr>
<td>Demonstrate target</td>
<td>MT</td>
<td>4–8 repetitions of target item(^b)</td>
</tr>
<tr>
<td>Singing</td>
<td>Child + SLP + MT</td>
<td>8–16 repetitions of target item(^b)</td>
</tr>
<tr>
<td>Rhythmic chanting</td>
<td>Child + SLP + MT</td>
<td>8–16 repetitions of target item(^b)</td>
</tr>
<tr>
<td>Simultaneously speaking</td>
<td>Child + SLP</td>
<td>4–8 repetitions of target item(^b)</td>
</tr>
<tr>
<td>Alternately speaking (direct imitation)</td>
<td>Child + SLP</td>
<td>5 repetitions of target item(^b)</td>
</tr>
<tr>
<td>Semi-spontaneous speaking (respond to question)</td>
<td>Child + SLP</td>
<td>3 repetitions of target item(^b)</td>
</tr>
</tbody>
</table>

Note. SLP = speech-language pathologist, MT = music therapist.

\(^a\)At the introduction the child is invited to participate when they are able to repeat the target item correctly.\(^b\)The number of repetitions is influenced by the length of the target item. A melody usually contains eight repetitions for words and four repetitions for sentences. The melody is sung and chanted at least two times. The numbers here are the minimal number of repetitions that are achieved during one exercise.

treatment with 2 months of no treatment focusing on speech production (at the level of phonological encoding as well as speech planning and programming). The research protocol included a pretest, baseline phase, treatment phase, posttest, no-treatment phase, and follow-up test. The protocol was approved by the research ethics committee at the University of Groningen (Ref. No. 77088008). Parents gave written informed consent before participation in the study. Treatment was provided at a rehabilitation center by an SLP and an MT who were both trained and experienced in providing SMTA. They followed the protocol for SMTA described in the introduction. Test administrations were conducted by another SLP, who was unaware of treatment progress. All tests at all timepoints were administrated by the same SLP, at the same location, using the same equipment for test administration and recording. Recordings of the test administrations were scored by the first author, who was blinded to the order of the recordings during scoring. After scoring was completed, results were matched to their date with the key provided by the SLP who conducted the test administrations.

Participant

The participant was a Dutch-speaking boy with CAS aged 5;8 (years;months). The diagnosis of CAS was confirmed following the protocol of Iuzzini-Seigel and Murray (2017), assessing several features of CAS in various speech tasks. The boy presented with inconsistent speech on a word and nonword repetition task and in spontaneous speech. Additional features included increasing problems with increasing complexity or length, which was shown in spontaneous speech, picture naming, nonword repetition, and DDK. Syllable segmentation, groping, and elongation of initial consonants were observed throughout the tasks. There were some consonant deletions and substitutions. In the phonological analysis of the child’s spontaneous speech, the consonants /l/ and /r/ were produced accurately less than 50% of occurrences in the initial position. In addition, the consonants /sl/, /dl/, and /sl/ (velar fricative) were produced correctly between 75% and 100% of the time in the initial position. For /dl/, there were several deletions in multisyllabic words. In monosyllabic contexts, /dl/ was realized accurately. In word-final position, the consonant /m/ was produced accurately in 33% of occurrences, and /l/, /kl/, /nl/, and /t/ were produced accurately in between 78% and 94% of occurrences. Other consonants were produced correctly in 100% of occurrences, in both initial and final positions. A full overview of Dutch phonemes and this boy’s phoneme acquisition is presented in Appendix A. The participant had voicing difficulties, leading to whispering partial and complete utterances across speech tasks. Intelligibility was negatively influenced by suprasegmental features, such as difficulties with voicing, dysfluency, low speech rate, and increasing difficulties with increasing length and/or complexity. Segmental errors impacted intelligibility to a lesser extent, with the absence of /h/ and /t/ in his speech being striking but also consistent. He showed awareness of his speech problems and a lack of self-confidence while speaking.

The medical history of the boy is described in detail in Appendix B. His medical history included gross motor difficulties, including delayed development of walking. His fine motor skills were age appropriate, and after physical therapy, his gross motor skills were age appropriate when he was at the age of 3;2. He had persistent colds and tonsil issues around the age of 18 months, resulting in the clipping of his tonsils and placement of tubes in his ears around his second birthday. His hearing was within normal limits when measured at the age of 20 months and again at the age of 3;7.

His speech and language were assessed multiple times from the age of 2 years to track progress and treatment effects. Word and sentence comprehension were within normal limits. Productive vocabulary showed a delay but was within normal limits from the age of 3 years. Sentence production was below normal limits. Treatment
was focused on increasing speech production, through targeting speech sounds and syllables. Nonverbal psychological assessment was conducted at the audiological center when the boy was at the age of 3;9, showing normal nonverbal psychological development.

The boy was placed in a specialized early education group, focusing on speech and language when he was at the age of 3;9. When he was at the age of 4;5, he went to school (which corresponds to preschool in the United States). He continued speech therapy in private practice. While there was progress in his phonological development, features of CAS became more apparent. Therefore, he was referred to the rehabilitation center for further assessment of suspected CAS and treatment at the age of 5;3.

**Intervention**

After the pretest, there was a 2-week baseline phase. In this period, five baseline measures (see below for a description of the measures) were taken. After the baseline phase, treatment started, consisting of two 30-min sessions of SMTA per week for 10 weeks, with additional homework using recordings of the targets that had been realized successfully during treatment. Ten target items on the sentence level were drawn up by the speech therapist and parents together. These items, which are presented in Appendix C, were both functionally relevant to the child as well as fitting with the outcomes of his speech assessment. Items were drawn up to target both the persistent segmental errors (/r/ and /h/) and the suprasegmental features, through choosing items at the sentence level, with multisyllabic words, and including numerous consonant clusters. Items were introduced over the treatment period as the SLP and MT saw fit.

