Treatment intensity and childhood apraxia of speech

Article *in* International Journal of Language & Communication Disorders · January 2015
DOI: 10.1111/1460-6984.12154

Aravind Kumar Namasivayam
University of Toronto

Jennifer Hard
ErinoakKids Centre for Treatment and Development

Frank Rudzicz
University of Toronto; Vector Institute; St Michael's Hospital; Surgical Safety Tech…

Toni Rietveld
Radboud University

International Journal of Language & Communication Disorders

INT J LANG COMMUN DISORD, JULY 2015,
VOL. 50, NO. 4, 529–546

Research Report

Treatment intensity and childhood apraxia of speech

Aravind K. Namasivayam†, Margit Pukonen‡, Debra Goshulak‡, Jennifer Hard§, Frank Rudzicz#, Toni Rietveld||, Ben Maassen¶, Robert Kroll‡ and Pascal van Lieshout†

†Department of Speech–Language Pathology, University of Toronto, Toronto, ON, Canada
‡The Speech and Stuttering Institute, Toronto, ON, Canada
§ErinoakKids Centre for Treatment and Development, Mississauga, ON, Canada
¶University of Groningen, Groningen, the Netherlands
||Radboud University, Nijmegen, the Netherlands
#Toronto Rehabilitation Institute, Toronto, ON, Canada

(*Received December 2013; accepted November 2014*)

Abstract

Background: Intensive treatment has been repeatedly recommended for the treatment of speech deficits in childhood apraxia of speech (CAS). However, differences in treatment outcomes as a function of treatment intensity have not been systematically studied in this population.
Aim: To investigate the effects of treatment intensity on outcome measures related to articulation, functional communication and speech intelligibility for children with CAS undergoing individual motor speech intervention.
Methods & Procedures: A total of 37 children (32–54 months of age) with CAS received 1×/week (lower intensity) or 2×/week (higher intensity) individual motor speech treatment for 10 weeks. Assessments were carried out before and after a 10-week treatment block to study the effects of variations in treatment intensity on the outcome measures.
Outcomes & Results: The results indicated that only higher intensity treatment (2×/week) led to significantly better outcomes for articulation and functional communication compared with 1×/week (lower intensity) intervention. Further, neither lower nor higher intensity treatment yielded a significant change for speech intelligibility at the word or sentence level. In general, effect sizes for the higher intensity treatment groups were larger for most variables compared with the lower intensity treatment group.
Conclusions & Implications: Overall, the results of the current study may allow for modification of service delivery and facilitate the development of an evidence-based care pathway for children with CAS.

Keywords: childhood apraxia of speech, treatment intensity, developmental motor speech disorders, speech intelligibility, speech–sound disorder, functional outcomes.

What this paper adds?
What is already known on the subject?
More intensive treatment has been recommended for the treatment of speech deficits in CAS. However, differences in treatment outcomes related to articulation, functional communication and speech intelligibility as a function of treatment intensity have not been systematically studied in this population.

Address correspondence to: Aravind Namasivayam, Department of Speech–Language Pathology, 160-500 University Avenue, Rehab Sciences Bldg, Toronto, ON M5G1V7, Canada; e-mail: a.namasivayam@utoronto.ca

International Journal of Language & Communication Disorders
ISSN 1368-2822 print/ISSN 1460-6984 online © 2015 Royal College of Speech and Language Therapists
DOI: 10.1111/1460-6984.12154

What this paper adds?
Findings from this study indicate that only higher intensity treatment (2×/week) yields significant outcomes and larger effect sizes for articulation and functional communication compared with 1×/week (lower intensity) intervention. Further, neither lower nor higher intensity treatment yielded a significant change for speech intelligibility at the word or sentence level. These findings may allow for modification of service delivery and facilitate the development of an evidence-based care pathway for children with CAS.

Introduction

Clinical management of childhood apraxia of speech (CAS) is challenging at many levels. These challenges extend from the aetiology and definition of CAS to the type, intensity, frequency and amount of treatment required to produce satisfactory outcomes. The aetiology of CAS is unknown, but it has been suggested to be a neurological sensorimotor based speech–sound disorder (SSD) subtype with a disruption of neurophysiological processes at the level of speech motor planning and/or motor programming of speech movement sequences (American Speech–Language–Hearing Association (ASHA) 2007, Caruso and Strand 1999). Since there are no definitive diagnostic markers that differentiate CAS from other SSDs (ASHA 2007), researchers and clinicians use a consensus approach to CAS diagnosis based on the presence of several behavioural characteristics typical of CAS. These characteristics are described in a position paper (ASHA 2007) and include inconsistent speech errors on repeated productions, lengthened and disrupted coarticulatory transitions and inappropriate prosody (ASHA 2007, Strand *et al.* 2006).

Based on the premise that the underlying issue(s) in CAS may be limitations in speech planning and motor programming, various researchers have recommended the use of motor learning principles (e.g., production frequency, conditions of practice, type and frequency of feedback etc.) in the treatment of CAS to address the disorder at its presumed origin (Edeal and Gildersleeve-Neumann 2011, Maas and Farinella 2012, Maas *et al.* 2012). As pointed out by Maas *et al.* (2012), although there are several studies in the speech and non-speech literature carried out in the adult population to support the use of the principles of motor learning (PML), there is limited empirical data that have verified the value of PML in CAS treatment. To date, only PML variables related to the amount of practice (production frequency; Edeal and Gildersleeve-Neumann 2011), practice schedule (random versus blocked practice; Maas and Farinella 2012), and feedback frequency (high-frequency versus low-frequency feedback; Maas *et al.* 2012) have been explicitly tested in the CAS population via single-subject experimental designs.

Edeal and Gildersleeve-Neumann (2011) demonstrated that the amount of practice is positively correlated with treatment outcomes in children with CAS, wherein greater amounts of practice (approximately 150 trials per session) improved speech–sound accuracy, retention and transfer relative to lesser amounts of practice (30–40 trials per session). The results of other studies testing PML variables in the CAS population have been mixed. For example, in Maas and Farinella (2012), utterance accuracy improved in two of four participants under the blocked practice condition relative to the random practice condition. In contrast, for another participant, there was a consistent advantage for the random over the blocked practice condition. Similarly, in Maas *et al.* (2012), a reduction in the frequency of augmented clinician-provided feedback (i.e., feedback provided in approximately 60% of all trials; low-frequency condition) enhanced learning in only two of four participants with CAS. One participant demonstrated an advantage for 100% feedback (high-frequency feedback condition) and one participant showed no clear preference for either feedback condition. These findings suggest that not all PML manipulations can be extrapolated from the motor learning literature based on adult studies and applied in the treatment of SSDs in the CAS population (Maas *et al.* 2012).

Current study

Another PML variable that is often cited as a critical factor for motor learning is treatment intensity (Edeal and Gildersleeve-Neumann 2011). ASHA (2007) indicates that there is preliminary evidence to support the provisioning of more intense treatment (i.e., three to five individual sessions/week) for CAS as opposed to less intense (1–2×/week) service delivery formats (Hall *et al.* 1993, Skinder-Meredith 2001, Strand and Skinder 1999). Although intensive treatment has been repeatedly recommended for the treatment of speech deficits (Maas *et al.* 2008, Strand *et al.* 2006), differences in treatment outcomes as a function of treatment intensity have not been systematically studied in the CAS population. Treatment intensity has a direct impact on service delivery and, as a consequence, has serious logistical and economic ramifications for caregivers, insurance companies and policy-makers. Typically, three to five individual sessions/week have been referred to as high-intensity treatment in the literature (ASHA 2007, Hall *et al.* 1993, Skinder-Meredith 2001, Strand and Skinder 1999), but current service delivery models in Ontario, Canada, allow only for one or two individual treatment sessions per week for children with SSDs. The current study works within the existing service delivery model in Ontario to assess speech and functional outcomes in a group of children with CAS subsequent to 1×/week (lower intensity) or 2×/week (higher intensity) individual motor speech treatment. In doing so, it aims to contribute to the limited corpus of treatment efficacy literature in this hard-to-treat speech disorder (ASHA 2007).

Research questions

- What are the effects of treatment intensity on outcome measures (related to speech production, speech intelligibility and functional communication) for children with CAS undergoing individual motor speech intervention?
- What is the magnitude of change as a function of treatment intensity across outcome measures for children with CAS undergoing individual motor speech intervention?

Methods

Participants

Children who presented with CAS features were extracted from a larger data set of 85 pre-school age children with SSDs who participated in a Ministry of Children and Youth Services (Province of Ontario)-funded research study. All 85 children had English as the primary language spoken at home and met several inclusion and exclusion criteria to participate in the larger research study. Inclusion criteria included prerequisite skills (social, play and attention skills) for direct speech intervention (clinical observation by a certified Speech-Language Pathologist (SLP) experienced in working with this population), hearing and vision within normal limits (parent reports), mild or greater delays in expressive language with age-appropriate or near age-appropriate receptive language, moderate to profound SSD severity and presence of indicators for motor speech involvement (e.g. vowel and consonant distortions, lateral jaw sliding, inappropriate jaw excursion, decreased lip rounding and retraction). Participants were excluded if they showed any signs and symptoms suggesting (1) global motor involvement (e.g. cerebral palsy), (2) presence of autism spectrum disorders, (3) presence of oral structural deficits (tongue, lips, palate), (4) presence of feeding impairments, or (5) presence of significant drooling. These exclusionary assessments were based on clinical referral forms, parental reports and clinical observation by an experienced and certified SLP.

