Model-driven treatment of childhood apraxia of speech: Positive effects of the Speech Motor Learning (SML) approach

Anita van der Merwe

Mollie Steyn

Department of Speech-Language Pathology and Audiology, University of Pretoria, Private Bag X20, Hatfield, 0028, Pretoria, South Africa.

Corresponding author:

Anita van der Merwe, PhD

Professor Emeritus

Department of Speech-Language Pathology and Audiology

University of Pretoria

Private Bag X20, Hatfield, 0028, Pretoria

South Africa

Cell phone: +27 82 8937672

E-mail: anita.vandermerwe@up.ac.za

Abstract

Purpose: To propose the SML approach (Van der Merwe, 2011) as a treatment for CAS and to determine if it will effect positive change in the ability of a 33 month old child to produce untreated nonwords and words containing treated age-appropriate consonants (Set 1 sounds), untreated age-appropriate consonants (Set 2) and untreated age-inappropriate consonants (Set 3); also to determine the nature and number of segmental speech errors before and after treatment.

Method: An A-B design with multiple target measures and follow-up was implemented to assess the effects of treatment of Set 1. Effect sizes for whole word accuracy (WWA) were determined and two criterion lines were generated following the conservative dual criterion method (CDC). Speech errors were judged perceptually.

Results: CDC analyses indicated no reliable treatment effect due to rising baseline scores. Effect sizes showed significant improvement in WWA of untreated non- and real words containing age-appropriate treated sounds and real words containing age-appropriate untreated sounds. The number of errors for all three sound sets declined. Sound distortion was the most frequent error type.

Conclusions: Preliminary evidence suggests potentially positive treatment effects. However, rising baseline scores limit causal inference. Replication with more children of different ages is necessary.

Key words: childhood apraxia of speech, speech motor learning (SML) treatment approach, four-level framework of speech sensorimotor control.

Theoretical models or frameworks which explicate speech motor control could drive hypotheses regarding the pathophysiology and core features of apraxia of speech and also have the potential to provide treatment guidelines. The purpose of the current paper is to propose a model driven treatment for childhood apraxia of speech (CAS) and to explore the effects in a single case with CAS. The construct of the speech motor learning (SML) approach (Van der Merwe, 1985, 2002, 2007, 2011) is grounded in the four-level framework (FLF) of speech sensorimotor control (Van der Merwe, 1997, 2009). To provide a frame of reference, theoretical models and theories which are implemented to explain CAS and guide treatment will be reviewed in the following sections followed by an overview of the FLF and the theoretical basis and methods of the SML approach. Application of the SML to a child with CAS will then be reported.

The American Speech-Language-Hearing Association (ASHA, 2007) technical report on CAS defined it as "a neurological childhood (pediatric) speech sound disorder in which the precision and consistency of movements underlying speech are impaired in the absence of neuromuscular deficits" due to "the core impairment in planning and/or programming spatiotemporal parameters of movement sequences" (ASHA, 2007, p. 2). The report's recommendation of the designation *childhood apraxia of speech* acknowledges shared core features between CAS and acquired apraxia of speech (AOS). Diagnostic markers were not proposed, but some consensus was gained regarding three features signifying a deficit in motor planning and programming. These are (a) inconsistent errors on sounds during repeated productions of syllables or words, (b) lengthened and disrupted coarticulatory transitions between sounds and syllables, and (c) inappropriate prosody (ASHA, 2007). While a validated list of diagnostic speech characteristics for the conclusive identification of CAS at all ages and across all cognitive levels, speech competence levels, and all etiologic contexts have yet to be established (Shriberg, Lohmeier, Strand, & Jakielski, 2012), contemporary diagnostic speech symptoms include speech sound distortions, disrupted coarticulation, struggling or groping articulation, syllable segregation, voicing errors, increasing difficulty with increased length of utterance, and equal stress or lexical stress errors (Hall, Jordan, & Robin, 2007; Iuzzini-Seigel, Hogan, Guarino, & Green, 2015; Nijland, Terband, & Maassen, 2015; Shriberg, Potter, & Strand, 2011). These signs underscore the probable motor-based nature of CAS and the need to explore the potential value of models and theories of speech motor control and motor learning.

Neurolinguists leverage models and theories of motor control from the field of human kinetics (Kawato, 1999; Romo & Schultz, 1992; Schmidt, 1975) to develop current neurolinguistic models of speech motor control (Guenther & Perkell, 2004; Hickok, Houde, & Rong, 2011; Van der Merwe, 1997, 2009). Central to these models is the concept of an internal model or representation of transformation within the central nervous system that converts inputs (aggregate of sensory feedback and an efference copy of the motor command) to output (motor commands). Transformations are bidirectional. A forward model indicates the causal direction - mapping of motor commands onto their sensory consequences, while an inverse model indicates the opposite transformation of a desired sensory consequence into the necessary motor commands. Motor learning requires acquisition by the brain of both forward and inverse internal models for different tasks *(*Kawato, 1999; Wolpert, Ghahramani, & Flanagan, 2001). Neurolinguistic frameworks have enabled theorists to conceptualise a motor planning stage of speech motor control which potentially can explain the notion of an apraxia of speech. Children with CAS proposedly suffer from weak internal models and disrupted feedforward motor control (Iuzzini-Seigel et al., 2015; Terband & Maassen, 2010; Terband, Maassen, Guenther, & Brumberg, 2009; Van der Merwe, 1997, 2009). A computational neural network model of speech acquisition and production, the Directions into Velocities of Articulators (DIVA) model (Guenther & Perkell, 2004; Maassen, Nijland, & Terband, 2010;

Terband & Maassen, 2010; Terband et al., 2009) has been implemented to explore the neuromotor deficits that underlie CAS. The model was linked to an articulatory speech synthesizer which allowed for simulation of deficits in a specific component. Results confirmed the hypothesis that feedforward commands are impaired in CAS (Terband & Maassen, 2010; Terband et al., 2009). A current trend is to integrate neurolinguistic (Guenther & Perkell, 2004; Van der Merwe, 2009) and psycholinguistic models (Levelt, Roelofs, & Meyer, 1999; Ziegler, 2009) as integrated models (Hickok, 2014; Hickok et al., 2011; Nijland et al., 2015) to also account for associated linguistic and cognitive problems present in some children with CAS. But, to treat the core motor impairment, theoretical models and theories of motor control and learning are most informative.

Motor skill learning theory (Guadagnoli & Lee, 2004; Maas, Robin, Hula, Freedman, Wulf, Ballard, & Schmidt, 2008; Schmidt, 1975; Schmidt & Lee, 2005; Wolpert et al., 2001; Wulf & Schmidt, 1997) does currently influence treatment of apraxia of speech. Principles of motor learning during speech skill acquisition have been applied with promising results (for example: Edeal & Gildersleeve-Neumann, 2011; Hula, Robin, Maas, Ballard, & Schmidt, 2008; Maas, Butalla, & Farinella, 2012; Maas & Farinella, 2012; Skelton & Hagopian, 2014) and are incorporated in treatment approaches for both AOS and CAS. Treatments for CAS that are motor-based and apply principles of motor learning include integral stimulation approaches like dynamic temporal and tactile cueing (DTTC) (Strand, Stoeckel, & Baas, 2006) and rapid syllable transition (ReST) treatment (Ballard, Robin, McCabe, & McDonald, 2010; Murray, McCabe, & Ballard, 2015). Principles that have been proven effective for long-term retention and generalization of skills are: distributed, random and variable practice, high number of trials, and also low-frequency and delayed knowledge of results feedback (Bislick, Weir, Spencer, Kendall, & Yorkston, 2012; Magill, 2007; Schmidt & Lee, 2005).

The four-level framework (FLF), which could be classified as a neurolinguistic model, characterises neuromotor speech disorders and typifies an apraxic impairment as primarily a motor planning disorder with additional (perhaps secondary) involvement of motor programming (Van der Merwe, 2009). The FLF was the first to differentiate between motor planning and programming levels of speech processing and to assign areas in the nervous system to specific speech production processes. Recent research supports this division (New et al., 2015). Though this model is basically an adult model, it does delineate the skills required for acquiring and producing normal speech and therefore provide treatment guidelines for both CAS and AOS. The main components of the FLF are submitted in the following section.

The four-level framework (FLF) of speech sensorimotor control

Figure 1 presents a synopsis of the FLF (Van der Merwe, 1997, 2009). The FLF differentiates an initial linguistic and three subsequent motor phases of processing (planning, programming and execution) during speech production. Speech production is a cognitive-motor process (Kent, 2004) in which linguistic contents is transformed into a code that is amenable to the motor system.

A central premise of the FLF (and SML treatment) is that the “building blocks” for speech motor planning is the acquisition of a core motor plan (CMP) for each speech sound. The primary role of single sounds is also acknowledged in the DIVA model (Guenther & Perkell, 2004). The CMP contains an inverse model of the spatial and temporal specifications for the production of a specific speech sound. Each plan contains several speech-structure-specific motor goals, which are adapted to the phonetic context in which the sound is to occur. Adaptation is guided by a critical acoustic configuration – a forward model – by internal predictive control and monitoring of the efference copy. Muscle tone, velocity, range, direction and force parameters are added to the code during the following motor programming phase. Such parameters may be sensitive to circumstances such as the need to speak loudly, to increase rate and to incorporate stress patterns. Accurate execution depends on normal muscle tone and well-coordinated movements.