Outcome Measures

A schematic overview of the outcome measures and timing of administration is presented in Table 3. The primary outcome measure was chosen to reflect intelligibility in daily communication, in line with the core objective of speech therapy, to support the child in communicative participation in society (Hustad, 2012). The selected measure was the Intelligibility in Context Scale–Dutch (ICS-Dutch; McLeod et al., 2012a). In this questionnaire, parents rate the intelligibility of their child in contact with various communication partners, such as family members, peers, teachers, and strangers, on a 5-point scale. Reliability and validity for this instrument were assessed for the original English version (McLeod et al., 2012b) and the Dutch version used in this study (McLeod, 2020; Van Doornik et al., 2018) and found to be adequate. Both parents filled out the ICS-Dutch independent of one another. Further outcome measures were selected to reflect speech motor abilities in various tasks, such as spontaneous speech, picture naming, nonword repetition, and DDK.

A speech sample was collected and analyzed using the Phonological Analyses for Dutch (Fonologische Analyse voor het Nederlands; Beers, 1995; Beers & Masereeuw, 2022). The sample was elicited through a series of standardized questions on topics such as school, hobbies, and vacation. The first 100 unique words in the sample were transcribed and analyzed. In this task, percent consonants correct in syllable-initial position (PCCI) and percent vowels correct (PVC) were calculated, as well as percentage of clusters correct in syllable-initial position (CCVC).

The Computer Articulation Instrument (CAI; Maassen et al., 2019) was used to assess speech in specific tasks. The CAI consists of the subtests Picture Naming, Nonword Imitation, Word- and Nonword Repetition, and DDK. For Picture Naming and Nonword Imitation, PCCI, PVC, and CCVC were calculated. In addition, the occurrence of cluster reduction was calculated. For Word- and Nonword Repetition, consistency was calculated, and for DDK, the measure was maximum repetition rate (MRR). The CAI is norm referenced and has been shown to have sufficient to good reliability and validity for the

### Table 3. Schematic representation of outcome measures administered per study phase.

<table>
<thead>
<tr>
<th>Pretest</th>
<th>Baseline (5 times)</th>
<th>Weekly testing</th>
<th>Posttest</th>
<th>Follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICS-Dutch</td>
<td>MDT</td>
<td>MDT</td>
<td>ICS-Dutch</td>
<td>ICS-Dutch</td>
</tr>
<tr>
<td>FAN</td>
<td>Figure Weights</td>
<td>Figure Weights</td>
<td>FAN</td>
<td>FAN</td>
</tr>
<tr>
<td>CAI</td>
<td></td>
<td></td>
<td>CAI</td>
<td>CAI</td>
</tr>
<tr>
<td>CAT</td>
<td>CAT</td>
<td>MDT</td>
<td>CAT</td>
<td>CAT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Figure Weights</td>
<td></td>
<td>MDT</td>
</tr>
</tbody>
</table>

Note. Figure Weights is from the Wechsler Intelligence Scale for Children–Fifth Edition, Dutch (Wechsler, 2018). Full descriptions of the measures are provided under “Outcome Measures” in the Method section. ICS-Dutch = Intelligibility in Context Scale–Dutch (McLeod et al., 2012a); MDT = Modified Diadochokinesis Test (Hurkmans et al., 2012); FAN = Phonological Analysis for Dutch (Beers, 1995); CAI = Computer Articulation Instrument (Maassen et al., 2019); CAT = Communication Attitude Test (Brutten & Vanryckeghem, 2003).
assessment of speech development in Dutch children ages 2–7 years (van Haaften et al., 2019). The CAI norms are divided into age groups that span 4 months for the younger groups and 6 months for the two oldest groups. For this study, raw scores for the pretest and posttest were compared with the norms for children aged 5:8–5:11, and raw scores for follow-up were compared to the norms for children aged 6:0–6:5 (van Haaften et al., 2019).

The Communication Attitude Test (CAT; Brutten, 1984; Brutten & Vanryckeghem, 2003) was used to assess the attitude of the child toward their own speech. In this norm-referenced instrument, children respond to statements about their speech with true or false. A higher score indicates a more negative attitude toward their own speech. The CAT was originally developed for children who stutter. The instrument’s reliability and validity have been studied in children who stutter and were sufficient (Vanryckeghem & Brutten, 1992).

The before-mentioned tasks were assessed at pretest, posttest, and follow-up. In addition, the Modified Diadochokinesia Test (MDT; Hurkmans et al., 2012) was used to assess speech motor planning and programming and establish a baseline and track treatment progress through weekly assessment. The MDT is a qualitative assessment of DDK. It consists of items of CV, CVC, CVCC, and CCVC structures, in which the vowel, or place or manner of the consonant, varies (see Appendix D for the items). The child is instructed to repeat each item 5 times, as accurately as possible. Responses are scored for accuracy, fluency, and consistency. Speech rate is not assessed. The reliability and validity of the MDT were assessed for use with adults with AoS and were sufficient for this group (Hurkmans et al., 2012). There are no records of the reliability and validity of the MDT for use with children with CAS.

A nonverbal control task was used to control for developmental progress. This was the task Figure Weights from the Wechsler Intelligence Scale for Children—Fifth Edition, Dutch (WISC-V-NL; Wechsler, 2018). The Figure Weights task was used in the baseline, weekly testing, and follow-up testing, alongside the MDT. Reliability and validity of the WISC-V-NL are adequate (Wechsler, 2018).