To identify participants with CAS, especially in the absence of clear-cut diagnostic markers or the availability of standardized assessment procedures for the diagnosis of CAS, clinical expert opinion remains the current accepted procedure (Maas *et al.* 2012, McCauley and Strand 2008, Shriberg *et al.* 1997, 2011). For the present study, an independent SLP (not providing assessment or treatment; SLP 1) with expertise in developmental motor speech disorders including CAS, screened all participants using a behavioural checklist specific to CAS that focused on three key areas: general motor control, segmental and supra-segmental features (see appendix A).

The CAS checklist used in the present study is a modified version of the behavioural criteria for CAS diagnosis as reported by Shriberg *et al.* (2011: 494). Additional items such as inconsistent production, limited variety of speech movements and preference for well-rehearsed sound sequences/words were included based on clinical features reported in the 2007 ASHA Technical Report on CAS (ASHA 2007), a careful review of the literature (Crosbie *et al.* 2005, Maassen *et al.* 2010, Maassen 2002, Ozanne 2005, Strand *et al.* 2006), a Ministry of Child and Youth Services Motor Speech Working Group document (February 2009), and through collaborative consultations with several researchers and clinicians with expertise on CAS. The original checklist reported by Shriberg *et al.* (2011) had a 40% (four of 10 features) cut-off for CAS categorization. Since the current check list contained additional features with some features weighted more strongly (e.g. groping is counted twice: (1) as a feature characterizing difficulty achieving initial articulatory configurations and (2) as the result of increasing length and complexity of utterance) it was decided to raise the cut-off point to a more conservative value of 58% for a positive CAS classification.

That is, for a positive CAS classification, at least seven of 12 (58%) features listed in appendix A must be present in three or more of the five assessments tasks. SLP 1 completed the CAS checklist based on video/audio recordings containing syllable repetition and sequencing tasks (Kaufman Speech Praxis Test; Kaufman 1995), word-level spontaneous picture labelling task (Goldman–Fristoe Test of Articulation, 2nd Edition—GFTA-2; Goldman and Fristoe 2000), word and phrase-level repetition tasks (Children's Speech Intelligibility Measure—CSIM; Wilcox and Morris 1999; and Beginner's Intelligibility Test—BIT; Osberger *et al.* 1994), and perceptually from a spontaneous speech sample, where available.

Using these procedures, 37 children (females = 9; males = 28; age range 32–54 months; mean = 40.22 months, SD = 5.60 months) were identified as having CAS from a larger data set of 85 pre-school age children who presented with moderate to profound SSDs with motor speech difficulties. The remaining 48 children did not meet specific criteria for CAS (i.e. scored < 58% on the checklist; see appendix A). The percentage of participants with CAS (37/85 = 43.5%) is much higher than that reported in the literature for children with general speech delays (3.4–4.3%; ASHA 2007). This difference can be attributed to the fact that we were working with a subpopulation of children with speech delays who also demonstrated motor speech involvement.

To establish robustness of the CAS classification across raters and clinical methods, a second SLP (SLP 2) specializing in the assessment and treatment of motor speech disorders in children reviewed videos of a therapy session (usually mid-point in treatment as a part of treatment fidelity assessment) and monitored for characteristics of CAS, as listed on the checklist (see appendix A). This clinician was blind to the child's assessment and diagnostic information, including the CAS classification by SLP 1. Cohen's Kappa was 0.669 as calculated between SLPs 1 and 2 based on 30% of the data set (26/85 children) that was randomly selected using the 'RANDBETWEEN' function in MS Excel. Cohen's Kappa scores below 0.40, 0.40–0.75 and over 0.75 are considered poor, fair to good, and excellent respectively (Fleiss 1981).

Approval from the Human Ethics Review Committee at the University of Toronto was obtained for the study. In addition, individual participating sites provided local ethics approval as required.

Experimental design

Pre-/post-design

The present study used a pre- and post-treatment design (pre-/post). This design was not intended to demonstrate causality but to study the effects of variations in treatment intensity on outcome measures (Dollaghan 2007). There were two paradigms: a lower intensity paradigm (45-min sessions, one session a week for 10 weeks = 10 sessions) and a higher intensity paradigm (45-min sessions, two sessions a week for 10 weeks = 20 sessions). Both paradigms had the same treatment duration of 10 weeks (this provides a control for time/maturation effects), but differed in the intensity of treatment. For this design, assessments were carried out before and after a 10-week treatment block. The pre- and post-treatment assessments were generally administered over two or three sessions scheduled within a span of 1–2 weeks. Thus, for example, a 10-week treatment block was preceded and succeeded by 1–2 weeks of assessments, totalling a 12–14-week study period.

Clinician training

All clinicians providing intervention were qualified SLPs who prior to the start of this study completed a survey regarding previous experience and training in working with children with motor speech disorders. Following this, the SLPs underwent two multi-day workshops on the assessment and treatment of children with SSD and motor speech issues, including CAS. The first two-day workshop focused on the assessment and treatment of children with motor speech disorders. Topics covered in this workshop included basic aspects of speech motor development, identifying motor speech disorders, speech motor hierarchy followed by assessment, goal selection, treatment techniques, caregiver training and participation, importance of home practice and resources for treatment activities. All clinicians were given 276 presentation slides as handouts.

The second two-day workshop covered the Motor Speech Treatment Protocol (MSTP) followed by group and individual case study activities that focused on clinical observation, goal setting, goal progression and other aspects of treatment planning and execution. All clinicians were given a 30-page manual and approximately 200 workshop presentation slides. The manual provided information on assessment and treatment. Specifically, the treatment section included information on (1) precursors to motor learning (motivation and attention), (2) treatment techniques to support speech skill development (multisensory cueing and motor learning considerations), (3) structure of practice to support acquisition, retention and generalization of motor speech skills (practice distribution and amount), (4) how to provide meaningful feedback that improves production and supports the development of self-monitoring skills (knowledge of results, knowledge of performance, frequency of feedback), (5) structure of treatment activities, (6) caregiver education/training strategies and home practice, and (7) assessing response to intervention.

Following this training, the SLPs completed two online video-based assignments with guided observations and learning tools. Finally, they piloted the MSTP as a case study in their own clinics. Treatment plans and activities for these activities were evaluated by a panel of experts (first, second and third authors) and feedback was given to the clinicians on their performance. In total, clinicians were provided with approximately 50 hours of training in the assessment and treatment of developmental motor speech disorders, including CAS. All clinicians were graded for the quality of their work and were given quantitative and qualitative feedback relating to the video-based assignments and case study. The competency levels ranged from 62.5% to 100% (median = 93.75%; mean = 88.50%; SD = 12.81%) following training prior to the start of the study.

Clinician assignment (stratified randomization approach)

To account for variability in clinician skill levels (scores on video assignments and case study; see the clinician training section above), differences in training (e.g. PROMPT certified) and the amount of experience in working with children with motor speech disorders (based on post-workshop survey results), a stratified randomization approach was used to assign approximately 50% of clinicians to the lower intensity paradigm and approximately 50% to the higher intensity paradigm. This was done to avoid assigning more (or less) experienced/skilled clinicians to one group and biasing results.

In total there were 60 clinicians from 45 different clinics/centres (representing 31 preschool speech and language service areas in Ontario) who participated in the training phase described above. Approximately 46 of these clinicians were available at the time of randomization (23 lower and 23 higher intensity group) to participate in the research phase of the study. The study design used an intention-to-treat principle. Although all clinicians/participants adhered to their allocated intensity groups maintaining 'per protocol' integrity, there was significant data attrition. At the time of analysis there were only completed data sets from 16 and 10 clinicians for the lower and higher intensity treatment paradigms, respectively. Clinician attrition was due to several factors: workplace restructuring, clinician resignation (unrelated to study), maternity and medical leave, undisclosed personal reasons, clinic/centre exempted from study participation by the Ministry of Children and Youth Services due to funding and/or logistical reasons, and inability to procure ethics approval (for some clinical sites) in time for data collection. Although unsystematic in nature, this attrition resulted in one group (higher intensity) having a higher number of clinicians (and consequently participants) than the other group (lower intensity). Participant data loss (loss to follow-up), on the other hand, was fairly equal for the treatment groups (nine for lower intensity; seven for higher intensity) and was mainly due to audio-video malfunction and illness.

Independent variable (intervention)

The MSTP was the treatment approach used in this study. It uses motor learning principles in combination with temporal and multisensory cueing strategies in an effort to increase speech intelligibility through improved speech motor control and speech–sound production (Namasivayam *et al.* 2013a). Motor learning principles such as using a combination of mass and distributed practice, incorporating multiple practice opportunities and providing feedback relating to both knowledge of performance (how a movement was executed) and knowledge of results (the movement's success) are integral to MSTP. Treatment progresses from simple single words to longer, more complex utterances in order to promote the acquisition and generalization of the new speech skills (Maas *et al.* 2008). During therapy, auditory ('say __'), visual ('watch me'), and tactile (physical prompt) cues are used directly to help the child alter their speech production pattern. By manipulating the rate of speech and the delay between the model and the child's production, the clinician can structure the amount of temporal support the child is receiving (Strand *et al.* 2006). Tactile cues are generally used to provide support when gaining control over a movement (e.g. jaw grading) or for cuing a movement (e.g. lip rounding) during speech production (Hayden *et al.* 2010). Treatment goals typically progress hierarchically beginning at the subsystem level (e.g., jaw, labio-facial, lingual, etc.; Hayden and Square 1994) at which the child is demonstrating difficulty. In therapy sessions, skills are practised in structured play activities using functional words and phrases that are meaningful to the child and family.