In the FLF, all phases of speech production are portrayed as context sensitive (Van der Merwe, 1997, 2009). The different contextual factors that could affect the dynamics of sensorimotor control are summarized in Figure 1. In the SML approach contextual factors drive choice of treatment stimuli and strategies in a bottom-up (easy to more difficult) fashion (e.g. initiation mode from imitated to self-initiated, short to longer utterances, slow production to increased speech rate).

Extrapolating from the depiction of speech motor planning and programming in the FLF, skills that need to be addressed in treatment of CAS are the following: acquisition of all motor plans of speech sounds of the language, coarticulation between sounds, keeping adapted movements within the critical acoustic configuration across repeated productions of the same utterance to avoid inconsistency and distortion, planning utterances of increasing length, executing internal feedback and predictive control of productions, and (as the child becomes a more “sophisticated” speaker) managing increased rate and manipulating, for example, force parameters during syllabic stress production. The SML approach was designed to guide the individual in acquiring, establishing, and controlling all the components of motor planning of speech and adding parameters of programming as conceptualised in the FLF, to achieve a sophisticated system capable of planning and programming continuous speech with its rhythmic and melodic properties (Van der Merwe, 2011).

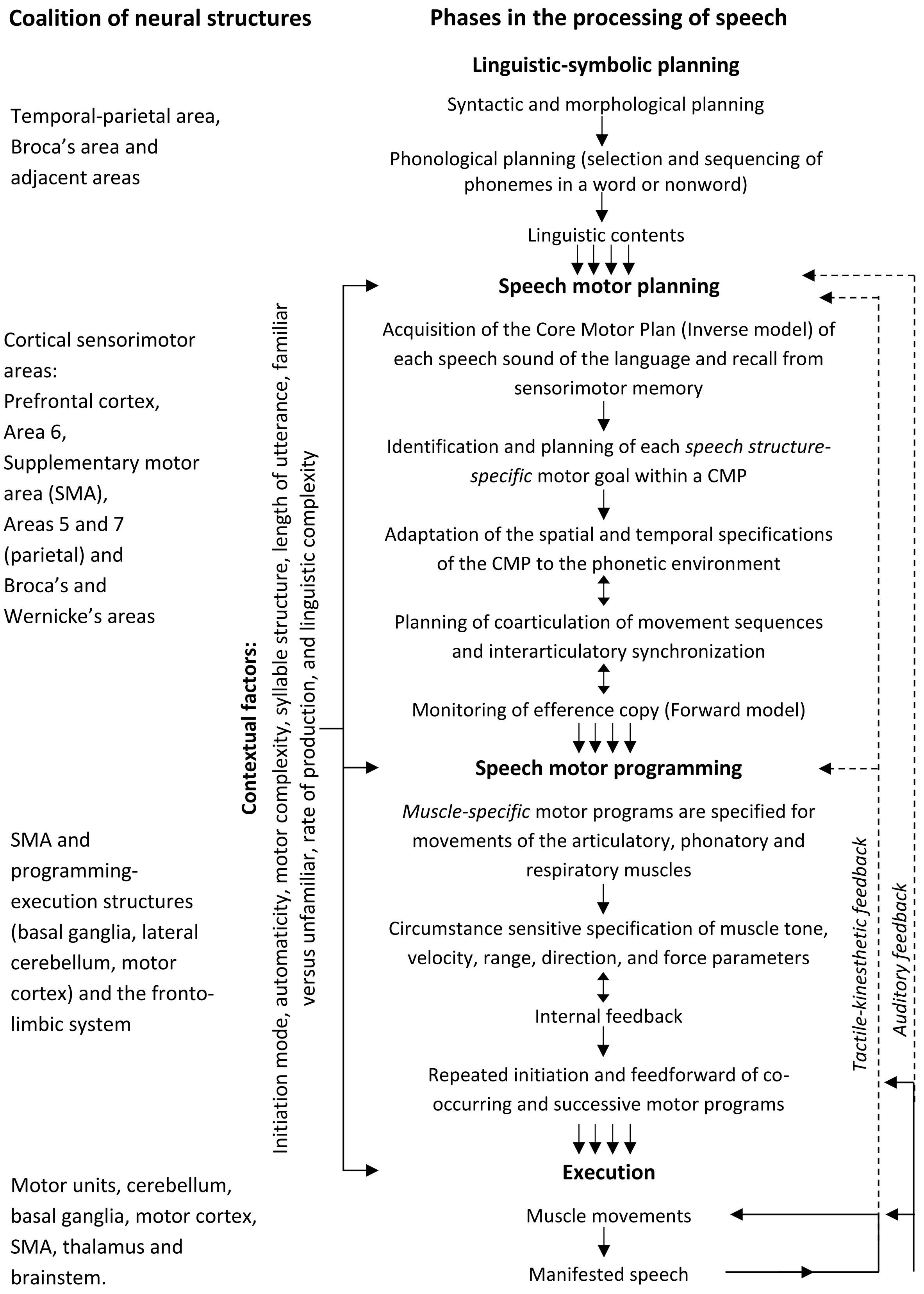

Figure 1. The Four-Level Framework (FLF) of speech sensorimotor control – a synopsis with additional clarifying terms and a greater focus on speech development (original framework in Van der Merwe, 1997, 2009).

Methods of the SML

Components of the SML approach resemble aspects of other interventions (for which evidence suggests effectiveness), like DTTC which uses intensive drill of movement gestures in words and phrases (Strand et al., 2006), ReST which entails rehearsal of a variety of multisyllabic nonwords containing sounds in the child's inventory with varied stress patterns (Ballard et al., 2010), stimulability training of sounds in isolation and CV shapes focusing on expanding the phonetic inventory (Iuzzini & Forrest, 2010; Powell, 1996) and the Nuffield Dyspraxia Programme (NDP), which first targets sounds in isolation and then in syllables and longer syllable sequences (Murray et al., 2015; Williams & Stephens, 2004). Like the SML approach, the latter three approaches implement nonwords as treatment stimuli and all approaches include visual and auditory stimulation, a high number of trials and the use of principles of motor learning. There are, however, features that are unique to the SML.

To build the speech production process step by step, treatment starts with a first set of easy sounds and gradually, in stages and across time, new sounds of increasing difficulty (from the particular speaker's perspective) are added to the already treated set. Pre-treatment assessment focuses on the ability to produce the sounds of the language in isolation and "sound sets" to be targeted across time are decided on at that point. The number of target sounds at any stage is determined by severity of the apraxia. Sounds are rehearsed in systemically varied phonetic contexts with increasing amounts of between-sound variation within a CVCV (C= consonant; V= vowel) nonword, and also across a series of CVCV (at later stages also other syllable shapes) nonwords. All sounds are rehearsed in different phonetic contexts (variable practice) and in different nonwords (random practice of tasks and targets). Both conditions address principles of motor learning (Schmidt & Lee, 2005).

The aim of systematically varied phonetic contexts is to facilitate stability of each core motor plan (inverse model), but also flexibility of motor goals within the plan (the ability to adapt spatial and temporal specifications to the phonetic context). The systematic changes in sound environment potentially enable the speaker to implicitly extract *generalizable rules* (Schmidt, 1975) for accurate production of a sound in different phonetic contexts and for coarticulation between sounds. The process of initially using easy sounds and gradually introducing more difficult sounds improves the probability of accurate speech production and could provide favourable conditions for the acquisition of motor planning and programming skills. Stimuli are not a set number of nonwords or words, as is the case in most other approaches, but multiple exemplars allowed by the possible combinations of target Cs and Vs. Real words and phrases, determined by the target sounds and syllable structure of a stage, are also introduced from the start of treatment. Pre-response delay periods with mental practice, a motor learning principle (Schmidt & Lee, 2005), are imposed to also strengthen the forward model of each sound and facilitate self-initiated production of a non- or real word. The steps followed during treatment further allow for the addition of parameters of programming like increased rate and stress production in series of nonwords and phrases. The NDP (Williams & Stephens, 2004) also appears to build speech production ability according to a long-term treatment plan, but no other approach implements multiple stimuli with systematically controlled phonetic environments and increasing variation between sounds, systematic lengthening of utterances, mental practice, or guidelines for treatment of all levels of severity as the SML does.

The SML, unlike most other approaches, is also suitable as treatment for AOS (Van der Merwe, 2011) with essentially similar stimuli and methods. In CAS, however, targets will be restricted to age-appropriate sounds and some steps may have to be adapted or omitted to accommodate the ability of the child. Approach-specific detail follows next.

Targets

During pre-treatment assessment, stimulability (ability to produce a sound in isolation from imitation), accuracy of production, and appropriateness for age are used to assign sounds to the different *stages* of treatment. The sounds of Stage 1 typically consist of three to four C's and three to five V's that are introduced simultaneously. During Stage 2 another one to three C's are added (number depends on severity of CAS). Across stages the C set is invariably expanded first and then the V set. Once the most stimulable Cs and Vs (about seven of each) have been treated, the CVC syllable structure containing these sounds is introduced. At later stages CVCVC and CVCVCVC structures are addressed. The most difficult sounds, which may be consonant clusters and diphthongs, are (gradually) added last. Supplemental File 1 provides an overview of an SML program across time.