**Analysis**

Kendall’s tau test was used to test for change on the MDT measures CV, CVC, CCVC, CVCC, accuracy, consistency, fluency, place, manner, and vowel. The level of significance was set at $p < .05$. Kendall’s tau test was also used to analyze changes in the nonverbal control task Figure Weights. The CAI and CAT norms were used to describe changes on these measures.

**Results**

The participant in this study received twenty 30-min SMTA sessions, over a period of 11 weeks instead of 10, due to 1 week of illness during the treatment period. The number of items (both full items at the sentence level and parts of items at the word level) that was trained in each session varied from four to seven. The SLP and MT implemented repetitions of items and breaks as needed, resulting in different numbers of items practiced across sessions. In one session, there were four trained items; in 12 sessions, there were five; in five sessions, there were six trained items; and in two sessions, there were seven. Each week, the items trained in the first session were repeated in the second. The first week, the more personally motivating item with the name of his stuffed animal was trained, building from practicing words to the full sentence. In the second and third weeks, items with the persistently difficult /h/ and /θ/ were practiced at the word and sentence levels. In the fourth through sixth week, items with consonant clusters were trained at the syllable, word, and sentence levels. From Week 7 on, focus was placed on multisyllabic words, as well as consonant clusters, building up to sentence level for items containing such words. Nine out of 10 of the formulated targets at the sentence level were trained. The boy was able to produce the fourth item correctly upon first request during treatment. Therefore, this item was not further trained. For all trained items, it was necessary to practice words or syllables separately. This concerned words with consonants that were difficult for this boy, such as /θ/ and /l/. Several consonant clusters were trained at the syllable level. Musical interventions were used to support speech production, such as decreased tempo for multisyllabic words and upbeat for the production of clusters. Speech therapy interventions included verbal instructions and oral examples. For example, the boy produced an interdental [n] as a substitute for /ŋ/. Verbal instruction and oral examples to keep his mouth wide open for the syllable /ʃən/ supported the production of this syllable correctly. This was subsequently integrated into the word “belangrijk” (important); /ʃŋ/ was easier achieved in the item “springen” (jump) that was trained later on. In addition to instances of feedback at the level of knowledge of performance, feedback on knowledge of results was provided through nonverbal signs. These included signals to increase attention and effort, as well as reinforcements for adequate productions. Frequency of feedback was decreased with increasing adequate productions. Homework was provided, without fixed guidelines or expectations. Recordings of items that the boy could produce correctly during practice were sent to parents through a secure e-health application. Parents reported having had limited opportunities for practice at home. The participant received no other speech treatment
Results for intelligibility, as measured with the ICS-Dutch, are presented in Table 4. Combining the scores of both parents, the results suggested some improvement over the study period. Scores of the father and mother individually show different patterns.

The analysis of a sample of spontaneous speech with the Phonological Analysis for Dutch suggested improvement on PCCI over the treatment period, as presented in Table 5. Improvement was found for /sl/ and /sl/, which were realized to be accurate in 100% of occurrences after treatment, which was maintained at follow-up. The realization of /ld/ improved after treatment and improved further over the follow-up period. For the consonant /sl/, there was some improvement after treatment, which was not maintained at follow-up. Improvement on PVC and the production of initial clusters, which was obtained after treatment, was not maintained at follow-up.

Tasks and measures from the CAI showed varying results. The CAI does not provide critical differences; therefore, a change of \( z > +0.5 \) was set as the norm for clinically relevant change. Scores were calculated using the means and SDs published by van Haften et al. (2019). In picture naming, presented in Figure 1, there was a clinically relevant change for PCCI over the study period (\( z = +0.51 \) at follow-up compared to pretest) but no clinically relevant changes for PVC (\( z = +0.42 \) over the treatment period). The production of initial clusters in picture naming improved after treatment (\( z = +5.46 \)), and the gains were mostly maintained at follow-up (\( z = +4.29 \)). Results for cluster reductions in the same task, for which scores were inverted to reflect that a higher score means fewer cluster reductions, showed a clinically relevant change over the treatment period, reaching ceiling level after treatment (\( z = +0.97 \)), which was maintained at follow-up (\( z = +1.16 \)). On Nonword Imitation, presented in Figure 2, results for PCCI showed a clinically relevant change after treatment (\( z = +1.22 \)), but this was not maintained at follow-up (\( z = +0.16 \)). Results for PVC showed an increase from below average to within normal limits directly after treatment (\( z = +0.85 \)), but this improvement was not maintained at follow-up (\( z = -1.0 \)). The production of initial clusters in Nonword Imitation did not change directly after therapy but did improve at follow-up (\( z = +1.34 \)). Results for cluster reduction in this task showed a clinically relevant change, reaching ceiling level after treatment (\( z = +2.8 \)), which was maintained at follow-up (\( z = +2.83 \)). Results for consistency are presented in Figure 3. Word consistency increased directly after treatment (\( z = +0.80 \)), but this improvement was not maintained at follow-up. Nonword consistency did not increase directly after treatment, but improvement was apparent at follow-up (\( z = +1.71 \)). The DDK task from the CAI, which measures MRR (syllables per second), showed no change for the sequential items /pa/, /ta/, and /ka/. Results for the alternating items are presented in Figure 4. For the items /pata/ and /taka/, results were lower directly after treatment (\( z = -0.48 \) and \(-2.42 \)) and at follow-up (\( z = -0.16 \) and \(-0.69 \)). For /pataka/, the scores improved after treatment (\( z = +3.59 \)) from not being able to perform this sequence at pretest to within normal limits at follow-up (\( z = +4.21 \)).