The overall structure of MSTP session is as follows: The caregiver is present in the room and participates in the therapy process. Each session follows the same routine and begins with the review of home practice. The clinician spends time speaking with the caregiver about home practice successes and challenges and then reviews the material with the child. The focus of the first therapy activity is the development of new speech skills. The clinician uses simple materials to reinforce the child's production attempts as they practise new target words/speech movement goals. The next three to four activities are more naturalistic in nature (e.g. book, game or craft) and the treatment words are practised as the clinician and child engage in the activity. The caregiver participates throughout the session and has the opportunity to practise strategies to support his/her child's speech development at home. The clinician explains relevant information, demonstrates and coaches the caregiver during the activity, and ensures the caregiver is able to do the home practice activity with the child. Strategies for incorporating the child's target words/speech movement goals into everyday routines are also discussed. All treatment sessions are carried out in rooms that are child friendly, quiet and comfortable with appropriate décor.

Dependent variables

Treatment-related change was measured based on changes in the speech–sound system, speech intelligibility and functional communication. These three domains and details on how measures were collected are described in the following subsections.

Change in the speech–sound system

Change in the speech–sound system was measured pre- and post-treatment using the Sounds-in-Words subtest of the Goldman–Fristoe Test of Articulation—2nd Edition (GFTA-2; Goldman and Fristoe 2000). While this test is primarily diagnostic, it also allows SLPs to track changes post-treatment. This assessment requires the child to label pictures spontaneously while the SLP transcribes their productions, noting the articulation errors at the single-word level. The SLPs administered the GFTA-2 according to the standardized administration instructions provided in the manual and scores are calculated based on the number of articulation errors the child makes.

Speech intelligibility

Measures. Speech intelligibility was measured at the word level using the Children's Speech Intelligibility Measure (CSIM; Wilcox and Morris 1999) and at the sentence level using the Beginner's Intelligibility Test (BIT; Osberger *et al.* 1994). Both assessments are often used to measure the effectiveness of speech therapy as they evaluate the impact of a SSD on a naive listener's ability to understand what a child is saying. The details of the speech intelligibility assessments are described in more detail below.

Procedure. Both the word (CSIM) and sentence (BIT)-level tasks require the child to imitate a word or sentence after the clinician's model. In the CSIM, the child is audio-recorded imitating the SLP model of 50 words. The edited recording is then played to naive listeners who must select what word the child said from the closed set of 12 phonetically similar words. A different 50-word list is used in the pre- and post-treatment assessments. For the BIT, the SLP models one of four 10-sentence lists and audio-records the child repeating the sentence directly after the therapist. The recording is then played to naïve listeners afterwards (described below) who must complete the open-set task by writing down the sentence they hear the child saying.

Stimulus materials. The audio-recordings of the CSIM and BIT were collected using Zoom digital recorders with a sampling rate of 44.1 kHz and a resolution of 16 bits/sample. The samples were saved as .wav files and digitally edited using the PRAAT program (Version 5.3.21; Boersma and Weenink 2012) to remove the SLP's model sentences, verbal instructions and any additional noise in the recording (e.g., laughing, other people talking, etc.). Playlists were generated by pairing the stimulus items preceded by instructional markers (e.g., 'List 1', 'Item 5') in order to help the listeners. Each stimulus item was played twice at an adjusted root mean square loudness of approximately 70 dB SPL.

Data collection. For each child there were four audio files (CSIM and BIT at pre- and post-) which resulted in a total of 148 audio files for the 37 children. These 148 audio files were divided into 25 playlists of approximately six randomly chosen files per playlist. Each playlist was played to a different group of three listeners, for a total of 75 listeners (25 playlists × 3 listeners). The naïve listeners (females: 53, males: 22; age range: 18–47 years; mean = 23.5 years; SD = 6.0) were recruited from the University of Toronto. To qualify, individuals had to report little to no exposure to children with SSDs and pass a pure tone hearing screening at 1, 2 and 4 kHz at 25 dB HL. Each assessment (BIT or CSIM at pre- or post-treatment) was judged by three listeners who were blind to the session type, participant information and disorder classification. To control for learning effects, the same listeners were never used for more than one assessment and all audio files were played in a random order. During the rating sessions, the playlists were played via headphones (Sony MDR-XD10) connected to a laptop computer (HP Touch Smart TM2—Intel Core Duo, 64-bit Windows 7 OS) and a multi-channel headphone amplifier (PreSonus HP60). This set-up allowed each listener to adjust the volume of the stimuli to their own personal comfortable listening level (Ertmer 2010, 2011). As mentioned above, the listeners had to either circle the word they heard from a list of 12 words (CSIM) or write down the sentence they heard (BIT). If listeners were unsure of what they heard, they were encouraged to guess in both tasks. Scores were calculated as the percent of words correctly circled in the CSIM and percent of target words correctly transcribed in the BIT, taken as an average across all three listeners.

Functional communication

Functional communication was measured using the Functional Outcomes for Children Under Six tool (FOCUS; Thomas-Stonell *et al.* 2010). This 50-item rating scale was completed both pre- and post-treatment by the child's parent as a measure of change in their child's ability to communicate in everyday life. The questionnaire correlates with quality-of-life measures and has established reliability and validity (Thomas-Stonell *et al.* 2010, 2013). It has been approved by the Ministry of Children and Youth Services as an outcome measure to be used with children under 6 years of age in the Province of Ontario.

Treatment fidelity and integrity

Treatment fidelity. A treatment fidelity checklist specific to the motor speech treatment protocol was developed (see appendix B) and includes three key aspects of fidelity, namely Section I. Competence, Section II. Adherence, and Section III. Treatment Differentiation. Sections I and II allow for quantitative analysis, and Section III involves a binary yes/no response.

- Section I. Competence: This section assessed the quality or skilfulness with which the therapist delivered the treatment and focused on questions such as whether the motor speech goals were appropriate for the child, whether target sounds, word shape and vocabulary were appropriate for the goal and if the clinician could adapt goals according to the child's needs. Section I evaluation was carried out by a clinician outside of the assessment and treatment process, from The Speech and Stuttering Institute, who has over 25 years of experience specializing in assessment and treatment of motor speech disorders in children (including those with CAS).
- Section II. Adherence: This concept is defined as the use of prescribed intervention techniques in the treatment session and the absence of proscribed practices. Section II captures key elements such as intervention setting, dosage, practice schedule, feedback amount, type and delay that have been previously recognized as necessary for successful motor speech treatment (e.g. Maas *et al.* 2008, McCauley *et al.* 2010). Five SLPs went through a 6-h training programme on completing Section II of the treatment fidelity checklist and practised on three case samples until reaching a 100% consensus on scoring prior to the start of the study. These SLPs assessed one video-taped treatment session (midpoint of the treatment block) and three therapy/lesson plans (beginning/mid/end) for each participant.
- Section III. Treatment Differentiation: This section accounted for whether or not the child received any other speech treatment other than the MSTP during the study (based on weekly parental report). There was only one participant who received additional treatment during the course of the study. This participant was removed from further analysis. The study-wide average fidelity score was 85.36% (SD = 8.9).

Treatment integrity. Treatment integrity is an aspect of treatment fidelity and refers to whether or not an intervention is delivered as planned or intended (Kaderavek and Justice 2010). The quantity and quality of treatment administered influences the internal validity and power of intervention studies and is strongly correlated with treatment gains (e.g. van Otterloo *et al.* 2006). Thus, to maintain a high level of treatment integrity the following measures were implemented:

- Quantity of treatment administered: average duration of each treatment session was fixed at 45 min and number of sessions attended was greater than/equal to 80% of the recommended number (i.e., 8/10 or 16/20 sessions for the lower versus higher intensity paradigm).
- Quality of treatment administered: this was monitored comprehensively at a number of different levels: All clinicians were required to submit (1) one videotape of a typical treatment session, (2) three therapy session plan forms (one at the beginning of treatment and one approximately half-way through treatment and one near the end of treatment), (3) a clinical progress report with information on treatment goals, overall success levels and recommendations, (4) three quality of parent–child interaction observation scales including one based on the session they videotaped, and (5) parent weekly log forms (completed by parents). Parents are partners in the therapy process and are integral to the success of therapy. In fact, it has been shown that treatment outcomes improve and outcome variability decreases with consistent home practice provided by the parent (Günther and Hautvast 2010). Since motor speech treatment depends heavily on motor practice and learning, monitoring the quantity and quality of practice is paramount. Quantity of home practice (frequency and intensity in minutes) was acquired from weekly parental logs and ranged from a minimum of 126 min to a maximum of 2897 min across 10 weeks; mean = 556.61; SD = 462.07). The quality of practice was monitored using the five-point parent–child interaction scale (van Otterloo *et al.* 2006), which was adapted to monitor (1) the child's focus during the therapy activity, enthusiasm and responsiveness to parent (Part 1—Child-centred) and (2) the parents' understanding of their child's speech difficulties, their observational skills, their ability to use a variety of strategies and provide appropriate feedback to support speech development (Part II—Parent-centred).

Data reliability

Five qualified SLPs performed blind assessments of approximately 40–50% of the entire data set pertaining to the GFTA-2 (Goldman and Fristoe 2000) based on audio/video recordings of the assessment sessions. A point-by-point agreement index was derived by comparing broad transcriptions (diacritics were excluded) obtained by each rater for the 'Sounds-in-Words' subtest of the GFTA-2. The average inter-rater reliability (per cent agreement) was 83.8%.