The "groundwork" for all C's and V's is done in CVCV nonwords. The CV structure and reduplication (CVCV) occurs in most languages and is present in infant babbling as well as the first 50 words of children (Edwards & Shriberg, 1983). Treatment stimuli are developed as *series* of about five nonwords. Increasing variation between sounds in a nonword and across a series is achieved by five *levels of variation*. Two examples of series of CVCV nonwords with Level 1 variation would be: /bæbi, bæbu, bæbɔ/ and/ bɛbi, bɛbu, bɛbɔ/ (C1V1C1V+). Plus indicates variation of sounds in that position across a series. Once criterion of in-session 80% correct as judged by the treating clinician (4 out of 5 without distortion, substitution, omission, trans-positioning, or addition of sounds and accurate coarticulation between sounds) during imitated production (in the case of a young child who cannot read) is reached, the same target set is rehearsed with Level 2: (C1V1C2V+ e.g. /bækɔ, bæku, bæki/), then Level 3 (C1V1C+V1, e.g. bɔkɔ, bɔtɔ, bɔfɔ), Level 4 (C1V+C2V1, e.g. /bɛkɔ, bukɔ, bikɔ /) and Level 5 (C1V+C+V+, e.g. /bɛki, bɔtu, bifɔ/) variation. Due to the programmatic nature of CVCV and CVC stimuli, computerized generation is possible and software is available at no cost (Van der Merwe, 2002).

Steps during treatment

The same steps are followed repeatedly. Pre-practice of each target sound in isolation or in CV or VC syllables may be necessary. Subsequently, production of a nonword is demonstrated three times while the child observes and listens. Imitation is attempted repeatedly (blocked practice). Multiple trials monitor consistently accurate production. Knowledge of results (KR) feedback is initially provided 100% of the time, but gradually faded. A three to four second pre-response delay period is imposed after 80% correct production has been achieved. After each nonword in a series has been rehearsed, the clinician produces them one after the other and the child imitates each consecutively. An individual who can read should read the series independently. A series is modelled rhythmically with stress on the syllables that vary. Rate of production is gradually increased within the limits of accuracy. Mental practice and fading KR feedback are both principles of motor learning (Magill, 2007; Schmidt & Lee, 2005) and are integrated into the steps.

Purpose of the research and research questions

The purpose was to explore the effect of the SML approach on the accuracy of production of an expanding corpus of speech sounds. Selected sounds were assigned to three sets which represent stages in SML treatment across time. Set 1 sounds (age-appropriate) were treated while the additional Cs in Set 2 (age-appropriate) and Set 3 (age-inappropriate) acted as control behaviours.

The following research questions were posed: Will treatment effects generalise to: 1) untreated non- and real-word stimuli containing four *treated age-appropriate* Cs and five Vs that have been rated as sounds with the highest measure of stimulability and accuracy (Set 1 sounds); 2) untreated non- and real-word stimuli containing three additional *untreated age-*

appropriate Cs rated as sounds with a lesser measure of stimulability and accuracy (Set 2 sounds); 3) untreated non- and real-word stimuli containing three additional *untreated age-inappropriate* Cs rated with the lowest measure of stimulability and accuracy (Set 3 sounds)?

Method

Participant

The participant was a 33 month old monolingual English-speaking boy with CAS.

Case history. The participant was referred to the first author by a speech-language practitioner in private practice. She had been treating the child since the age of 21 months. Early intervention following the Hanen approach to language therapy (Manolson, 1992), was provided once weekly for approximately one year (age 21 to 33 months). Articulation was not directly treated and showed limited development. According to parental report he had achieved most developmental milestones at the expected ages, but demonstrated reduced babbling as a baby and limited word productions at a later age.

Hearing. An oto-acoustic emission (OAE) test performed by a hospital-based audiologist shortly after birth indicated normal middle ear and cochlear functioning. His hearing was screened again by another audiologist at 24 months and found to be within normal range. There was no history of ear infections.

Oral mechanism. Assessment revealed no overt structural or functional impairments. Occlusion was normal with no dental deviations. No tongue thrust during rest or speech, no abnormality of the hard and soft palate and no signs of a sub-mucous cleft palate were evident. The face, lips, tongue and velum were symmetrical during both rest and movement. He was able to imitate single lingual, lip, and jaw movements and muscle strength, tone, and range of movement appeared normal. Sucking, chewing and swallowing were normal, though occasional mild drooling while playing was observed. No signs of dysarthria due to paresis or involuntary movements of the oral structures (or limbs) were observed. Voice quality was good and respiratory support was adequate.

Language. Due to limited and unintelligible verbal output, expressive language could not formally be assessed and attempts to perform a formal norm-referenced test of auditory comprehension were unsuccessful as the child did not consistently cooperate. We then opted for the Language Comprehension and Language Expression subtests of the Rossetti Infant-Toddler Language Scale (Rossetti, 2006), which are criterion-referenced evaluations in a natural communication context. The participant scored within his age range (33-36 months). Scores on the Action Picture Test and Word Finding Vocabulary Test (Renfrew, 2010) at 4 years 9 months (two years after this study) confirmed age-appropriate language skills.

Development. The Developmental Assessment Schema checklist (Anderson, Fowler, & Nelson, 1978) was implemented and play-based assessment revealed appropriate symbolic and imaginative play skills, as well as age-appropriate scores on the personal-social, perceptual-cognitive and self-help skills subtests.

Picture naming and spontaneous speech. Verbal output during an informal picture naming task comprising 50 familiar words and a 15-minute spontaneous speech-language sample was recorded. Analysis rendered 243 words and revealed the use of one- to two-word phrases, with reduced intelligibility (35% words understood). Utterances consisted mainly of nouns and verbs, with V, CV, VC, VCV, CVCV and CVC syllable structure. Longer words were simplified to these structures. Frequency of occurrence (%) and accuracy (%) of sounds were calculated (see Supplemental File 2). Six out of 13 age-appropriate Cs (Arlt & Goodban, 1976; Dodd, Holm, Hua, & Crosbie, 2003; Prather, Hedrick, & Kern, 1975) had a frequency of 3% or less.

Speech production and diagnosis. During assessment a narrow phonetic approach, which allows for the notation of phoneme size errors, as well as of distortions, was followed.

Stimulability of all English Vs and single Cs in isolation was assessed by eliciting direct imitation of an auditory and visual model without instructions or feedback. Production of stimulable Cs was also assessed in combination with stimulable Vs in CV, VC, CVCV, and CVC combinations to determine coarticulatory ability. For assessment of consistency of production, the child was prompted to repeat sounds, syllables and words 3 to 5 times consecutively. All activities were play-based. Diadochokinesis and lexical stress tasks (Shriberg et al., 2011) were omitted due to limited and unintelligible speech production and the possibility that diadochokinesis could be influenced by factors other than apraxia.

The diagnosis was made by the first author who has more than 40 years' experience in differential diagnosis of motor speech disorders. To meet criteria for a CAS diagnosis the child first, had to display compromised stimulability for, or the inability to produce at least two age-appropriate sounds in isolation (absent or not yet stable core motor plans). Second, he had to display the three features proposed by ASHA (2007) across two or more of the tasks mentioned above. More specifically the child had to display the inability to coarticulate stimulable Cs and Vs in CV, VC, CVCV or CVC contexts and inconsistency in phonetic features of sounds across three consecutive trials in isolation, in the assessed syllable shapes or in real words. Third, the child had to display at least four surface behaviors from Strand's 10-point list (Shriberg et al., 2011) across two or more of the tasks mentioned above and also during the conversation sample. The assessment revealed: compromised stimulability of nine out of 13 age-appropriate Cs; inability to coarticulate stimulable Cs in combination with different stimulable Vs (or a particular V with different Cs) without distortion or pauses; inconsistent perceptual features of sounds in isolation, in syllables and in words across consecutive trials; syllable segregation in CVCV utterances; inconsistent voicing errors across consecutive trials; C and V distortions; distorted substitutions; and the inability to produce words with more than four phonemes. Two of the three ASHA (2007) consensus

Table 1. List of symptoms of CAS used in the diagnosis.*

Symptom	Present	Examples**	Perceived errors
Inconsistent or variable articulation of consonants and vowels across trial-to-trial repeated productions of the same sound, syllable, or word	Yes	*no*: /nəʊ/→ [nʌ; nəʊ; nɑʊ] *hop*: /hɒp/ → [_ɒ_; _ɒp; _ɒt]	Variable vowel quality Final /p/ omitted; correct /p/; /p/ as [t] with weak plosive release
Inappropriate co-articulatory transitions between sounds and syllables	Yes	*foot*: /fʊt/ → [pʊ_t]	Pause between /ʊ/ and /t/
Inappropriate prosody, such as difficulties in lexical stress	No		
Reduced phonetic repertoire in spontaneous speech	Yes	/ f, k, n, g, ŋ, j /	Age-appropriate, but occurrence 3% or less
Vowel and diphthong errors	Yes	*back*: /bæk/ → [tʌ _] *here*: /hiəɹ/ → [jiæ _]	/æ/ perceived as [ʌ] /iə/ perceived as [iæ]
Consonant distortions	Yes	*this*: /ðəs/ → [dɛs] *foot*: /fʊt/ → [pʊ_ ts]	Distortion of /s/ Distortion of /t/
Distorted substitutions of consonants	Yes	*bunny*: /bʌni:/ → [bʌdi:]	Substitution with a distorted [d] (elements of plosive and nasal release)
Omission of consonants or vowels from words	Yes	*sit*: /sət/ → [sə_] *doctor*: /dɒktɝ/ → [dɒ_ tɑ]	Omission of /t/ Omission of /k/
Error production increases with increased length of utterance	Yes	*Mom, pass my dudu*: /mɒm pɑs mai dudu/ → [mɒ_ bæ_ bai du_ ti:]	Omission of second /m/, voicing of /p/, omission of /s/, /m/ perceived as [b] with elements of nasal release, pause between /u/ and /d/, devoicing of second /d/, /u/ vowel error
Non-speech groping behaviours	No		
Voicing errors	Yes	*too big*: /tu: bɪg/ →[du: pɪx]	Voicing of /t/ and devoicing of /b/
Intermittent hyper-nasality	Minor		See examples above
Difficulties in imitating sounds and words or reduced willingness to attempt	Yes		
Oral apraxia for single and/or sequential movements	Mild	Consecutively puff cheeks and move tongue left-to-right	Slow movements with disrupted sequence.