![Picture Naming](figure1.png)

Table 4. Scores on the Intelligibility in Context Scale–Dutch (ICS-Dutch) as scored by the parents of the participant.

<table>
<thead>
<tr>
<th>ICS-Dutch</th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Father</td>
<td>3.7</td>
<td>3.4</td>
<td>3.9</td>
</tr>
<tr>
<td>Mother</td>
<td>3.6</td>
<td>4.0</td>
<td>4.1</td>
</tr>
<tr>
<td>Parents combined</td>
<td>3.6</td>
<td>3.7</td>
<td>4.0</td>
</tr>
</tbody>
</table>

Note. \( T1 = \) pretest; \( T2 = \) posttest; \( T3 = \) follow-up (after 2 months of no treatment).

Table 5. Mean scores of PCCI, PVC, and CCVC in a spontaneous speech sample.

<table>
<thead>
<tr>
<th></th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCCI</td>
<td>85</td>
<td>92</td>
<td>91</td>
</tr>
<tr>
<td>PVC</td>
<td>89</td>
<td>97</td>
<td>94</td>
</tr>
<tr>
<td>CCVC</td>
<td>40</td>
<td>64</td>
<td>52</td>
</tr>
</tbody>
</table>

Note. FAN = Phonological Analysis for Dutch; T1 = pretest; T2 = posttest; T3 = follow-up (after 2 months of no treatment); PCCI = percent consonants correct in initial position; PVC = percentage of vowels correct; CCVC = percentage of initial clusters correct.

Figure 1. Raw scores and \( z \) scores for measures of PCCI, PVC, CCVC, and Cred in picture naming. The star symbol (*) indicates \( z > +0.5 \), clinically relevant change. PCCI = percent consonants correct in initial position; PVC = percentage of vowels correct; CCVC = percentage of initial clusters correct; Cred = cluster reduction; T1 = pretest; T2 = posttest; T3 = follow-up (after 2 months of no treatment).
Results for communication attitude as measured with the CAT are presented in Table 6. There was no change in scores directly after treatment. At follow-up, scores were lower, showing a more positive attitude toward speech.

Results on several measures of the MDT showed change after treatment and over the study period. For syllable structure, there was a significant gradual improvement on CV (Kendall $\tau = 0.633$, $p < .01$) and CCVC (Kendall $\tau = 0.396$, $p < .05$) structures, as shown in Figure 5. Results for the structures CVC (Kendall $\tau = 0.290$, $p > .05$) and CVCC (Kendall $\tau = 0.132$, $p > .05$) showed no significant change over the study period. Significant gradual improvement was found for the measures of accuracy (Kendall $\tau = 0.433$, $p < .05$) and consistency (Kendall $\tau = 0.447$, $p < .05$), but not for fluency (Kendall $\tau = 0.211$, $p > .05$) as shown in Figure 6. Figure 7 shows significant gradual improvement for the measures place (Kendall $\tau = 0.513$, $p < .01$), manner (Kendall $\tau = 0.656$, $p < .01$), and vowel (Kendall $\tau = 0.356$, $p < .05$). For the scores on the nonverbal control task Figure Weights, which is also shown in Figure 7, there was no significant improvement over the study period (Kendall $\tau = -0.210$, $p > .05$).

Discussion

In this article, we introduced SMTA as a new method in the treatment of CAS. SMTA combines speech therapy and music therapy and is designed to support speech production at the level of motor planning and programming. Its potential for the treatment of CAS is supported by evidence of its effectiveness in adults with AoS (Hurkmans et al., 2015) as well as theoretical frameworks of the neural processing of rhythm in speech and music.

Table 6. Results for the Communication Attitude Test (CAT).

<table>
<thead>
<tr>
<th>CAT</th>
<th>T1 (Form A)</th>
<th>T2 (Form B)</th>
<th>T3 (Form B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw scores</td>
<td>7</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>$SD$</td>
<td>+0.25</td>
<td>+0.43</td>
<td>−0.52</td>
</tr>
</tbody>
</table>

Note. A higher score indicates a more negative attitude toward their own speech. T1 = pretest; T2 = posttest; T3 = follow-up (after 2 months of no treatment); $SD$ = standard deviation.
(Fujii & Wan, 2014), similarities between speech and music at the level of prosody (Hurkmans, 2016; Terband et al., 2019), and the positive effects of music on mood and motivation (Merrett et al., 2014). As a first investigation of this potential, we evaluated the effectiveness of SMTA in the treatment of CAS in a multiple-baseline single-subject design. First and foremost, this study showed that SMTA can be administered to 5- to 6-year-old children with CAS. Furthermore, the treatment yielded positive outcomes in speech production and intelligibility for the participant in this case study.

The study was designed to evaluate SMTA on a range of outcome measures, including intelligibility in daily communication and communication attitude, as well as a variety of measures of speech motor planning and programming in specific speech–motor tasks. Tasks included picture naming, nonword imitation, word and nonword repetition, and DDK, with outcome measures such as percentage of initial consonants correct, production of clusters, consistency, and fluency. Overall results showed progress on measures of speech motor planning and programming occurring directly after treatment.

(Fujii & Wan, 2014), similarities between speech and music at the level of prosody (Hurkmans, 2016; Terband et al., 2019), and the positive effects of music on mood and motivation (Merrett et al., 2014). As a first investigation of this potential, we evaluated the effectiveness of SMTA in the treatment of CAS in a multiple-baseline single-subject design. First and foremost, this study showed that SMTA can be administered to 5- to 6-year-old children with CAS. Furthermore, the treatment yielded positive outcomes in speech production and intelligibility for the participant in this case study.