Statistical analysis

The assumption of homogeneity of variance (Levene's Test of Equality of Variances) was violated for some variables (e.g., for GFTA-2), hence a repeated-measures analysis of variance (ANOVA) was not completed. Further, data transformation procedures (log, square root or arcsinh) to compensate for unequal variances can be applied only if difference scores are non-negative (Osborne 2002, Zumel and Mount 2013). In the present study, some participants did not show improvements with treatment, hence had zero or even negative scores. Data transformation of such scores would not result in meaningful values (Zumel and Mount 2013). Furthermore, the higher intensity group had almost twice the number of participants ($N = 21$) compared with the lower intensity group ($N = 12$). Given these unequal group sizes, it was decided to split the higher intensity group into two smaller groups (randomly (RND) split into RND 1 ($N = 10$) and RND 2 ($N = 11$) and run separate paired t-tests for each treatment intensity (Nimon 2012). The two higher intensity groups were then compared with the lower intensity group. This procedure is similar in concept to that of bootstrapping and statistical resampling methods reported in the literature (e.g. Efron 1979, Liu and Singh 1992). Similar results across two randomly selected higher intensity groups would strengthen the robustness of the findings and the generalizability of the results. Further, to assess group (lower versus higher) similarity at baseline independent sample t-tests were conducted for all variables.

Several statistical assumptions must be met before t-tests can be applied (Rietveld and Van Hout 2013, Wilcox 2010). These include assumption of normality of the difference scores, assumption of equal variances between the groups (Pitman's test; Pitman 1939), non-significant skewness and kurtosis of difference scores and assumption of non-significant interaction between participants and treatment (Tukey's Test of Additivity). The results of the tests indicated normality of the difference scores, non-significant interaction between participants and treatment, equal variances between the groups for most variables (with the exception of BIT showing a decrease in post-treatment SD; $p = 0.04$) and non-significant skewness and kurtosis for difference scores. With these assumptions met, paired two-tailed t-tests were carried out for each variable across lower and higher intensity groups (RND lists 1 and 2 separately). Since multiple significance tests were conducted for four dependent variables, a Bonferroni correction was applied ($0.05/4 = 0.0125$) to maintain the alpha level at 0.05. To assess magnitude of change as a function of treatment intensity, effect sizes were calculated using Cohen's d statistic corrected for dependence between the means (Cohen 1988, Morris and DeShon 2002) where 0.2, 0.5 and 0.8 are typically referred to as small, medium and large effects, respectively.

Results

Independent sample t-tests indicated that GFTA-2 standard score ($t(31) = 1.2$, $p = 0.22$), FOCUS ($t(30) = -0.54$, $p = 0.58$), CSIM ($t(23) = 0.57$, $p = 0.57$), and BIT ($t(20) = -0.16$, $p = 0.87$) baselines were similar across lower and higher intensity groups. Results from paired two-tailed t-tests indicate that lower intensity treatment (table 1) does not yield any significant pre-/post-change across any of the dependent variables. On the other hand, both higher intensity groups (RND lists 1 and 2; tables 2 and 3) yield significant results for changes in articulation (GFTA-2 standard score) and functional communication (FOCUS scores). Neither lower nor higher intensity treatment yielded a significant change for speech intelligibility at the word or sentence level. Variables with significant pre-/post-change are indicated by $*$ in table 2 and 3. Additionally, for higher intensity treatment, 95% confidence intervals (CI) do not encompass zero indicating robust effects for most variables, except for sentence-level speech intelligibility scores (BIT). For lower intensity treatment, 95% CI encompass zero for two (GFTA-2 standard scores and BIT) of the four variables. Figures 1 and 2 depict pre-/post-treatment changes across dependent variables for both lower and higher intensity (RND lists 1 and 2) groups. Overall, these results indicate that for participants with CAS, 1×/week MSTP intervention does not result in positive and significant treatment benefits. In contrast, there are clear, robust and significant treatment benefits for this population with the higher intensity treatment.

Cohen's d was calculated for the mean difference between pre-/post-treatment for the lower and higher intensity groups (RND lists 1 and 2). In general, MSTP treatment yielded large positive effect sizes for most variables (except for sentence-level speech intelligibility). Figure 3 indicates large effect sizes for articulation (GFTA-2 standard score) and functional communication outcomes (FOCUS scores) in the higher intensity treatment groups relative to the lower intensity group. The speech intelligibility outcome measures yielded mixed results. With regards to word-level speech intelligibility (CSIM scores), both the lower intensity treatment group and one of the higher intensity treatment groups (RND 1) had similar and large effect sizes, whereas the RND 2 higher intensity group had a smaller effect size. For sentence-level speech intelligibility (BIT

Table 1. Data for the lower intensity group across all variables: paired samples test—lower intensity group

	Paired differences							
				95% Confidence interval of the difference				
	Mean	SD	Standard error of the mean	Lower	Upper	*t*	d.f.	*p* (two-tailed)
GFTA-2 Pre-/Post-standard score	−4.0833	7.8446	2.2645	−9.0676	0.9009	−1.803	11	0.099
FOCUS Pre-/Post	−27.727	35.003	10.554	−51.243	−4.212	−2.627	10	0.025
BIT Pre-/Post	−0.08833	7.03181	2.87073	−7.46777	7.29110	−0.031	5	0.977
CSIM Pre-/Post	−7.39286	7.19651	2.72003	−14.04852	−0.73719	−2.718	6	0.035

Note: Bonferroni corrected p-value = 0.0125.

Table 2. Data for the higher intensity group (RND list 1) across all variables: paired samples test—higher intensity RND List 1

	Paired differences							
				95% Confidence interval of the difference				
	Mean	SD	Standard error of the mean	Lower	Upper	*t*	d.f.	*p* (two-tailed)
GFTA-2 Pre-/Post-standard score	−7.5000	4.1966	1.3271	−10.5020	−4.4980	−5.652	9	0.000*
FOCUS Pre-/Post	−30.900	29.233	9.244	−51.812	−9.988	−3.343	9	0.009*
BIT Pre-/Post	−6.12833	12.78660	5.22011	−19.54705	7.29038	−1.174	5	0.293
CSIM Pre-/Post	−9.42833	7.39128	3.01748	−17.18500	−1.67166	−3.125	5	0.026

Note: *Significant at Bonferroni corrected p-value = 0.0125.

Table 3. Data for the higher intensity group (RND list 2) across all variables: paired samples test—higher intensity RND List 2

	Paired differences							
				95% Confidence interval of the difference				
	Mean	SD	Standard error of the mean	Lower	Upper	*t*	d.f.	*p* (two-tailed)
GFTA-2 Pre-/Post-standard score	−4.818	3.894	1.174	−7.434	−2.202	−4.104	10	0.002*
FOCUS Pre-/Post	−36.364	32.113	9.682	−57.938	−14.790	−3.756	10	0.004*
BIT Pre-/Post	−0.33500	7.74900	2.45045	−5.87830	5.20830	−0.137	9	0.894
CSIM Pre-/Post	−7.88800	8.41439	2.66086	−13.90729	−1.86871	−2.964	9	0.016

Note: *Significant at Bonferroni corrected p-value = 0.0125.

scores) a moderate effect size was evident for only one of the higher intensity groups (RND 1).

Discussion

The primary purpose of the present study was to explore the effects of treatment intensity on outcome measures for children with CAS undergoing motor speech intervention. A secondary purpose was to assess the magnitude of change as a function of treatment intensity across outcome measures in this population.

Results indicate that both higher intensity groups (RND lists 1 and 2; tables 2 and 3) yielded significant results for changes in articulation (GFTA-2 standard score) and functional communication (FOCUS scores). However, lower intensity treatment (table 1) did not yield any statistically significant results. Importantly, there were no significant changes in speech intelligibility scores (for word or sentence level) across either treatment intensity. In general, higher intensity treatment produced large effect sizes for the articulation (GFTA-2 standard score) and functional communication variables (FOCUS scores) and moderate effect sizes for sentence-level speech intelligibility (BIT for higher intensity RND list 1; figure 3). On the other hand, lower intensity treatment yielded relatively smaller effect sizes than higher intensity treatment for articulation and functional communication, but for word-level speech intelligibility (CSIM) both lower and higher intensity (RND 1) treatments resulted in similar magnitude of effect sizes. For sentence-level speech intelligibility (BIT scores) only one higher intensity treatment group (RND 1) showed moderate effect sizes.