*Speech tasks were: 15 minute spontaneous speech sample, picture naming, imitation of words, imitation of consonants (C) and vowels (V) in isolation and nonword combinations with CV, VC, CVCV, CVC shapes, and three to five times self-initiated production of words and nonwords by providing visual prompts.

** Underlined sounds indicate error being described, but other errors may also be present in the example. An underscore (_) represents an omission of a sound or an inappropriate pause between sounds or syllables.

signs were noted (lexical stress errors and other prosodic errors were not apparent) and six from the 10-point list (Shriberg et al., 2011). As transcriptions were not made of all utterances, the frequency of occurrence of the six behaviors is not available, but each was noted as being evident in the participant's speech. The different CAS symptoms with which the participant presented and some examples of his speech errors are summarized in Table 1.

Ethical considerations. The parents granted informed consent for the study and the participant expressed his willingness to take part in the speech- and play activities prepared by the clinician. A Faculty ethics committee gave ethical clearance for the research.

Design

An A-B (baseline followed by treatment) design with multiple target measures and follow-up (Barlow, Nock, & Hersen, 2009) was implemented. SML treatment was the independent variable and accurate production of untreated non- and real words each containing one of three sound sets acted as dependent variables.

Three pre-treatment baseline probes were performed followed by eighteen 30-minute treatment sessions (9 hours) across nine weeks. Repeated probes of all three sound sets were performed across the nine week treatment period followed by two follow-up probes. No other intervention was provided during the study period.

Experimental stimuli

Selection of three sets of sounds. Three pre-treatment assessments of stimulability and production accuracy of all English Vs and single Cs were performed on three different days. The participant was given five opportunities per day to imitate Vs and Cs in isolation, while the authors made online judgements of measure of stimulability and accuracy. The following 4-point rating scale was implemented: 1-The sound was imitated correctly all or most of the time (four or five attempts correct); 2- three correct; 3- one or two correct; 4- The sound was consistently produced incorrectly or omitted, or no attempt was made to imitate the sound (0 correct). The test-retest ratings were considered in the allocation of Cs to the three target sets.

Given the young age of the participant, the assignment of sounds to each set was further guided by developmental norms for typical English speech sound acquisition. The sound sets were as follows: *Set 1* sounds included /b, m, n, t/ (rated “1”) which are typically mastered by 3;0 years. *Set 2* sounds consisted of Set 1 sounds plus three additional Cs (/d, f, p/), also typically acquired by 3;0 years but rated “2”. *Set 3* sounds consisted of Set 1 and 2 sounds plus three additional Cs (/l, s, v/), typically mastered between 3;0 and 5;0 years and rated “4”. Dodd et al. (2003) report that 90% of the children in their study had already acquired the sounds of both Set 1 (/b, m, n, t/) and 2 (/d, f, p/) by the age of 3;0, while Arlt and Goodban (1976) and Prather et al. (1975) found that these sounds were acquired before or by 3;0 years in 75% of their participants. Both Set 1 and 2 sounds were regarded as attainable targets, but only Set 1 was treated. Set 3 sounds (/l, s, v/) were developed between the ages of 3;0 and 4;6 years in 75% of children sampled by Arlt and Goodban (1976) as well as Prather et al. (1975). Five Vs (/i, u, ɔ, ɛ, æ/) rated “1” remained consistent across sets.

Treatment stimuli. CVCV nonword stimuli, containing Set 1 Cs and Vs (see Supplemental File 3), were compiled using the SML software (Version 1.0, Van der Merwe, 2002). Only Variation Levels 1 and 2 were addressed due to time constraints. The software generated 103 nonwords (using Set 1 sounds) for Level 1 and 303 for Level 2. During each session a selection of series of nonwords, containing different target sounds in the initial position (as a way to systemize the selection), were treated. Nonwords varied across sessions. Depending on cooperation, 10 to 25 nonwords were practiced in each session.

Probe stimuli. Stimuli comprised untreated CVCV nonwords and real words. Stimuli of Sets 2 and 3 each included an untreated sound. Each set contained 10 nonwords

Table 2. Target sounds, selected on the basis of stimulability in isolation, accuracy, and age-appropriate norms, and the corresponding nonword and real word probe stimuli for the three sound sets

Set 1		**Set 2**		**Set 3**	
Target sounds: C: /b, m, n, t/ V: /i, u, ɔ, ɛ, æ/		*Target sounds:* C: /b, m n, t, d, f, p/ V: /i, u, ɔ, ɛ, æ/		*Target sounds:* C: / b, m, n, t, d, f, p, l, s, v/ V: /i, u, ɔ, ɛ, æ/	
Untreated nonwords	*Untreated real words*	*Untreated nonwords*	*Untreated real words*	*Untreated nonwords*	*Untreated real words*
/bibæ/	beanie	/bɔdu/	daddy	/bisu/	belly
/bubi/	Betty	/bifɛ/	“dudu”	/bɔvi/	loony
/bɔni/	Mannie	/dɔmɛ/	Daffy	/dilɔ/	Lulu
/mɛtæ/	many	/dɛni/	fatty	/dɔvɛ/	messy
/mɔtu/	nanny	/ditɔ/	meaty	/lubi/	movie
/næmɛ/	naughty	/fibɔ/	nappy	/lɛti/	Nelly
/nitɛ/	Tammy	/fumɛ/	penny	/sænɛ/	Sally
/tɛbɔ/		/fænu/	patty	/sɔtu/	seesaw
/tɔmu/		/pubæ/	teddy	/vimæ/	valley
/tænu/		/pætu/	tutu	/vɛnɔ/	

Note: Sounds were selected based on the pre-treatment stimulability and accuracy of production ratings. Target sound allocation was also guided by developmental norms for English speech sound acquisition (Arlt & Goodban, 1976; Dodd et al., 2003; Prather et al., 1975). Vowels remained constant across sets. Newly introduced consonants are underlined. Only Set 1 sounds were treated.
C = consonant, V = vowel.

representing all five levels of variation. Set 1 had seven real words, Set 2 had 10, and Set 3 had nine. Due to the limited number of sound combinations that could form meaningful words, the number of real words per set differed. When the available sounds limited meaningful combinations, proper names and word approximations (for example, “dudu” for “sleep” - a nursery word used by all South African speakers) were included if they were meaningful to the participant. For the same reason real words were not treated during the study. The probe stimuli are presented in Table 2.

Probe procedures

The second author conducted all probes, either at the participant’s home, or at his speech-language therapist’s rooms. In total, 14 probe sessions were conducted. First, three pre-treatment baseline probes (B1-B3) were performed on different days over two weeks. Repeated probes were performed once weekly across the treatment period of nine weeks (T1-T9 for Set 1, B4-B12 for Sets 2 and 3). Thereafter, two follow-up probes (F1-F2 for Set 1, B13-B14 for Sets 2 and 3) were performed, at one week and two weeks post-treatment. Treatment probes were administered on the same day of a week and no treatment was provided on those days. Each baseline and follow-up probe included all three sets. During the treatment period, Set 1 stimuli were probed at every probe session (T1 through T9). After the fourth and fifth baseline probe sessions, Sets 2 and 3 stimuli were probed at alternate probe sessions (B6, B8, B10, B12 and B7, B9, B11 respectively) to accommodate the participant’s attentional and motivational capacity.

Identical collection procedures were followed. All items were elicited once (imitation; visual and auditory model provided), with fixed instructions and no feedback. Nonwords were randomized within each set so that those beginning with the same C were not all presented successively. First, the nonword stimuli from Sets 1, 2, and 3 were presented verbally. Thereafter, the real words from each set were presented verbally with an accompanying picture. The order of presentation of nonwords remained the same for all probes, while that of real words occasionally varied to maintain the child’s cooperation. Various play activities were incorporated to ensure that the participant remained engaged. All productions were recorded on a digital voice recorder (Olympus DM-450) and transferred to a laptop computer (Acer Aspire 3002WLCi) for later analysis (described below).