The study was designed to evaluate SMTA on a range of outcome measures, including intelligibility in daily communication and communication attitude, as well as a variety of measures of speech motor planning and programming in specific speech–motor tasks. Tasks included picture naming, nonword imitation, word and nonword repetition, and DDK, with outcome measures such as percentage of initial consonants correct, production of clusters, consistency, and fluency. Overall results showed progress on measures of speech motor planning and programming occurring directly after treatment.
Improvement on intelligibility and communication attitude was mostly obtained over the entire study period and became apparent at follow-up.

As the goal of treatment in clinical practice lies in optimizing communication in daily life, we chose intelligibility in daily communication as the main outcome measure. In the scores of both parents on the ICS-Dutch combined, improvement was found at follow-up, but no effect was found directly after treatment. Examining the individual data of the two parents revealed that the father scored lower intelligibility directly after treatment. A possible explanation for this might be an increased awareness of his son’s speech problems through the pretest and treatment phases. Participation in the study and the parental questionnaires filled out at the beginning of the study might have drawn attention and raised awareness of the speech problems, leading to lower scores directly after treatment.

A similar trend was found in the scores of the CAT, with improvement in the boy’s attitudes toward his own speech apparent at follow-up, but not directly after treatment. This was observed in both the results on the CAT and clinical observations by the SLP who conducted the test administrations as well as by the first author who judged and scored the video recordings of spontaneous speech and test administrations. Clinical observations included changes in posture, communicative initiative, and speaking more freely. Two underlying mechanisms may have contributed to the pattern of improvement at follow-up rather than directly after treatment. First, during treatment, focus is put on speech and the difficulties with speech. This may be confronting and lead to a more negative attitude toward speech. While scores at pretest and posttest were the same, there were different responses on several items, showing an increased feeling of his speech being regarded “different” and feeling that “words don’t come out easy.” At follow-up, a growth in self-confidence became apparent in items like “speaking is easy for me” and “I won’t let others speak for me.” The follow-up period, in this case, was a summer break. The boy may have had fewer negative experiences with intelligibility, as he spent more time with family members and had no negative reactions at school or in other social situations. Second, changes in attitude toward own speech may take longer than improvement at the functional level of speech production, as new experiences with improved intelligibility may arise at the end of or even after treatment.

Intelligibility and attitude toward own speech are known to be impacted by inappropriate prosody, consonant and vowel errors, and voicing and nasality errors (Chenausky et al., 2022; Hustad, 2012; Klopfenstein, 2009; McCabe et al., 2014). This suggests that the improvement in intelligibility and attitude in this study may be caused by changes in this type of features related to speech motor planning and programming. In this study, there were positive changes in these features, which will be discussed next.

Secondary research questions concerned the effect of SMTA on speech motor planning and programming, such as the production of consonants, vowels, and clusters in spontaneous speech; picture naming; nonword imitation; and DDK measures of speed, accuracy, consistency, and...
 fluency. The analysis of the spontaneous speech sample showed progress in the production of consonants, vowels, and consonant clusters, suggesting generalization from trained items to spontaneous speech production. Post hoc analysis showed that consonants were roughly divided into three categories, one category of consonants that was pronounced accurately at pretest and remained that way, a small category of consonants that showed improvement, and another small category that was persistently difficult throughout the study period. For /l/ and /h/, it is shown in the treatment reports that the boy did produce /l/ and /h/ correctly during practice, but there was no generalization to spontaneous speech in the posttest. Potentially, a larger overall dose for these specific segments, supported with homework, would have resulted in generalization, like the generalization that occurred for cluster production.

Generalization from trained items directly after treatment was found in measures of the CAI. Changes in z scores on these measures show that the progress for this boy exceeded growth that would be expected with development. This suggests an effect of treatment that contributed to growth that allowed the boy to (partially) catch up with his peers.

Improvement was found on measures of the CAI that are related to specific features of CAS, that is, consistency and the production of clusters in picture naming. At pretest, the production of consonant clusters in the CCVC syllable structure was below normal limits. There were deletions, resulting in a low score on cluster reductions, just within normal limits. (Note that scores on cluster reduction were inverted, so that a higher score means less cluster reduction.) After treatment, scores on both the production of clusters and cluster reduction reached ceiling level. At follow-up, the ceiling level performance on cluster reduction was maintained, reflecting that the boy had now fully acquired this syllable structure. A small drop in scores for the production of clusters reflects that some substitutions were present in the realizations at follow-up.

Consistency and the production of clusters also changed in the production of nonwords, but this change occurred over the follow-up period, with no improvement directly after treatment. Differences in scores on tasks with words and nonwords might be influenced by auditory skills, which are necessary in the nonword imitation task, but not in picture naming. However, it is unclear why auditory skills would play a role at the posttest and not at follow-up. This would imply an improvement of auditory skills over the follow-up period, for which there are no further indicators. Another explanation might be that new skills need to be automated and therefore take longer to show up in test results. In that case, it would be expected that all scores for nonword imitation would progress at follow-up, but scores for PCCI and PVC improved directly after treatment. Additional analyses into features of the tasks and items might provide a more robust explanation for the different timing of improvement in consistency and production of clusters in picture naming versus nonword imitation.