In general, these data indicate that individual motor speech treatment (MSTP) delivered 2×/week is effective in improving articulation and functional

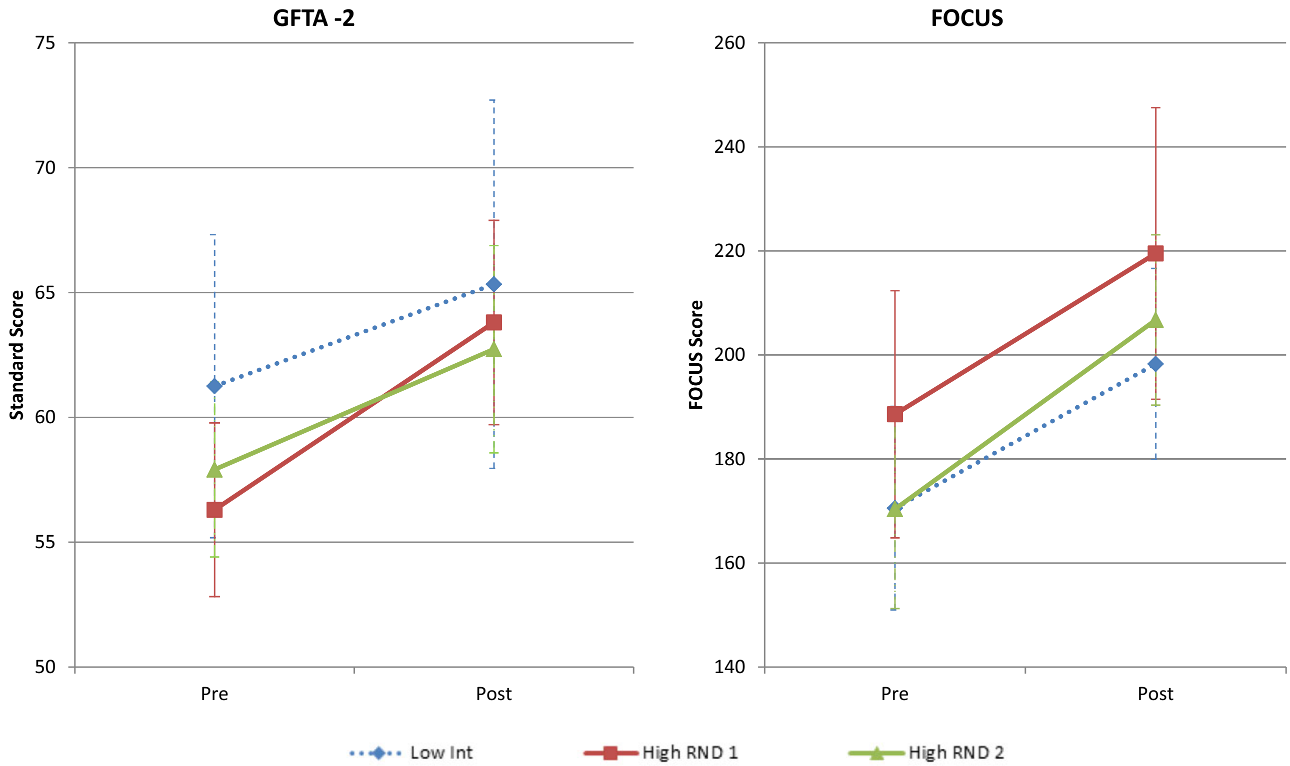

Figure 1. Pre-/post-treatment changes in GFTA-2 standard score and functional outcomes (FOCUS score) across lower and higher intensity treatments (RND lists 1 and 2).

communication in children with CAS. Additionally, children with CAS who received 2×/week individual treatment significantly outperformed those that received 1×/week treatment. This finding is consistent with recent data reported from various treatment approaches with other populations (SSD/developmental delay/autism spectrum disorders (ASD)) and provides evidence to support the general idea that a more intensive intervention may yield better outcomes (Allen 2013, Schooling *et al.* 2010).

With regards to articulation scores, the mean pre-treatment GFTA-2 standard scores and the amount of change in scores following MSTP intervention in the current study are similar to those reported in the literature. In the present study, the mean pre-treatment articulation (GFTA-2) standard score of 58.63 (figure 1) for children with CAS is similar to those reported in the literature (e.g. mean GFTA-2 standard score of 53; Preston *et al.* 2013: 630). These scores are slightly more severe than those reported for children with SSD with motor speech involvement without CAS (e.g. mean GFTA-2 standard score of 63.67; Namasivayam *et al.* 2013b). However, a more careful examination of the raw data from these original studies (Preston *et al.* 2013: 630, table 1; Namasivayam *et al.* 2013b: table 3) indicates an overlap of GFTA-2 standard scores between those children with CAS and those with SSD with motor speech involvement without CAS. This overlap in scores is not surprising as the GFTA-2 scores represent an assessment of consonant speech–sounds at the single word level and are less likely to capture the difficulties in motor planning and/or sequencing that children with CAS tend to demonstrate (Namasivayam *et al.* 2013b). With regards to treatment change, only higher intensity treatment yielded significant changes in articulation scores following MSTP intervention (tables 2 and 3). Mean change following MSTP intervention in GFTA-2 standard scores was 4.0 for the lower intensity group and 7.5 and 4.8 for higher intensity groups (RNDs 1 and 2, respectively).

In terms of functional communication outcomes (FOCUS scores), effect sizes for both the lower and higher intensity groups were generally large (> 0.8), with greater magnitudes for the higher intensity groups (> 1.5) compared with the lower intensity group (1.12). The mean difference (pre-/post) in FOCUS scores in the present study is 27 points for the lower intensity treatment group and 30–36 points for the higher intensity treatment groups (tables 1–3). These mean differences are well above what is considered the minimal clinically important difference (MCID = delta FOCUS > 16 points; Thomas-Stonell *et al.* 2013). Delta FOCUS

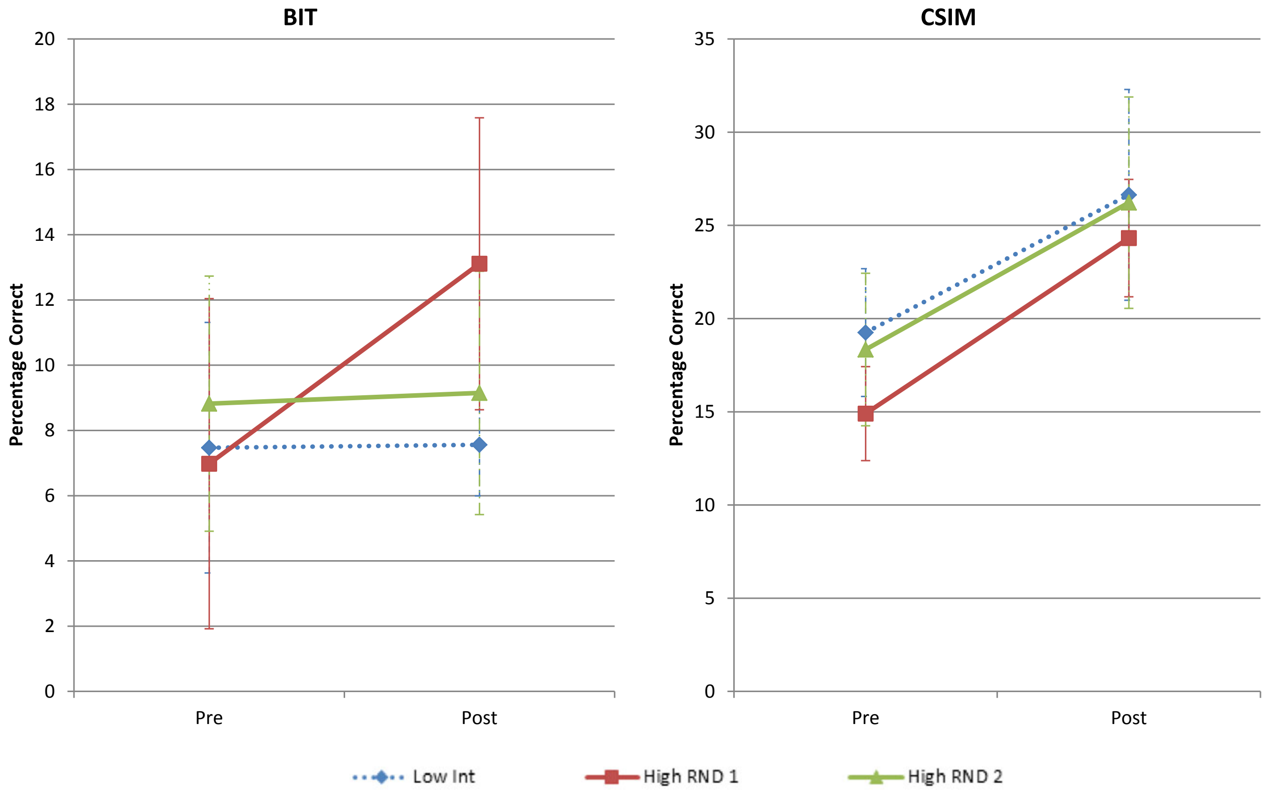

Figure 2. Pre-/post-treatment changes for word (CSIM) and sentence-level (BIT) speech intelligibility across lower and higher intensity treatments (RND lists 1 and 2).

scores from the present study are also larger than those reported recently by Thomas-Stonell *et al.* (2013). In their paper, children with speech impairments (mean age = 3.75 years; SD = 0.78) comprising of children with cleft palate, cerebral palsy and other syndromes receiving an average of 8.7 hours of treatment targeting articulation/phonology and intelligibility were found to have delta FOCUS scores of 18 points. The difference in scores between the current study and Thomas-Stonell *et al.*'s is to be expected due to several reasons. First, there is a difference in hours of treatment received with 10–20 hours (lower and higher intensity treatments, respectively) in the current study versus an average of 8.7 hours in Thomas-Stonell *et al.*'s study. Also, the treatment delivery varied between the studies with the current study delivering individual treatment and Thomas-Stonell *et al.*'s study reporting a combination of group and individual treatments. Finally, the populations between the two studies are different with Thomas-Stonell *et al.*'s study including children with medical diagnoses of structural and neurological disorders and the current study excluding such participants. Despite these differences, it is still noteworthy that the present findings indicate that children with CAS receiving only 10 hours of motor speech treatment can produce a minimally clinically important difference or functional change.