Treatment procedures

Treatment was provided by the participant’s speech-language therapist during the set number of sessions. Treatment stimuli were supplied and her role was to follow the steps of the SML approach. The first author explained the steps to her and also gave her a written manual to study. The second author, who was trained across three years by the first author to apply SML treatment, provided coaching to the therapist during the first treatment session and observed all sessions. The second author also made video recordings of all sessions. Together, the authors reviewed recordings of sessions 2, 3 and 4. All treatment steps were found to be executed correctly, according to a check sheet. These steps were: demonstration of a nonword, imitation (three to five times), pre-response delay period imposed at 80% criterion, sequential imitation of a series, and fading KR feedback.

Outcome measure analysis and description

All probe stimuli were perceptually analysed by consensus of two experienced listeners (the first author and an independent experienced listener), after repeated listening. The order of presentation of probe sessions was randomized and the listeners were blinded to time points, but not to sound sets. The primary outcome measure was whole word accuracy (Newbold, Stackhouse, & Wells, 2013) as the aim was to determine if treated sounds could be produced and coarticulated accurately in a non- or real word. Whole word accuracy (WWA) was scored perceptually on a binary (correct or incorrect) scale. A word was judged as “correct” if all the component sounds were produced accurately, with no distortions, substitutions, omissions, additions, or trans-positioning and if all sounds were co-articulated well with no inappropriate pauses or any other prosodic inaccuracy.

Processing and analysis of the data

Raw scores of WWA were processed to determine percentage WWA, as the number of real words differed across sets. Effect sizes (ESs) for WWA were calculated to determine the relative degree of the treatment effects. ESs were calculated as follows: (mean of post-treatment scores minus mean of baseline scores) divided by the pooled standard deviation of Set 1 and 2 baseline data. Set 3 baseline scores (B1-B3) demonstrated zero variance, with which ESs cannot be calculated. To allow for direct comparison across Sets, the pooled standard deviation of Sets 1 and 2 baseline data was used as a consistent denominator for all ES calculations. As benchmarks for interpreting the magnitude of ESs have not been established for CAS treatment research, a significant improvement was operationally defined as an ES greater than 1.00 (as in Maas and Farinella, 2012).

To enhance analysis of treated Set 1 data, two criterion lines were generated following the conservative dual criterion (CDC) method (Fisher, Kelly, & Lomas, 2003). A linear regression line (trend) was generated from baseline data, using the ordinary least squares method, and a level line was generated from the mean of the baseline data. The height of these two lines was raised by 0.25 standard deviations (using baseline data), as this “represented a reasonable compromise between Type I and Type II errors” (Fisher et al., 2003, p. 392). The conservative baseline (B) trend ($\text{Trend}^{+}_{\text{B}}$) and mean ($\text{Mean}^{+}_{\text{B}}$) lines were superimposed on the treatment phase of Set 1 graphs (see Figure 2A and B). Additional conservative treatment (T) trend (Trend^{+}_{T}) and mean (Mean^{+}_{T}) lines were generated using treatment data and superimposed on the follow-up phase for Set 1 graphs (see Figure 2A and B). To determine whether a reliable treatment and retention effect was present, the number of data points which fell above the conservative trend and mean lines was compared according to the guidelines provided by Fisher et al. (2003, p. 399).

Post-hoc data analysis

Following the WWA analysis, the researchers were interested in the nature of segmental (sound) level speech errors and a possible change in the number of errors from pre- to post-treatment of Set 1. The type of errors may provide a view to the salient features of CAS, while a decrease in the number of errors may point to improved probe word production. A perceptual analysis was performed to augment WWA data. All stimuli which were scored as “incorrect” were further analysed by narrow phonetic transcription. Errors that were noted were classified into three groups: *First*, C distortion due to disruption in temporal or spatial features of a sound or distorted coarticulation with other sounds. Distortions include, for example, voicing errors i.e. sounds not clearly voiced or devoiced, weak stop, partial nasalization of a non-nasal sound and plosive elements added to a continuant. All these errors can be explained from a motor perspective. *Second*, perceived substitutions or distorted substitutions of Cs**,** as well as all V errors, were grouped together. Vowel errors could not always be classified as substitutions, distortions, or distorted substitutions. Errors in this group cannot be confidently explained as being either motor or linguistic in nature. *Third,* sound omissions were classified as a separate group as the frequency of occurrence took on a differentiable trend particularly in Set 3 words. A fourth group, *no score* (NS), was specified for isolated incidences when word production was not attempted (WWA was noted as incorrect) or when words were distorted to such an extent that sounds were not identifiable.

Items scored as NS (no score) were excluded from the error analysis. The total number of errors per error category across baseline (B1 to B3 total) and the follow-up period (F1 + F2 for Set 1; B13 + B14 for Sets 2 and 3) was counted separately for nonwords and words. The average number of errors per probe for each period was also calculated.

Reliability

Twenty per cent of recordings were re-analysed six to seven weeks after the initial analysis to determine the intra-reliability of WWA and error group analyses. Point-to-point agreement between the result of the initial analysis and the re-analysis for each item was determined. For WWA, reliability was 87%. For the post-hoc analysis, reliability for type of error was 92% for omissions, 80% for distortions, and 78% for substitutions and V errors.

Results

Whole word accuracy (WWA)

Treated Set 1 sounds. Baseline scores for Set 1 nonwords and words did not remain stable and displayed a rising trend. Effect size analyses showed a significant improvement in WWA for Set 1 nonwords (ES = 4.91) as well as real words (ES = 10.27) (see Figure 2 graphs A and B). Performance for both nonwords and real words remained variable. According to the CDC method (Fisher et al., 2003), if there are nine data points in the treatment phase, at least eight data points must be above both criterion trend (Trend^{+}_{B}) and mean (Mean^{+}_{B}) lines to confirm a reliable treatment effect. For Set 1 nonwords, four data points in the treatment phase are above both criterion lines (see Figure 2A), while five data points are above both lines for Set 1 real words (see Figure 2B). Based on this analysis the treatment effect is not reliable. For the follow-up phase, both data points are above the treatment criterion lines (Trend^{+}_{T} and Mean^{+}_{T}), suggesting a reliable retention effect.

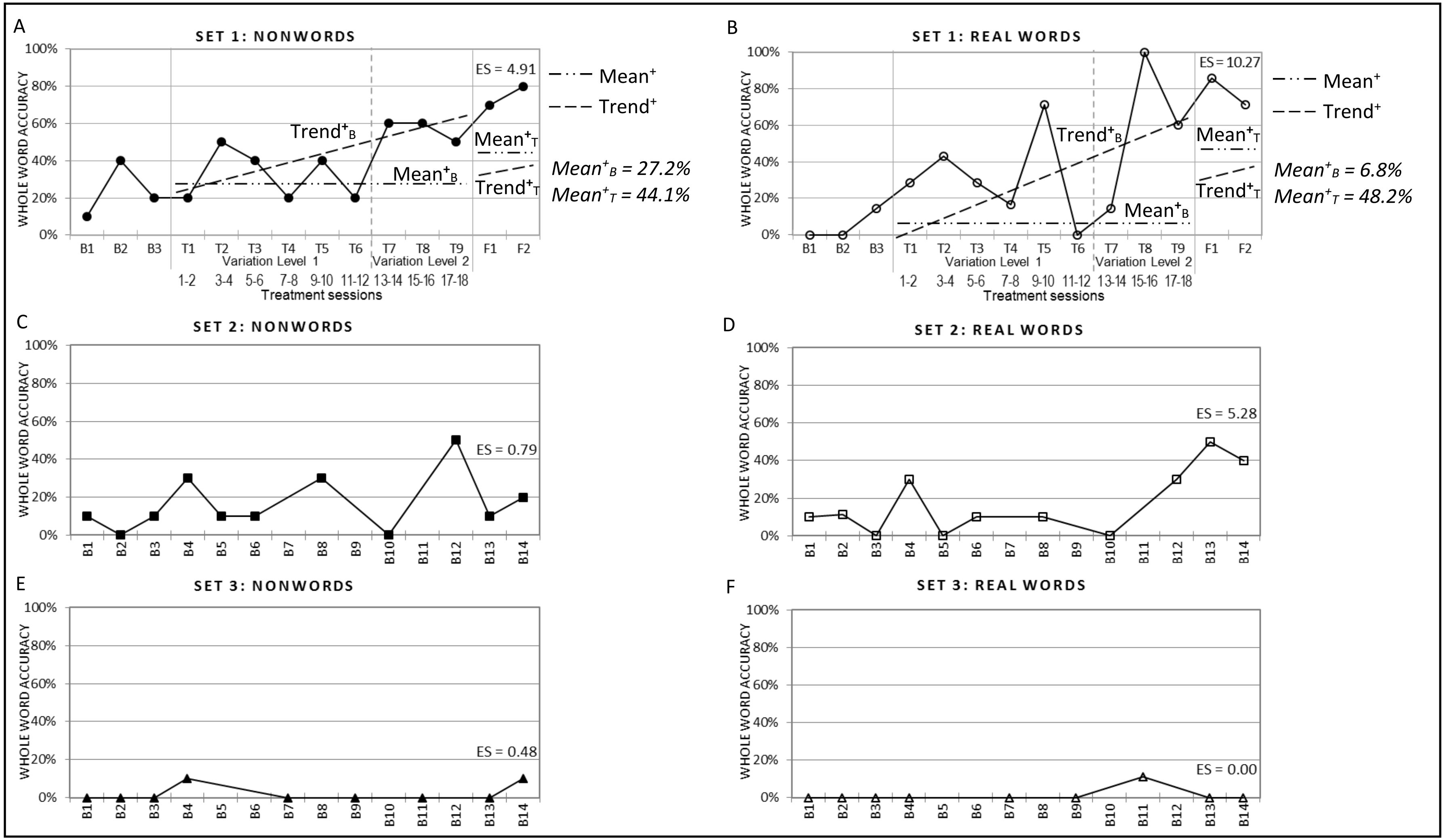

Figure 2. Whole word accuracy (% correct) for: Set 1 nonwords (A) and real words (B) containing treated age-appropriate sounds; Set 2 nonwords (C) and real words (D) containing an untreated age-appropriate consonant; and Set 3 nonwords (E) and real words (F) containing an untreated age-inappropriate consonant. Effect sizes (ES), conservative baseline trend (Trend$^{+}_{B}$) and mean (Mean$^{+}_{B}$) lines superimposed on the treatment phase of Set 1 graphs, and treatment trend (Trend$^{+}_{T}$) and mean (Mean$^{+}_{T}$) lines superimposed on the follow-up phase of Set 1 graphs, are also indicated.