On the DDK task, scores on MRR for both /pata/ and /taka/ sequences showed a clinically relevant decline. This may be influenced by the repeated administrations of the MDT during the treatment phase. The DDK tasks in the MDT and in the CAI measure DDK in a different way. In weekly testing during treatment, the child was instructed to produce syllables as accurate as possible in the MDT. Producing syllables at optimal speed, as required in the CAI at posttest, would then go against a trained habit. The production of /patakai/, which was impossible for the child at pretest, did change after treatment. After treatment, production was possible, but the score for speed was below normal limits. Similar results were found for /pata/ and /taka/. The habit of optimal accuracy was broken at follow-up, and speed increased again in DDK in the CAI for all alternating sequences.

In addition to the pretest, posttest, and follow-up, a baseline with subsequent weekly testing was conducted of the MDT and a control task. The analysis of different syllable structures and features of speech sounds on the MDT provide detailed insight into speech motor skills at the level of motor planning and programming. Most MDT measures showed a significant trend of improvement over the study period, including measures that relate to specific difficulties in CAS, such as consistency and the production of clusters. The improvement in DDK, a task that places a high demand on speech motor planning and programming, corresponds to the results of the study by Hurkmans et al. (2015).

In this study, it was shown that SMTA impacts speech production at the level of speech planning and programming. In a study by Chenausky et al. (2016), minimally verbal children with autism spectrum disorder, some of whom presumably also had CAS, improved in speech production after an intervention that used intoning and rhythm. They hypothesized about the effect of unison production and slowed production rate combined with intoning and tapping as facilitators for speech production. For SMTA, similar mechanisms could explain the effect on speech motor planning and production, but further research is needed to establish which components of SMTA may be considered the working mechanisms in the treatment of CAS.

The nonverbal, cognitive control task that was administered alongside the MDT showed no significant improvement over the study period, suggesting that
improvement on the MDT was not caused by developmental progress.

Overall, improvement was obtained on measures that reflect features of CAS and motor planning and programming. These are also measures that reflect more suprasegmental features (cf. co-articulation) of speech. Improvement on these suprasegmental features is in line with both the difficulties this boy experienced in his speech and the rate/rhythm approach of SMTA. The results in this case study may therefore provide a first indication of what the target group for SMTA might be, but further studies are needed to gain more insight into the effect of SMTA in the treatment of different stages and severities of CAS.

Limitations and Directions for Future Research

In this first case study of a series within a proof-of-principle study, there were some methodological limitations, especially in the choice of outcome measures. This single-subject design study represents a low level of evidence by itself. However, methodology was optimized, through comparison with a no-treatment period and additional control through multiple baseline measurements and a nonspeech control task. As the first single-subject design into SMTA, this study provides sufficient encouragement for follow-up studies.

The measurements in this study were chosen to reflect intelligibility in daily communication as well as speech motor planning and programming. Objectifying change in intelligibility in daily communication proved to be difficult. In addition to concerns about the scoring of the ICS-Dutch by the father in this specific case, this instrument is vulnerable to bias, as parents are aware of the timing of treatment. In addition, this instrument provides insight into intelligibility in daily communication but does not measure whether communicative participation has changed. Changes in communicative participation could be expected when both intelligibility and attitudes toward speech improve, but the current measures did not provide insight into participation. An additional measure, such as Focus on the Outcomes of Communication Under Six (Thomas-Stonell et al., 2013), might be useful in future studies to get a broader insight into changes in communicative participation after treatment.

Results on the CAI show that, for some measures, a small change in the raw score may result in a far greater change in $z$ score. This is caused by ceiling effects on these measures. Typically developing children at the age of the participant score close to 100% on most measures, reflecting nearly completed motor–speech development (although refinement of these skills continues for longer; Ballard et al., 2012). Therefore, one error affecting the raw score may cause a large drop in the $z$ score for such a measure (or one more item correct may cause a large increase in the $z$ score).

The results on the MDT showed that maintenance of the treatment gains mostly occurred for measures that reached proficiency levels of 75% or higher during the treatment phase, such as the correct realization of syllable structures CV and CVC, as well as place and manner of articulation, and the consistency of productions. This suggests that a child should demonstrate a minimum increase in performance during treatment (in this case, a 75% or higher proficiency level) to expect learning (the ability to apply a skill without support; Olswang & Bain, 1994), as reflected by our assessment of maintenance. In addition, the results on the MDT represent generalization of treatment effects to untreated items. Maintenance of generalized treatment gains holds great potential for improved intelligibility in daily communication. In this study, the dose of treatment may have been too low to obtain a 75% proficiency level on all measured features. An extended treatment period or increased intensity may be necessary to obtain the proficiency levels required for broader maintenance of improvement.

Measurements were also chosen to objectify speech motor planning and programming. This was specifically measured using the MDT, including consistency, fluency, and accuracy. For most measures of the MDT, there was a strong dip in the scores toward the end of the treatment period. At this time, the boy turned out to be sick. While results on following test administrations did not completely recover, overall, improvement was still apparent.