Despite the large effect sizes for articulation, functional outcomes and word-level speech intelligibility, connected speech intelligibility (e.g. at the sentence level) demonstrates negligible change even with a 2×/week 10-week treatment block (e.g. as seen in the higher intensity RND 2 group). These speech intelligibility results are in contrast to children with SSDs with motor speech involvement without CAS (Namasivayam *et al.* 2013b). In the study by Namasivayam *et al.* (2013b), pre-treatment CSIM and BIT scores for children with SSDs with motor speech involvement without CAS were approximately 50% and 40%, respectively, whereas children with CAS in the present study demonstrated pre-treatment scores between 15% and 20% for CSIM and between 7% and 9% for BIT (figure 2). Summarizing the data from the current study and that from Namasivayam *et al.* (2013b), the ratio of word to sentence-level speech intelligibility scores is higher in children with CAS (15–20/7–9 = 2.14–2.22) than for children with SSD with motor speech involvement without CAS (50/40 = 1.25). A similar word to sentence-level speech intelligibility ratio of 1.6 was also found in a group of children with moderate to severe SSD with motor speech involvement without CAS (Namasivayam *et al.* 2013a). This ratio has important ramifications: the higher ratio may indicate an increased

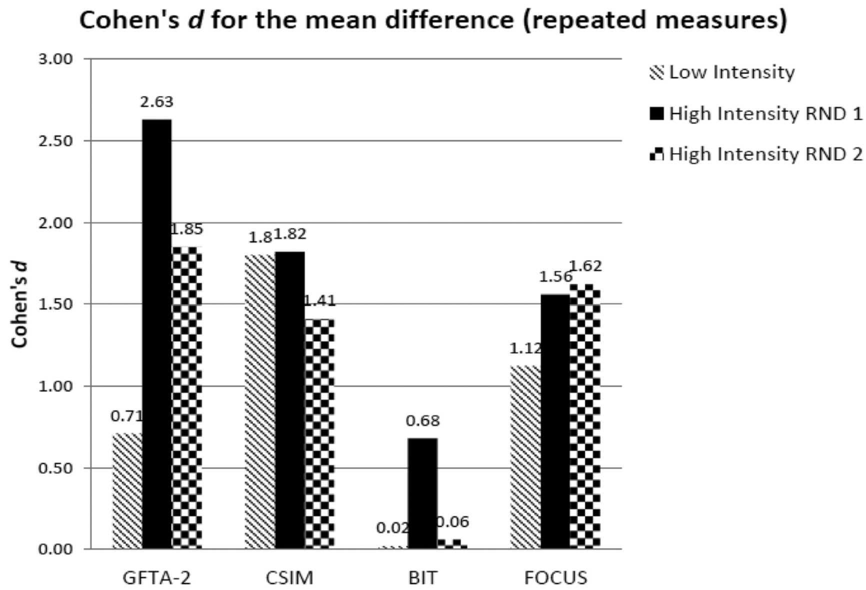

Figure 3. Within-group effect size (Cohen's statistic for various measures pre- and post-treatment for lower intensity and higher intensity (RND lists 1 and 2) groups.

demand on the speech motor system when producing longer and more complex utterances (Maner *et al.* 2000, Pennington *et al.* 2009). This increased demand on the speech motor system may have more adverse effects on connected speech (sentence level) intelligibility scores in children with CAS than in children with SSD with motor speech involvement without CAS.

Importantly, even with a 2×/week 10-week block of intervention children with CAS are not yet at an acceptable speech intelligibility level for their age. Typically developing children between 3 and 6 years old are expected to have sentence-level intelligibility (BIT) scores that approximate 71–99% (Chin *et al.* 2003). Several researchers have pointed out that speech that is < 60% intelligible is difficult for unfamiliar listeners to understand and may negatively impact a child's interaction in social settings (Ertmer 2010, Monsen 1981, Monsen *et al.* 1988). Further, Gordon-Brannan and Hodson (2000) have indicated that children with speech intelligibility scores of < 66% should be considered potential candidates for intervention. The speech intelligibility outcomes from the present study indicate that children with CAS may need additional intervention beyond the 2×/week 10-week block. In fact, a second consecutive block may be recommended for those who continue to present with a severe SSD and poor intelligibility following the initial intervention block.

There are several limitations in the present study. First, although these data were from a large-scale study, there was a significant clinician attrition that resulted in an imbalance in participant numbers between lower and higher treatment intensity groups. Second, there may be potential power issues underlying the lack of statistical significance for some variables. The achieved (post-hoc) power was estimated based on effect size, sample size and alpha levels using G*Power 3.1.9.2 (Faul *et al.* 2009). The analysis revealed low power for speech intelligibility at the sentence level (BIT) for both lower (5%) and higher treatment intensity groups (RND 1 = 21%; RND 2 = 5%) and for articulation scores (56%; GFTA-2) in the lower intensity treatment group. The achieved power was > 85% for all other variables across treatment intensity groups. Thus, not all non-significant results in the study are due to statistical power issues. For variables such as sentence-level speech intelligibility (BIT scores), the effect sizes were also extremely small, suggesting limited change due to treatment.

Third, overall the data from the present study indicates that children with CAS who received 2×/week motor speech treatment improved significantly in relation to those who received 1×/week treatment. However, it is important to note that in the current study we have controlled for 'total intervention duration' (10 weeks) while manipulating 'dose frequency' (number of hours of intervention over specified period of time: 1× or 2×/week; Warren *et al.* 2007). Future studies might be directed towards investigating the possibility of whether more sessions for the lower intensity group accumulated over a longer period of time (e.g. 20 sessions over 20 weeks) might yield similar effects to the higher intensity group (20 sessions over 10 weeks). These studies would have to carefully design control conditions to address confounding factors such as increased total intervention duration and changes due to maturation.

Conclusions

While intensive treatment has been previously recommended for the treatment of children with CAS, differences in treatment outcomes as a function of treatment intensity have not been systematically studied in this population. The present study contributes to the limited evidence base on treatment efficacy in this population by investigating speech and functional outcomes in a group of preschool age children with CAS subsequent to 1×/week (lower intensity) or 2×/week (higher intensity) individual motor speech treatment. Results from this study corroborate previous recommendations regarding the benefits of higher intensity treatment programmes for children with CAS (ASHA 2007, Maas *et al.* 2008, Strand *et al.* 2006). Overall, children with CAS who received 2×/week (higher intensity) individual MSTP intervention for 10 weeks demonstrated significantly better outcomes for articulation and functional communication compared with those who received 1×/week (lower intensity) intervention. Further, lower intensity intervention yields smaller effect sizes than higher intensity treatment for most variables (except CSIM). However, a single, 10-week block of treatment delivered 1× or 2×/ week does not significantly improve word- or sentence-level speech intelligibility in children with CAS. This suggests that speech intelligibility may need more than one block of intervention to significantly improve in this population. In the future, the results of the current study may allow for modification of service delivery and facilitate the development of an evidence-based care pathway for children with CAS.

Acknowledgements

The study was carried out by the Ministry of Children and Youth Services (Ontario, Canada) via The Speech and Stuttering Institute and a treatment research grant from the Childhood Apraxia of Speech Association of North America (CASANA) awarded to the first author (AN). Thank you to the families who participated in the study. The authors would like to thank the agencies and children's treatment centres and their clinicians who participated in the data collection for this study: Northeastern Mental Health Centre, Northwestern Health Unit, Thunder Bay District Health Unit, One Kids Place, Porcupine Health Unit, Algoma Health Unit, Middlesex London Health Unit, Thames Valley Children's Centre, Elgin tykeTALK, Landsdowne Children's Treatment Centre, Grey Bruce PSLS, Pathways Health Centre for Children—Lambton PSLS, Children's Treatment Centre of Chatham Kent, Huron Perth Health Care Alliance—Stratford General Hospital, St. Mary's Family Centre, John McGivney Children's Centre, ErinoakKids, KidsAbility, Wellington Dufferin PSL Program, Niagara Peninsula Children's Centre, Affiliated Services for Children & Youth, Hamilton PSLS—Hamilton Health Sciences, St. Joseph's Health Care Hospital—Hamilton, Markham Stouffville Hospital, York Region PSLS, Grandview Children's Centre, Royal Victoria Hospital, Haliburton-Kawartha Pine Ridge PSLS: Five Counties Children's Centre, Leeds, Grenville, Lanark Health Unit, PSL: Renfrew, Champlain Community Care Access Centre, Quinte Health Care, Kingston, Frontenac Lennox and Addington Public Health Unit—Hotel Dieu Hospital, Eastern Ontario—PSLS, One Kids Place, Pinecrest-Queensway Health and Community Services, Ottawa PSLS—Children's Hospital of Eastern Ontario (CHEO), Toronto Public Health & Rouge Valley Health System—Centenary site, The Scarborough Hospital—General Campus, The George Hull Centre, The Hanen Centre, The Etobicoke Children Centre, North York General Hospital—Branson site, and The Speech and Stuttering Institute. The authors would also like to thank the speech–language pathologists and research assistants who assisted with this study: Samanta Baker, So Hyung (Erica) Park, Margi Patel, Mariam Komeili, Nicole Richard, Talya Wolff, Elaine Kearney, Katia Klymenko, Katherine Kovler, Edith Zhang, Maia Rotin, Israa Nasir, Carolyn Spavor, Dafna Saltzman, Sepideh Kojori, James Lee, Rene Jahnke, Elena Huang, Beverley Ho, Kelly Ross-Schraibman, Janiece DeJong, Emily Goshulak, Lauren Greenwood, Anita Abesek, Sarah Dindyal, Charla Maye Cooper, Maylin Kanter, Soniya Siva, Stephanie Weleschuk, Rotem Yoffe, Allie Gallinger and Shamera Sathiaruban. ***Declaration of interest:*** The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

ALLEN, M. M., 2013, Intervention efficacy and intensity for children with speech–sound disorder. *Journal of Speech, Language and Hearing Research*, **56**, 865–877.