Untreated Set 2 and 3 sounds. Set 2 non- and real word scores fluctuated, but evidenced slightly higher scores in WWA of real words towards the end of the study (Figure 2 graphs C and D). ES was significant for real words (ES = 5.28), but not for nonwords (ES = 0.79). No change in WWA was observed for Set 3 nonwords (ES = 0.48) and real words (ES = 0.00) (Figure 2 graphs E and F).

Post-hoc results: Number and nature of errors

Data from the post-hoc analysis are displayed in Table 3. The average number of errors per probe declined post-treatment (of Set 1) in all data sets except for Set 3 real words. For example, on Set 1 nonwords, an average of 11.3 errors per probe (across 10 nonwords) were made during B1 to B3, while an average of 4.0 errors were made during follow-up. A predominance of distortions (47, 90, 93 total errors across stimuli and period, for Set 1, 2 and 3, respectively) was observed, followed by substitutions and vowel errors (34, 51, 71). Both types of errors declined in frequency from pre-treatment to follow-up for all three data sets. Only a few omissions were recorded, with most incidences noted for Set 3 real words.

The results of the *no score* (NS) group were not included in the analysis of nature of errors, but are briefly described here. A total of 23 NSs were recorded over the duration of the study. Eighteen were recorded during baseline (nine due to no attempted production and nine due to whole word distortion), four during the treatment phase and one during follow-up. NSs were most often recorded for Set 3 real words, specifically the words ‘valley,’ ‘Lulu,’ ‘Sally,’ and ‘belly,’ which were completely distorted. The number of NS scores for Set 3 real words decreased across B1-B3 (n = 9), B4-B12 (n = 4) and B13-B14 (n = 1).

Table 3. Number of segmental level errors in the three error groups and average number of errors per probe during baseline (B1 - B3) and follow-up period (F1 + F2 for Set 1; B13 + B14 for Sets 2 and 3) for each of the three sound sets.

Stimuli	Period	Consonant Distortions	Substitutions and vowel errors	Omissions	Total number of errors and average per probe in brackets
Set 1 (Treated)					
Nonwords (N=10)	B1 to B3 total	17	17	0	34 (11.3)
	F1 + F2	*5*	*3*	*0*	*8 (4)*
Real words (N=7)	B1 to B3 total	23	13	1	37 (12.3)
	F1 + F2	*2*	*1*	*0*	*3 (1.5)*
Total across stimuli and periods		47	34	1	
Set 2 (Untreated)					
Nonwords (N=10)	B1 to B3 total	31	25	1	57 (19)
	B13 + B14	*13*	*14*	*1*	*28 (14)*
Real words (N=10)	B1 to B3 total	34	8	2	44 (14.7)
	B13 + B14	*12*	*4*	*0*	16 (8)
Total across stimuli and periods		90	51	4	
Set 3 (Untreated)					
Nonwords (N=10)	B1 to B3 total	39	24	1	64 (21.3)
	B13 + B14	*18*	*9*	*1*	*28 (14)*
Real words (N=9)	B1 to B3 total	20	17	1	38 (12.7)
	B13 + B14	*16*	*21*	*7*	*44 (22)*
Total across stimuli and periods		93	71	10	

Discussion

The effects of SML treatment (Van der Merwe, 1985, 2002, 2011) on the speech of a 33 month old child with CAS was explored across a nine week treatment period. Effect size (ES) analyses of whole word accuracy (WWA) showed a significant improvement of untreated non- and real words containing treated Set 1 sounds and untreated real words containing untreated Set 2 sounds. The CDC analyses (Fisher et al., 2003), however, indicated an unreliable treatment effect of Set 1 stimuli due to rising baseline scores. The perceptual analysis revealed a decline in number of sound level errors. This was true for all data sets except for Set 3 real words. In all data sets most errors were classified as distortions.

Treatment effects

Improved scores of Set 1 stimuli point to a change in behavior during the second (on nonwords) and third (on words) baseline probes. For experimental reasons this is unfortunate, since improved WWA accuracy across treatment cannot confidently be ascribed to the effect of treatment. Maturation may have caused improved baseline performance, but it could also have resulted from the foregoing assessment process. The three pre-baseline assessment sessions and the first baseline probe may have acted as “intervention” opportunities. Observation and then imitation of sounds in isolation and in non- and real words may have stimulated the mirror neuron system (Arbib, 2006; Kent, 2004) and initiated development or reinforcement of emerging internal models for guiding motor planning of sounds, particularly those of stimulable sounds. Mirror neurons respond to the actions of others and to self-generated actions (Wolpert et al., 2001). The acquisition process could have been supported by the human ability of “complex imitation” which presupposes the capacity for “complex action analysis.” The latter is the ability to analyze a model as a combination of actions and then to add new actions to the repertoire (Arbib, 2006, p. 22). Imitated production of a sound afforded the opportunity to experience tactile-kinesthetic feedback and perception of the auditory outcome. Conditions were favorable to build an inverse model (core motor plan) of the transformation from the “desired movements” to the “motor commands required to attain these movement goals” (Kawato & Gomi, 1992, p. 446). Response produced auditory feedback could also contribute towards development of a forward model for monitoring the efference copy for production of a sound (see discussion in Van der Merwe, 1997, 2009). Of significance is the fact that production of nonwords and words containing Set 1 sounds only improved as from the second and third baseline probe during which non- and real words were imitated. This may suggest that exposure to the sound in those contexts was a necessary prerequisite for improved production of probe stimuli. Imitation of the most stimulable sounds in isolation and in non- and real word probe stimuli could have reinforced these core motor plans which then acted as “building blocks” (a central concept in the FLF and SML approach), for motor planning of probe stimuli. Improved production of untreated probe stimuli during the subsequent treatment probes supports this explanation. During treatment, nonwords practiced across sessions differed and the child was continually confronted with the task of online planning of novel movement sequences, utilizing the target core motor plans.

Imitation and rehearsal of sounds in isolation as early treatment strategy for CAS has been recommended (for example by Hall et al., 2007) and tested experimentally. Early broadening of the phonetic inventory with sounds that are not stimulable is usually the focus (Iuzzini & Forrest, 2010; Powell, 1996). By teaching the child to be stimulable, Powell (1996) addressed absent aspects of the phonological system and stabilized emerging skills of his participant. Iuzzini and Forrest (2010) investigated a combined approach of stimulability training and a core vocabulary with complex phonological targets while treating four children between the ages 3,7 and 6,10. All children in their study demonstrated inventory expansion and increased percentage consonants correct (phonemic accuracy). Other approaches, like the DTTC integral stimulation approach (Strand et al., 2006), do not exclude practice of single sounds or syllables, but the main focus is on shaping movement gestures in longer utterances. Functional words and phrases are usually chosen as treatment targets in the DTTC. In the SML approach, production of sounds in isolation are assessed to determine if these core motor plans are yet acquired, but then CVCV nonwords containing the most stimulable sounds are utilized as early targets. Recent research in childhood speech disorders supports the generalisation potential of nonword treatment stimuli (Gierut, Morrisette, & Ziemer, 2010) and of less complex stimuli **(**Rvachew & Bernhardt, 2010)**.** The CV sequence is given a central role in the theory of evolution of speech. Prototypical babbling consists of a repeated rhythmic alternation between open and closed mouth configurations and reduplication of CV sequences (MacNeilage, Davis, Kinney, & Matyear, 2000), rendering the CVCV shape a logical choice. Stimuli of the SML approach presumably represent attainable, less complex, goals to the young child with CAS and provide a basis from which to work towards more difficult targets and an increasingly sophisticated speech motor system. Treatment methods, in other words "how children practice" are important, but treatment targets, or "what children practice", also need careful consideration (Maas, Gildersleeve-Neumann, Jackielski, & Stoeckel, 2014, p. 199).

Untreated Set 2 stimuli did not show sustained improvement during or shortly after baseline. Even though the target Cs were age-appropriate and the participant could achieve a score of three out of five correct in isolation across three assessments, performance remained variable with no significant change in nonword accuracy. Real words showed improvement during the last three baseline probes (B12-B14). Exposure to these words in his environment may have promoted learning. Another explanation is that knowledge was gained during treatment on how to achieve certain motor goals, for example, voicing contrast and tongue placement. Two of the Set 2 consonants (/d, p/) were voiced or voiceless cognates of Set 1 consonants (/t, b/). Furthermore, a stronger auditory model for real words may exist in sensorimotor memory which can guide auditory to articulatory inversion. Improved ability to plan this inversion successfully could have been an effect of treatment.