Measures of speech planning and programming also provide insight into changes in the core features of CAS. Consistency was assessed specifically not only via the repetition of words and nonwords but also in the repetition of syllables on the MDT. Co-articulatory transitions are assessed via the production of clusters. Prosody was to be assessed through the measure of fluency in the MDT. However, this measure was insufficient in this case to objectify prosody. In the definition of CAS as stated by ASHA (2007), the realization of lexical and phrasal stress is named as a marker for inappropriate prosody. As the MDT uses nonword syllables, lexical stress is not assessed. In addition, in this case, scores for fluency approached ceiling levels very early on, as the boy adopted a strategy where he would produce segmentally simplified, inaccurate sequences in a fluent manner. The structure of the task and the strategy adopted by this boy both contribute to a positive but incorrect reflection of his prosodic skills in the task results.
The assessment of prosody in the evaluation of SMTA in the treatment of CAS is relevant and should be expanded in future studies. First, assessing all three core features of CAS will provide insight into the effect of treatments on these features. This could lead to better choices in treatment planning, choosing a method that is best for a child at a given point in time. Second, SMTA might be especially effective in the treatment of prosody through the support of music. Speech and music are highly related with respect to prosodic features such as frequency (pitch), duration (rhythm), and intensity (dynamics; Boutsen, 2003; Hurkmans, 2016; Terband et al., 2019). Because of the similarities between music and speech in prosody and the focus of SMTA on the level of speech planning and programming, it is hypothesized that prosody in children with CAS will improve after treatment with SMTA. To evaluate this potential effect, specific measures of prosody at the level of lexical and phrasal stress are needed. Such tasks will be developed to be included in outcome measurements in future studies.

Conclusions

SMTA, a treatment method that combines speech therapy and music therapy, was introduced, and its efficacy in the treatment of CAS is now being evaluated in a single-subject–design study. This study shows that speech production in the participant improved after treatment, specifically on tasks that relate to CAS and motor speech planning and programming such as consistency and the production of clusters. While intelligibility improved over the study period, objectifying changes in daily communication proved to be difficult. Additional measures may be necessary to gain more insight into treatment effects at the level of communicative participation. Overall, the results of this first single-subject–design study provide sufficient support and important encouragement for further evaluation of SMTA in the treatment of CAS in a proof-of-principle study.

Author Contributions

Mirjam van Tellingen: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Visualization, Writing – original draft. Joost Hurkmans: Conceptualization, Methodology, Supervision, Writing – review & editing. Hayo Terband: Conceptualization, Formal analysis, Methodology, Visualization, Supervision, Writing – review & editing. Anne Marie van de Zande: Conceptualization, Investigation, Project administration, Writing – review & editing. Ben Maassen: Conceptualization, Methodology, Supervision, Writing – review & editing. Roel Jonkers: Conceptualization, Methodology, Supervision, Writing – review & editing.

Data Availability Statement

The data can be obtained by contacting the first author.

Acknowledgments

This study was performed without external funding. We thank the participant and his parents for their contribution. We thank the speech therapists and music therapists at the rehabilitation centers “Revalidatie Friesland” and “Rijndam Revalidatie” for their work in providing treatment and collecting data.

References


Hurkmans, J. J. S. (2016). The treatment of apraxia of speech: Speech and music therapy, an innovative joint effort [Thesis fully internal (DIV), University of Groningen].


Appendix A

Phonetic Repertoire

<table>
<thead>
<tr>
<th>Phoneme (IPA)</th>
<th>Mean age of acquisition (y;m)</th>
<th>PICC in spontaneous speech</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>1;3–1;8</td>
<td>100</td>
</tr>
<tr>
<td>t</td>
<td>1;3–1;8</td>
<td>100</td>
</tr>
<tr>
<td>m</td>
<td>1;3–1;8</td>
<td>100</td>
</tr>
<tr>
<td>n</td>
<td>1;3–1;8</td>
<td>100</td>
</tr>
<tr>
<td>j</td>
<td>1;3–1;8</td>
<td>100</td>
</tr>
<tr>
<td>k</td>
<td>1;9–1;11</td>
<td>100</td>
</tr>
<tr>
<td>s</td>
<td>2;0–2;2</td>
<td>95</td>
</tr>
<tr>
<td>x</td>
<td>2;0–2;2</td>
<td>96</td>
</tr>
<tr>
<td>h</td>
<td>2;0–2;2</td>
<td>50</td>
</tr>
<tr>
<td>b</td>
<td>2;3–2;5</td>
<td>100</td>
</tr>
<tr>
<td>f</td>
<td>2;3–2;5</td>
<td>n.a.</td>
</tr>
<tr>
<td>v</td>
<td>2;3–2;5</td>
<td>100</td>
</tr>
<tr>
<td>l</td>
<td>2;6–2;8</td>
<td>100</td>
</tr>
<tr>
<td>r</td>
<td>2;6–2;8</td>
<td>0</td>
</tr>
<tr>
<td>d</td>
<td>2;9–2;11</td>
<td>81</td>
</tr>
<tr>
<td>z\a</td>
<td>Unknown</td>
<td>n.a.</td>
</tr>
<tr>
<td>v\a</td>
<td>Unknown</td>
<td>100</td>
</tr>
<tr>
<td>g</td>
<td>Unknown</td>
<td>100</td>
</tr>
</tbody>
</table>

Note. Data were from the pretest. Mean age of acquisition was according to Beers (1995). IPA = International Phonetic Alphabet; y;m = years;months; PICC = percentage of initial consonants correct; n.a. = not applicable (meaning, there were no occurrences of this phoneme in this sample).

\aVoicing in /z/ and /v/ is not meaning-distinguishing in Dutch; therefore, they are increasingly realized as /s/ and /f/, which is now considered correct when analyzing a speech sample.

Appendix B

Medical History of the Participant

The medical history of the boy included a delayed development of walking due to tiptoeing, which was treated with splints between the age of 15 and 20 months. At the age of 20 months, his gross motor skills were below age level, but his fine motor skills were age appropriate and of good quality. He received physical therapy to support the development of his gross motor skills. At the age of 3;2, his gross motor skills were age appropriate.

He had persistent colds and tonsil issues around the age of 18 months, resulting in the clipping of his tonsils and placement of tubes in his ears around his second birthday. His hearing was within normal limits when measured at the age of 20 months and again at the age of 3;7.