AMERICAN SPEECH AND HEARING ASSOCIATION (ASHA), 2007, *Childhood Apraxia of Speech* [Position Statement] (available at: www.asha.org/policy) (accessed on 16 August 2013).

BOERSMA, P. and WEENINK, D., 2012, *PRAAT: Doing Phonetics by Computer Version* 5.3.21 [computer program] (available at: http://www.praat.org/).

CARUSO, A. J. and STRAND, E. A., 1999, *Clinical Management of Motor Speech Disorders in Children* (New York, NY: Thieme).

CHIN, S. B., TSAI, P. L. and GAO, S., 2003, Connected speech intelligibility of children with cochlear implants and children with normal hearing. *American Journal of Speech–Language Pathology*, **12**, 440–451.

COHEN, J., 1988, *Statistical Power Analysis for the Behavioral Sciences*, 2nd edn (Hillsdale, NJ: Lawrence Erlbaum Associates).

CROSBIE, S., HOLM, A. and DODD, B., 2005, Intervention for children with severe speech disorder: a comparison of two approaches. *International Journal of Language and Communication Disorders*, **40**, 467–491.

DOLLAGHAN, C. A., 2007, *The Handbook for Evidence-Based Practice in Communication Disorders* (Baltimore, MD: Paul H. Brooks).

EDEAL, D. and GILDERSLEEVE-NEUMANN, C., 2011, The importance of production frequency in therapy for childhood apraxia of speech. *American Journal of Speech–Language Pathology*, **20**, 95–110.

EFRON, B., 1979, Bootstrap methods: another look at jackknife. *Annals of Statistics*, **7**, 1–26.

ERTMER, D. J., 2010, Relationships between speech intelligibility and word articulation scores in children with hearing loss. *Journal of Speech, Language, and Hearing Research*, **53**, 1075–1086.

ERTMER, D. J., 2011, Assessing speech intelligibility in children with hearing loss: toward revitalizing a valuable clinical tool. *Language, Speech, and Hearing Services in Schools*, **42**, 52–58.

FAUL, F., ERDFELDER, E., LANG, A.G., BUCHNER, A., 2009, G∗Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, **39**, 175–191.

FLEISS, J. L., 1981, *Statistical Methods for Rates and Proportions*, 2nd edn (New York, NY: Wiley).

GOLDMAN, R. and FRISTOE, M., 2000, *Goldman–Fristoe Test of Articulation—2* (Circle Pines, MN: American Guidance Service).

GORDON-BRANNAN, M. and HODSON, B., 2000, Intelligibility/severity measurements of prekindergarten children's speech. *American Journal of Speech–Language Pathology*, **9**, 141–150.

GÜNTHER, T. and HAUTVAST, S., 2010, Addition of contingency management to increase home practice in young children with a speech–sound disorder. *International Journal of Language and Communication Disorders*, **45**, 345–353.

HALL, P. K., JORDAN, L. S. and ROBIN, D. A., 1993, *Developmental Apraxia of Speech* (Austin, TX: PRO-ED).

HAYDEN, D., EIGEN, J., WALKER, A. and OLSEN, L., 2010, PROMPT: a tactually grounded model. In A. L. Williams, S. McLeod and R. J. McCauley (eds), *Interventions for Speech–Sound Disorders in Children* (Baltimore, MD: Paul H. Brookes), pp. 453–474.

HAYDEN, D. A. and SQUARE, P., 1994, Motor speech treatment hierarchy: a systems approach. *Clinics in Communication Disorders*, 4 [P. A. Square (ed.), 'Developmental Apraxia of Speech: Assessment'], 162–174.

KADERAVEK, J. N. and JUSTICE, L. M., 2010, Fidelity: an essential component of evidence-based practice in speech–language pathology. *American Journal of Speech–Language Pathology*, **19**, 369–379.

KAUFMAN, N., 1995, *Kaufman Speech Praxis Test for Children* (Detroit, MI: Wayne State University Press).

LIU, R. Y. and SINGH, K., 1992, Efficiency and Robustness in re sampling. *Annals of Statistics*, **20**, 370–384.

MAAS, E., BUTALLA, C. E. and FARINELLA, K. A., 2012, Feedback frequency in treatment for childhood apraxia of speech. *American Journal of Speech–Language Pathology*, **21**, 239–257.

MAAS, E. and FARINELLA, K., 2012. Random versus blocked practice in treatment for childhood apraxia of speech. *Journal of Speech, Language, and Hearing Research*, **55**, 561–578.

MAAS, E., ROBIN, D. A., HULA, S. N. A., FREEDMAN, S. E., WULF, G., BALLARD, K. and SCHMIDT, R.A., 2008, Principles of motor learning in treatment of motor speech disorders. *American Journal of Speech–Language Pathology*, **17**, 277–298.

MAASSEN, B., 2002, Issues contrasting adult acquired versus developmental apraxia of speech. *Seminars in Speech and Language*, **23**, 257–266.

MAASSEN, B., NIJLAND, L. and TERBAND, H., 2010, Developmental models of childhood apraxia of speech. In B. Maassen and P. Van Lieshout (eds), *Speech Motor Control: New Developments in Basic and Applied Research* (Oxford: Oxford University Press).

MANER, J. K., SMITH, A. and GRAYSON, L., 2000, Influences of utterance length and complexity on speech motor performance in children and adults. *Journal of Speech, Language, and Hearing Research*, **43**, 560–573.

MCCAULEY, R. J., WILLIAMS, A. L. and MCLEOD, S., 2010, Future directions. In A. L. Williams, S. McLeod and R. J. McCauley (eds), *Interventions for Speech–Sound Disorders in Children* (Baltimore, MD: Paul H. Brookes), pp. 601–613.

MCCAULEY, R. J. and STRAND, E. A., 2008, A review of standardized tests of nonverbal oral and speech motor performance in children. *American Journal of Speech–Language Pathology*, **17**, 81–91.

MONSEN, R., B., 1981, A usable test for the speech intelligibility of deaf talkers. *American Annals of the Deaf*, **126**, 845–852.

MONSEN, R. B., MOOG, J. and GEERS, A., 1988, *Picture Speech Intelligibility Evaluation* (St. Louis, MO: Central Institute for the Deaf).

MORRIS, S. B. and DESHON, R. P., 2002, Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods*, **7**, 105–125.

NAMASIVAYAM, A. K., PUKONEN, M., GOSHULAK, D., YU, V. Y., KADIS, D. S., KROLL, R., PANG, E. W. and DE NIL, L. F., 2013b, Changes in speech intelligibility following motor speech treatment in children. *Journal of Communication Disorders*, **46**, 264–280.

NAMASIVAYAM, A. K., PUKONEN, M., HARD, J., JAHNKE, R., KEARNEY, E., KROLL, R. and VAN LIESHOUT, P. H. H. M., 2013a, Motor speech treatment protocol for developmental motor speech disorders. *Developmental Neurorehabilitation*. doi:10.3109/17518423.2013.832431

NIMON, K., 2012, Statistical assumptions of substantive analyses across the general linear model: a mini-review. *Frontiers in Psychology*, **3**(**322**), 1–5. doi:10.3389/fpsyg.2012.00322

OSBERGER, M. J., ROBBINS, A. M., TODD, S. L. and RILEY, A. I., 1994, Speech intelligibility of children with cochlear implants. *Volta Review*, **96**, 169–180.

OSBORNE, J., 2002, Notes on the use of data transformations. *Practical Assessment, Research and Evaluation*, **8**(**6**) (available at: http://PAREonline.net/getvn.asp?v=8&n=6.) (accessed on 27 December 2013).

OZANNE, A., 2005, Childhood apraxia of speech. In B. Dodd (ed.), *Differential Diagnosis and Treatment of Children with Speech Disorder* (London: Whurr), pp. 71–82.

PENNINGTON, L., MILLER, N. and ROBSON, S., 2009, Speech therapy for children with dysarthria acquired below three years of age. *Cochrane Database of Systematic Reviews*, **4**, CD006937.

PITMAN, E. J. G., 1939, A note on normal correlation. *Biometrika*, **31**(**1–2**), 9–12.

PRESTON, J. L., BRICK, N. and LANDI, N., 2013, Ultrasound biofeedback treatment for persisting childhood apraxia of speech. *American Journal of Speech–Language Pathology*, **22**, 627–643.

RIETVELD, T. and VAN HOUTM, R., 2013, The *t*-test and beyond: recommendations for testing the central tendencies of two (in)dependent samples in research on speech, language and hearing pathology (submitted).

SCHOOLING, T., VENEDIKTOV, R. and LEECH, H., 2010, *Evidence Based Systematic Review: Effects of Service Delivery on the Speech and Language Skills of Children from Birth to 5 Years of Age* (available at: http://www.asha.org/uploadedFiles/EBSR-Service-Delivery.pdf) (accessed on 16 December 2013).