A significant result was the higher scores that were attained for Set 1 stimuli and Set 2 real words shortly after Set 1 Variation Level 2 stimuli were introduced in treatment (see T7 in Figure 2 graph A and B, and B12 in Figure 2 graph D). These nonwords (C1V1C2C+) require greater variation in articulatory features than Level 1 (C1V1C1C+) and presumably represent increased functional task difficulty and higher contextual interference in practice conditions (Guadagnoli & Lee, 2004; Magill, 2007). The number of targets (sounds) in a nonword increased from three to four, representing a greater amount of random practice. Also, coarticulatory challenges increased, and also varied across different treatment stimuli, contributing to higher variability in practice. Both conditions challenge motor planning and could foster speech motor learning. The challenge point theory proposes that learning can only occur when the learner is challenged, but not to the extent that learning is prohibited (Guadagnoli & Lee, 2004). Variation Level 2 treatment stimuli appeared to have posed appropriate challenges which fostered transfer and retention of acquired skills.

Post-hoc analysis: Number and nature of errors

The average number of segmental level errors declined from baseline to the follow-up period for all data sets except for Set 3 real words (see Table 3). Although WWA scores, which reflect accurate individual sound production and coarticulation between sounds, had not yet reached 100% correct, there were fewer errors per word or nonword. Fewer errors could by inference lead to improved probe word intelligibility. The decrease in average errors of Set 3 nonwords (21.3 to 14.0) in the presence of a WWA accuracy ES of 0.48 implies that accuracy of Sets 1 and 2 sounds (easier sounds) that appeared in these nonwords had improved while the Set 3 sound remained incorrect. Conversely, errors on Set 3 real words increased. Five of the nine words contained a medial /l/. This distribution was unintentional and was determined by the number of meaningful CVCV probe words which could be formed given the available sounds. Initially these words were noted as *no score* (NS) as the participant was unable to produce the /l/, and also unable to coarticulate the different sounds in the words, resulting in whole word distortion. During the follow-up period an increase in the number of omissions of medial /l/ was noted. His ability to separate the /l/ from the other sounds improved to the extent that he was able to produce the treated sounds in these words, while omitting the /l/. The increase in omissions therefore reflects attempts to produce Set 3 words more accurately. Omission errors are reportedly prominent in younger children with CAS and could be a strategy to lessen complexity (Hall et al., 2007).

The analysis of nature of errors indicated that the participant made more distortion errors than errors of the other types. Distortions mainly included incidents where sounds could not be judged as either clearly voiced or clearly voiceless. Voicing errors could be attributed to interarticulatory synchronization which was mistimed within milliseconds. Two other types of distortions that were noticed were occasional inappropriate nasalization (probably due to velar movement mistiming) and inadequate lip closure for plosive sounds (spatial disruption). Such errors could be the result of the underlying deficit in motor planning. Errors of the second group (substitutions and vowel errors) consisted mostly of vowel errors. Distortion and substitution of vowels were perceptually virtually indistinguishable. Errors judged as substitutions or distorted substitutions rarely occurred. No transpositioning or additions of phonemes were noted. The latter result does not support the observation of additions by Shriberg et al. (2012). Conclusions that are drawn regarding the underlying nature of surface speech errors depend on theoretical orientation. Within the theoretical framework of this study, distortions are regarded as phonetic-motoric in nature, while well-articulated substitutions, additions, and trans-positioning of phonemes may point to encoding process dysfunction. The complexity of speech development and the interface between motor and phonological development (Rvachew & Bernhardt, 2010) render a judgement on level of dysfunction speculative, but the high proportion of distortions does seem to support the notion that the underlying deficit in *idiopathic* CAS (Shriberg et al., 2012) is motor-based (ASHA, 2007; Hall et al., 2007).

Conclusions, limitations and future considerations

Collectively the results appear to suggest that application of the SML approach probably had positive effects on the speech of a 33 month old child with CAS. In view of a history of slow speech development it is unlikely that the observed gains were attributable to maturation alone. Improving performance during baseline was likely due to pre-treatment assessment procedures acting as intervention. The decline in number of errors per nonword/word as judged perceptually points to improved production of Set 1 treated sounds as they occurred in nonwords and words of Sets 1, 2, and 3. Fewer errors per word could enhance word intelligibility. The results suggest gradually improving motor planning skills.

Only one young child participated in the present study. Future research should include more participants, including a wider range of age groups. Higher levels of evidence should be pursued, particularly cohort, case-control studies and randomised control trials. However, this study contributes to the growing literature on the treatment of CAS and proposes SML treatment (Van der Merwe, 2002, 2011) as a promising approach.

Acknowledgements

Financial assistance of the National Research Foundation (NRF) of South Africa towards the Master’s studies of Mollie Steyn is acknowledged. Opinions expressed and conclusions presented are not to be attributed to the NRF. The authors declare no conflicts of interest. The authors wish to thank the participant, his parents, and speech-language pathologist, and to express appreciation to Dr Diane Kendall for her assistance during transcription and processing of the data.

References

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

Anderson, D., Fowler, S., & Nelson, J. (1978). Developmental assessment schema. In W. J. Northcott (Ed.), *Curriculum guide: Hearing impaired children (0-3 years) and their parents*. Washington DC, MD: Alexander Graham Bell Association for the Deaf.

Arbib, M. A. (2006). The mirror system hypothesis on the linkage of action and languages. In M. A. Arbib (Ed.), *Action to language via the mirror neuron system* (pp. 3-47). Cambridge: Cambridge University Press.

Arlt, P. B., & Goodban, M. T. (1976). A comparative study of articulation acquisition as based on a study of 240 normals, aged three to six. *Language, Speech, and Hearing Services in Schools, 7,* 173-180. doi:10.1044/0161-1461.0703.173

Ballard, K. J., Robin, D. A., McCabe, P., & McDonald, J. (2010). A treatment for dysprosody in childhood apraxia of speech. *Journal of Speech, Language & Hearing Research, 53*(5), 1227-1245. doi:10.1044/1092-4388(2010/09-0130)

Barlow, D. H., Nock, M. K., & Hersen, M. (2009). *Single case experimental designs: Strategies for studying behavior change* (3rd ed.). Boston, MA: Pearson Education.

Bislick, L. P., Weir, P. C., Spencer, K., Kendall, D., & Yorkston, K. M. (2012). Do principles of motor learning enhance retention and transfer of speech skills? A systematic review. *Aphasiology, 26*(5), 709-728. doi:10.1080/02687038.2012.676888

Dodd, B., Holm, A., Hua, Z., & Crosbie, S. (2003). Phonological development: A normative study of British English-speaking children. *Clinical Linguistics & Phonetics, 17*(8), 617-643. doi: 10.1080/0269920031000111348

Edeal, D. M., & Gildersleeve-Neumann, C. (2011). The importance of production frequency in therapy for childhood apraxia of speech. *American Journal of Speech-Language*

Pathology, 20(2), 95-110. doi:10.1044/1058-0360(2011/09-0005)

Edwards, M. L., & Shriberg, L. D. (1983). *Phonology: Applications in communicative disorders*. San Diego, CA: College-Hill Press Inc.

Fisher, W. W., Kelley, M. E., & Lomas, J. E. (2003). Visual aids and structured criteria for improving visual inspection and interpretation of single-case designs. *Journal of Applied Behavior Analysis, 36*(3), 387-406. doi:10.1901/jaba.2003.36-387

Gierut, J. A., Morrisette, M. L., & Ziemer, S. M. (2010). Nonwords and generalization in children with phonological disorders. *American Journal of Speech-Language Pathology, 19*(2), 167-177. doi:10.1044/1058-0360(2009/09-0020)

Guadagnoli, M. A., & Lee, T. D. (2004). Challenge point: A framework for conceptualizing the effects of various practice conditions in motor learning. *Journal of Motor Behaviour, 36*(2), 212-224. doi:10.3200/jmbr.36.2.212-224

Guenther, F. H., & Perkell, J. S. (2004). A neural model of speech production and its application to studies of the role of auditory feedback in speech. In B. Maassen, R. Kent, H. Peters, P. Van Lieshout, & W. Hulstjin (Eds.), *Speech motor control in normal and disordered speech* (pp. 29-49). Oxford: Oxford University Press.

Hall, P. K., Jordan, L. S., & Robin, D. A. (2007). *Developmental apraxia of speech: Theory and clinical practice* (2nd ed.). Austin, TX: Pro-Ed.