His speech and language were assessed multiple times from the age of 2 years to track progress and treatment effect. Word and sentence comprehension were within normal limits. Sentence production and vocabulary (productive) were difficult to assess, due to limited and unintelligible speech at the age of 2;8. From the age of 3 years, productive vocabulary was well within normal limits, while sentence production was below normal limits. The boy started speech therapy around the age of 2;4. Treatment was focused on increasing speech production, through targeting speech sounds and syllables. Additionally, sentence production was targeted, while it was noted that insufficient sentence production may be caused by the speech production difficulties.

Treatment led to insufficient developmental progress in speech production. Therefore, his speech and language development were assessed at an audiological center when he was 3;7. He was diagnosed with a delayed syntactic development due to the omission of function words in his sentences. Additionally, he was diagnosed with a serious phonological delay, due to substitutions and omissions of several speech sounds. Nonverbal psychological assessment was conducted at the audiological center when the boy was 3;9, showing a normal nonverbal psychological development.

After these assessments, the boy was placed at a specialized early education group, focusing on speech and language. From the age of 4;0, the boy attended both regular school (which corresponds to preschool in the United States) and the specialized early education group. When he was 4;5, he went to preschool full-time. He continued speech therapy in private practice. While there was progress in his phonological development, as processes like devoicing and lateralization decreased, features of childhood apraxia of speech became more apparent. Therefore, he was referred to the rehabilitation center for further assessment of suspected CAS and treatment at the age of 5;3.
Appendix C
Trained Items

1. Rudo gaat straks naar groep vijf en ik ga naar groep drie
Rudo will attend fifth grade later and I will go to third grade

2. Ik hou van knutselen, kleien en strijkkralen
I like to do arts and crafts, play-doh and ironing beads

3. Mijn liefste knuffel heet Buk de Beer
My dearest stuffed animal is named Buk the bear

4. Mijn kleine nichtje heet June
My little cousin is called June

5. Mijn juffen heeten juf Jeltsje en juf Greet
My teachers names are miss Jeltsje and miss Greet

6. Ik ga op vakantie naar camping het kuierpad in drenthe
I’m going on a vacation to campsite the kuierpad in Drenthe

7. Ik leer skaten en ik kan al schaatsen
I’m learning to (roller)skate and I can ice skate already

8. Het is belangrijk om je gordel om te doen
It is important to fasten your seatbelt

9. Het is gevaarlijk om zonder helm te fietsen
It is dangerous to cycle without a helmet

10. Ik kan hoog springen op de trampoline
I can jump high on the trampoline
## Appendix D

Items and Scoring of the Modified Diadochokinesis Test

<table>
<thead>
<tr>
<th>Item</th>
<th>Consistency</th>
<th>Accuracy</th>
<th>Fluency</th>
<th>Total structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>pa pa pa</td>
<td>0–4</td>
<td>0–15</td>
<td>0–5</td>
<td></td>
</tr>
<tr>
<td>pa ta ka (p)</td>
<td>0–4</td>
<td>0–15</td>
<td>0–5</td>
<td></td>
</tr>
<tr>
<td>da na la (m)</td>
<td>0–4</td>
<td>0–15</td>
<td>0–5</td>
<td></td>
</tr>
<tr>
<td>pa po pu (v)</td>
<td>0–4</td>
<td>0–15</td>
<td>0–5</td>
<td>(CV, 96)</td>
</tr>
<tr>
<td>paf paf paf</td>
<td>0–4</td>
<td>0–15</td>
<td>0–5</td>
<td></td>
</tr>
<tr>
<td>paf taf kaf (p)</td>
<td>0–4</td>
<td>0–15</td>
<td>0–5</td>
<td></td>
</tr>
<tr>
<td>daf naf laf (m)</td>
<td>0–4</td>
<td>0–15</td>
<td>0–5</td>
<td></td>
</tr>
<tr>
<td>paf pof puf (v)</td>
<td>0–4</td>
<td>0–15</td>
<td>0–5</td>
<td>(CVC, 96)</td>
</tr>
<tr>
<td>paks paks paks</td>
<td>0–4</td>
<td>0–15</td>
<td>0–5</td>
<td></td>
</tr>
<tr>
<td>paps pats paks (p)</td>
<td>0–4</td>
<td>0–15</td>
<td>0–5</td>
<td></td>
</tr>
<tr>
<td>pats pans pals (m)</td>
<td>0–4</td>
<td>0–15</td>
<td>0–5</td>
<td></td>
</tr>
<tr>
<td>paks poks puks (v)</td>
<td>0–4</td>
<td>0–15</td>
<td>0–5</td>
<td>(CVCC, 96)</td>
</tr>
<tr>
<td>spag spag spag</td>
<td>0–4</td>
<td>0–15</td>
<td>0–5</td>
<td></td>
</tr>
<tr>
<td>spag stag skag (p)</td>
<td>0–4</td>
<td>0–15</td>
<td>0–5</td>
<td></td>
</tr>
<tr>
<td>stag snag slag (m)</td>
<td>0–4</td>
<td>0–15</td>
<td>0–5</td>
<td></td>
</tr>
<tr>
<td>spag spog spug (v)</td>
<td>0–4</td>
<td>0–15</td>
<td>0–5</td>
<td>(CCVC, 96)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>(64)</td>
<td>(240)</td>
<td>(80)</td>
<td></td>
</tr>
<tr>
<td><strong>Total p/m/v</strong></td>
<td>(p, 96)</td>
<td>(m, 96)</td>
<td>(v, 96)</td>
<td></td>
</tr>
</tbody>
</table>