SHRIBERG, L. D., ARAM, D. M. and KWIATKOWSKI, J., 1997, Developmental apraxia of speech: II. Toward a diagnostic marker. *Journal of Speech, Language and Hearing Research*, **40**, 286–312.

SHRIBERG, L. D., POTTER, N. L. and STRAND, E. A., 2011, Prevalence and phenotype of childhood apraxia of speech in youth with galactosemia. *Journal of Speech, Language, and Hearing Research*, **54**, 487–519.

SKINDER-MEREDITH, A., 2001, Differential diagnosis: developmental apraxia of speech and phonologic delay. *Augmentative Communication News*, **14**, 5–8.

STRAND, E. A. and SKINDER, A, 1999, Treatment of developmental apraxia of speech: integral stimulation methods. In A. Caruso and E. Strand (eds), *Clinical Management of Motor Speech Disorders in Children* (New York, NY: Thieme), pp. 109–148.

STRAND, E. A., STOECKEL, R. and BAAS, B., 2006, Treatment of severe childhood apraxia of speech: a treatment efficacy study. *Journal of Medical Speech–Language Pathology*, **14**(**4**), 297–307.

THOMAS-STONELL, N., ODDSON, B., ROBERTSON, B. and ROSENBAUM, P., 2010, Development of the FOCUS (Focus on the Outcomes of Communication Under Six): a communication outcome measure for preschool children. *Developmental Medicine and Child Neurology*, **52**, 47–53.

THOMAS-STONELL, N., WASHINGTON, K., ODDSON, B., ROBERTSON, B. and ROSENBAUM, P., 2013, Measuring communicative participation using the FOCUS: Focus on the Outcomes of Communication Under Six. *Child: Care, Health and Development*, **39**(**4**), 474–480.

VAN OTTERLOO, S. G., VAN DER LEIJ, A. and VELDKAMP, E., 2006, Treatment integrity in a home-based pre-reading intervention programme. *Dyslexia*, **12**, 155–176.

WARREN, F. S., FEY, M. E. and YODER, P. J., 2007, Differential treatment intensity research: a missing link to creating optimally effective communication interventions. *Mental Retardation and Developmental Disabilities Research Reviews*, **13**(**1**), 70–77.

WILCOX, K. and MORRIS, S., 1999, *Children's Speech Intelligibility Measure* (San Antonio, TX: Psychological Corporation).

WILCOX, R. R., 2010, *Fundamentals of Modern Statistical Methods: Substantially Improving Power and Accuracy*, 2nd edn (New York, NY: Springer).

ZUMEL, N. and MOUNT, J., 2013, *Practical Data Science with R* (Greenwich, CT: Manning), ch. 4.

Appendix A: CAS checklist

CAS Study Inclusion / Exclusion Criteria

THE SPEECH & STUTTERING INSTITUTE

Section C :		
GENERAL SPEECH MOTOR CONTROL	Yes	No
1. Limited variety of speech motor movements (e.g. uses jaw as primary articulator)	☐	☐
2. Preference for well rehearsed sound sequences/words (stereotypic responses) (i.e. use of favorite sound, sound sequences, words).	☐	☐
3. Productions are variable (may produce different sound combinations for the same word).	☐	☐
4. Difficulty in achieving initial articulatory configurations (as evidenced by initial hesitations, pausing, groping).	☐	☐
AND/OR		
Difficulty in transitioning between movement gestures (i.e. lengthened and disrupted coarticulatory transitions between sounds and syllables, and words- e.g. sound prolongation)	☐	☐
SEGMENTAL		
Vowels:		
5. Spatiotemporal vowel errors: Child has limited vowel repertoire and/or reduced vowel space/vowel distortions (i.e. decreased vowel differentiation) and/or abnormal vowel duration.	☐	☐
6. Presence of intrusive schwa (e.g. balack – black; bata – bat)	☐	☐
Consonants:		
7. Child has limited consonant repertoire and/or consonant distortions within age appropriate boundaries *and inconsistent nonstandard productions on the same word e.g. dentalization, glottal stop substitutions, bilabial fricative, idiosyncratic cluster reductions etc.*	☐	☐
8. Child demonstrates age inappropriate phonological processes (e.g. persistence of early developing phonological processes such as final consonant deletion, reduplication, syllable deletion) and/or atypical processes (e.g. backing, initial consonant deletion, stops produced as fricatives).	☐	☐
Syllable Shape and Complexity:		
9. Child has limited syllable and word shapes.	☐	☐
10. Child has difficulty maintaining sound and syllable integrity with increased length and complexity of utterance (e.g. Multi-syllabic words, phrases and sentences). There is evidence of: (check all that apply).		
a. Increased variability / inconsistency of errors.	☐	☐
b. Groping.	☐	☐
c. Decreased Intelligibility.	☐	☐
d. Increased omission e.g. tada – television; capa – caterpillar	☐	☐
SUPRA SEGMENTAL		
Prosody:		
11. Child has atypical intonation (e.g. flat, choppy, staccato speech), Staccato speech (or Syllable Segregation): Variations in speech rate due to prolonged sounds or pauses between sounds/syllables/words.	☐	☐
12. Equal stress or lexical stress errors (e.g. stress on wrong syllables).	☐	☐

Appendix B: Treatment fidelity checklist

TREATMENT FIDELITY EVALUATION FORM

THE SPEECH & STUTTERING INSTITUTE

For Office Use ONLY

Subject ID____________ **Clinician providing Treatment:**____________________

Evaluating Clinician:________________________ **Date:** ________________

Characteristics of Motor Speech Treatment

Section I. Competence: Defined as the quality or skillfulness with which the therapist delivered the treatment.

Goal Setting				
	Are **motor speech goals appropriate** for child	**0** ☐	**1** ☐	**2** ☐
	Are **target sounds, word shape** (CVC. CV. CVCV etc) and **core vocabulary** appropriate for the goal and proposed level of motor speech control.	**0** ☐	**1** ☐	**2** ☐
	Online goal adjustments: Can the clinician adapt goals according to the child's current needs, performance and progress in real time (i.e. the clinician is "tuned" in to the child's motor system) *Does the clinician observe child's motor speech control on-line and adapt goals according to child's performance.*	***No*** **0** ☐	***Sometimes*** **1** ☐	***Yes*** **2** ☐
	TOTAL =		**/ 6**	
Key: 0 = none are appropriate; 1 = few or some are appropriate; 2 = All are appropriate				

Section II. Adherence: Defined as the usage of prescribed intervention techniques implemented in session and the absence of proscribed practices.

PREREQUISITIES & SESSION STRUCTURE		
Prerequisites		
1	**Motivation and Attention:** Clinical setting is child-friendly, quiet and comfortable with appropriate décor. Session and activities are structured to support motivation, engagement, participation. (e.g. appropriate seating, developmentally appropriate activities, predictable routine, positive interaction between clinician and child)	☐
Session Structure	*(The treatment session should follow this sequence of events)*	
2	**Caregiver participation:** *(At least 2 of 4 should be present)* • As appropriate, session begins with a review of home practice successes and challenges with the parent and review of targets with the child. • The caregiver participates in a practice activity.	

TREATMENT FIDELITY EVALUATION FORM

	• The Clinician may explain relevant information, demonstrate for the caregiver and coach him/her to ensure the caregiver is able to do the home practice activity with the child. • Strategies for incorporating the child's target words/speech movement goals into everyday routines are discussed with caregiver.	☐
3	**New skill development:** The focus of the first therapy activity is the development of new speech skills. The Clinician uses simple materials to reinforce the child's production attempts as they practice new target words/speech movement goals.	☐
4	**Skill practice:** The next 3 to 4 activities are more naturalistic in nature (e.g. book, game or craft) and the treatment targets are practiced as the Clinician and child engage in the activity.	☐
	TOTAL = Add ***PREREQUISITIES & SESSION STRUCTURE =*** ***/ 4***	
	TECHNIQUES AND PRINCIPLES *(Active Ingredients)*	
	Therapy Techniques	
1	**Watch and listen**: When Clinician is providing speech models - Child is watching the Clinician's face.	☐
2	**Reduced speech rate:** Clinician modifies his/her rate of production and the Child's production to facilitate accuracy of production	☐
3	**Hierarchical cueing:** Clinician uses hierarchical cueing strategies (e.g. Simultaneous production, imitation, delayed imitation) to improve the quality of the child's production	☐
4	**Multi-sensory cueing:** Clinician uses multi-sensory cueing (e.g. auditory, visual, tactile *as appropriate)* to improve the quality if the child's production *(cross reference with Tx/Lesson plan sheet).*	☐
	Motor Learning Principles	
5	Clinician creates opportunities for multiple repetitions of core vocabulary.	☐
6	Clinician creates opportunities for first massed and then distributed practice within the session.	☐
7	Clinician provides feedback to give child *knowledge of results* (e.g. overall success = "that's right!"; reinforce attempts = "good try") and reinforce participation/ maintain motivation.	☐
8	Clinician provides feedback to give child *knowledge of performance* (e.g. make a bigger mouth, close your lips tighter, you used a round mouth, close your mouth a little bit etc)	☐
	TOTAL = Add sections **Therapy Techniques + Motor Learning principles** = **/ 8**	

Section III. Treatment differentiation:

During the study, did the child **receive any other speech treatment** other than the Motor Speech Treatment Protocol *(check parental weekly report)*	**Yes** ☐ **No** ☐