Hickok, G. (2014). The architecture of speech production and the role of the phoneme in speech processing. *Language, Cognition and Neuroscience, 29*(1), 2-20. doi:10.1080/01690965.2013.834370

Hickok, G., Houde, J., & Rong, F. (2011). Sensorimotor integration in speech processing: computational basis and neural organization. *Neuron*, 69, 407-422. doi:10.1016/j.neuron.2011.01.019

Hula, S. N. A., Robin, D. A., Maas, E., Ballard, K. J., & Schmidt, R. A. (2008). Effects of feedback frequency and timing on acquisition, retention, and transfer of speech skills in acquired apraxia of speech. *Journal of Speech, Language, and Hearing Research*, *51*, 1088–1113. doi:10.1044/1092-4388(2008/06-0042)

Iuzzini, J., & Forrest, K. (2010). Evaluation of a combined treatment approach for childhood apraxia of speech. *Clinical Linguistics & Phonetics*, *24*(4-5), 335-345. doi:10.3109/02699200903581083

Iuzzini-Seigel, J., Hogan, T. P., Guarino, A. J., & Green, J. R. (2015). Reliance on auditory feedback in children with childhood apraxia of speech. *Journal of Communication Disorders, 54*, 32-42. doi:10.1016/j.jcomdis.2015.01.002

Kawato, M. (1999). Internal models for motor control and trajectory planning. *Current Opinion in Neurobiology*, *9*(6), 718-727. doi:10.1016/s0959-4388(99)00028-8

Kawato, M. & Gomi, H. (1992). The cerebellum and VOR/OKR learning models. *Trends in Neuroscience*, *15*, 445-453.

Kent, R. (2004). Models of speech motor control: Implications from recent developments in neurophysiological and neurobehavioral science. In B. Maassen, R. Kent, H. Peters, P. Van Lieshout, & W. Hulstjin (Eds.), *Speech motor control in normal and disordered speech* (pp. 3-28). Oxford: Oxford University Press

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(1), 1-75. doi:10.3115/992628.992631

Maas, E., Butalla, C. E., & Farinella, K. A. (2012). Feedback frequency in treatment for childhood apraxia of speech. *American Journal of Speech-Language Pathology, 21*(3), 239-257. doi:10.1044/1058-0360(2012/11-0119)

Maas, E., & Farinella, K. A. (2012). Random versus blocked practice in treatment for childhood apraxia of speech. *Journal of Speech, Language, and Hearing Research, 55*(2), 561-578. doi:10.1044/1092-4388(2011/11-0120)

Maas, E., Gildersleeve-Neumann, C., Jakielski, K., & Stoeckel, R. (2014). Motor-based intervention protocols in treatment of childhood apraxia of speech (CAS). *Current Developmental Disorders Reports, 1*(3), 197-206. doi:10.1007/s40474-014-0016-4

Maas, E., Robin, D. A., Hula, S., Freedman, S. E., Wulf, G., Ballard, K. J., & Schmidt, R. A. (2008). Principles of motor learning in treatment of motor speech disorders. *American Journal of Speech-Language Pathology, 17*(3), 277-298. doi:10.1044/1058-0360(2008/025)

Maassen, B., Nijland, L., & Terband, H. (2010). Developmental models of childhood apraxia of speech. In B. Maassen, & P. Van Lieshout (Eds.), *Speech motor control: New developments in basic and applied research* (pp. 243-258). New York, NY: Oxford University Press.

MacNeilage, P. F., Davis, B. L., Kinney, A., & Matyear, C. L. (2000). The motor core of speech: A comparison of serial organization patterns in infants and languages. *Child Development, 71(1)*, 153-163. doi:10.1111/1467-8624.00129

Magill, R. A. (2007). *Motor learning and control: Concepts and applications* (8th ed.). Boston, MA: McGraw-Hill.

Manolson, A. (1992). *It takes two to talk.* Toronto: Hanen Centre.

Murray, E., McCabe, P., & Ballard, K. J. (2015). A randomized controlled trial for children with childhood apraxia of speech comparing Rapid Syllable Transition treatment and the Nuffield Dyspraxia Programme-Third Edition. *Journal of Speech, Language, and Hearing Research, 58*, 669-686. doi:10.1044/2015_JSLHR-S-13-0179

New, A. B., Robin, D. A., Parkinson, A. L., Duffy, J. R., McNiel, M. R., Piguet, O., … Ballard, K. J. (2015). Altered resting-state network connectivity in stroke patients with and without apraxia of speech. *NeuroImage: Clinical, 8*(2015), 429-439. doi:10.1016/j.nicl.2015.03.013

Newbold, E. J., Stackhouse, J., & Wells, B. (2013). Tracking change in children with severe and persisting speech difficulties. *Clinical Linguistics & Phonetics, 27*(6-7), 521-539. doi: 10.3109/02699206.2013.790479

Nijland, L., Terband, H., & Maassen, B. (2015). Cognitive functions in childhood apraxia of speech. *Journal of Speech, Language and Hearing Research, 58*, 550-565. doi:10.1044/2015_JSLHR-S-14-0084

Powell, T. W. (1996). Stimulability considerations in the phonological treatment of a child with a persistent disorder of speech-sound production. *Journal of Communication Disorders, 29*, 315-333. doi:10.1016/0021-9924(96)00015-9

Prather, E., Hedrick, D., & Kern, C. (1975). Articulation development in children aged two to four years. *Journal of Speech and Hearing Disorders, 40*, 179-19. doi:10.1044/jshd.4002.179

Romo, R., & Schultz, W. (1992). Role of primate basal ganglia and frontal cortex in the internal generation of movements. *Experimental Brain Research, 91*(3), 396-407. doi:10.1007/bf00227834

Rvachew, S., & Bernhardt, B. M. (2010). Clinical implications of dynamic systems theory for phonological development. *American Journal of Speech-Language Pathology*, *19,* 34-50. doi: 10.1044/1058-0360(2009/08-0047)

Renfrew, C. (2010). The Renfrew language scales: Revised edition. London: Speechmark Publishing Company.

Rossetti, L. R. (2006). *The Rossetti infant-toddler language scale: A measure of communication and interaction*. East Moline, IL: Lingui Systems.

Schmidt, R. A. (1975). A schema theory of discrete motor skill learning. *Psychological Review*, *82*(4), 225-260. doi:10.1037/h0076770

Schmidt, R. A., & Lee, T. D. (2005). *Motor control and learning: A behavioral emphasis* (4th ed.). Champaign, IL: Human Kinetics.

Shriberg, L. D., Lohmeier, H. L., Strand, E. A., & Jakielski, K. J. (2012). Encoding, memory, and transcoding deficits in childhood apraxia of speech. *Clinical Linguistics & Phonetics, 26*(5), 445-482. doi:10.3109/02699206.2012.655841

Shriberg, L. D., Potter, N. L., & 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. doi:10.1044/1092-4388(2010/10-0068)

Skelton, S. L., & Hagopian, A. L. (2014). Using randomized variable practice in the treatment of childhood apraxia of speech. *American Journal of Speech-Language Pathology, 23*(4), 599-611. doi:10.1044/2014_AJSLP-12-0169

Strand, E. A., Stoeckel, R., & Baas, B. (2006). Treatment of severe childhood apraxia of speech: A treatment efficacy study. *Journal of Medical Speech-Language Pathology, 14*(4), 297-307.

Terband, H., & Maassen, B. (2010). Speech motor development in childhood apraxia of speech (CAS): Generating testable hypotheses by neurocomputational modeling. *Folia Phoniatrica et Logopaedica, 62,* 134-142. doi:10.1159/000287212

Terband, H., Maassen, B., Guenther, F. H., & Brumberg, J. (2009). Computational neural modeling of childhood apraxia of speech (CAS). *Journal of Speech, Language and Hearing Research, 52*(6), 1595-1609. doi:10.1044/1092-4388(2009/07-0283)

Van der Merwe, A. (1985). *Treatment program for developmental apraxia of speech and other speech disorders [title translated]*. Pretoria: University of Pretoria Press.

Van der Merwe, A. (1997). A theoretical framework for the characterization of pathological speech sensorimotor control. In M. R. McNeil (Ed.), *Clinical management of sensorimotor speech disorders* (pp. 1-25). New York, NY: Thieme.

Van der Merwe, A. (2002). Speech motor learning (SML) approach software (Version 1.0) [Computer software]. Retrieved from http://www.apraxia-anitavandermerwe.co.za.

Van der Merwe, A. (2007). Self-correction in apraxia of speech: The effect of treatment. *Aphasiology, 21*(6-8), 658-669. doi:10.1080/02687030701192174

Van der Merwe, A. (2009). A theoretical framework for the characterization of pathological speech sensorimotor control. In M. R. McNeil (Ed.), *Clinical management of sensorimotor speech disorders* (2nd ed., pp. 3-18). New York, NY: Thieme.

Van der Merwe, A. (2011). A speech motor learning approach to treating apraxia of speech: Rationale and effects of intervention with an adult with acquired apraxia of speech. *Aphasiology, 25*(10), 1174-1206. doi:10.1080/02687038.2011.582246

Williams, P., & Stephens, H. (2004). *Nuffield Dyspraxia Programme* (3rd ed.). Windsor, England: The Miracle Factory.

Wolpert, D. M., Ghahramani, Z., & Flanagan, J. R. (2001). Perspectives and problems in motor learning. *Trends in Cognitive Sciences, 5*(11), 487-494. doi:10.1016/s1364-6613(00)01773-3

Wulf, G., & Schmidt, R. A. (1997). Variability in practice and implicit motor learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*(4), 987-1006. doi:10.1037/0278-7393.23.4.987

Ziegler, W. (2009). Modelling the architecture of phonetic plans: Evidence from apraxia of speech. *Language and Cognitive Processes, 24*, 631-661. doi:10.1080/01690960802327989