Journal of Communication Disorders 54 (2015) 32–42

Contents lists available at ScienceDirect

Journal of Communication Disorders

Reliance on auditory feedback in children with childhood apraxia of speech

Jenya Iuzzini-Seigel, Tiffany P. Hogan, Anthony J. Guarino, Jordan R. Green *

MGH Institute of Health Professions, United States

ARTICLE INFO

Article history:
Received 6 February 2014
Received in revised form 7 November 2014
Accepted 8 January 2015
Available online 19 January 2015

Keywords:
Childhood apraxia of speech
Feedforward programs
Auditory feedback
Voice onset time
Vowels
Acoustical analyses

ABSTRACT

Children with childhood apraxia of speech (CAS) have been hypothesized to continuously monitor their speech through auditory feedback to minimize speech errors. We used an auditory masking paradigm to determine the effect of attenuating auditory feedback on speech in 30 children: 9 with CAS, 10 with speech delay, and 11 with typical development. The masking only affected the speech of children with CAS as measured by voice onset time and vowel space area. These findings provide preliminary support for greater reliance on auditory feedback among children with CAS.

Learning outcomes: Readers of this article should be able to (i) describe the motivation for investigating the role of auditory feedback in children with CAS; (ii) report the effects of feedback attenuation on speech production in children with CAS, speech delay, and typical development, and (iii) understand how the current findings may support a feedforward program deficit in children with CAS.

1. Introduction

Childhood apraxia of speech is a pediatric speech disorder characterized by a broad range of features that includes inconsistent speech sound production (ASHA, 2007; Iuzzini, 2012), voicing errors (Iuzzini, 2012; Lewis, Freebairn, Hansen, Iyengar, & Taylor, 2004), disrupted prosody (Shriberg, Aram, & Kwiatkowski, 1997), and disrupted coarticulatory transitions (Maassen, Nijland, & Van Der Meulen, 2001; Nijland et al., 2002). The neurologic basis of these speech difficulties is unknown, but a recent proposal implicates an impairment of the feedforward motor control system (Terband & Maassen, 2010). Feedforward programs putatively contain articulatory motor commands for acoustic (Guenther, 2006) or vocal tract goals (Browman & Goldstein, 1989), and it is posited that weak feedforward programs result in the imprecise and inconsistent speech errors that are associated with CAS (Terband & Maassen, 2010).

Children's dependence on sensory feedback is purported to decline as they master motor skills (Forssberg, Eliasson, Kinoshita, Johansson, & Westling, 1991; Guenther, 2006; Haas, Diener, Rapp, & Dichgans, 1989). During the early stages of speech development, for example, auditory and somatosensory feedback is used to establish robust neural programs that encode the acoustic consequences of articulator movements (Guenther, 2006). Computational models of speech have simulated how that talkers refine these feedforward programs over time by resolving discrepancies between intended speech goals and actual speech output (Guenther, 2006; Haruno, Wolpert, & Kawato, 2001; Houde & Nagarajan, 2011). Once

* Corresponding author at: Speech & Feeding Disorders Lab, MGH Institute of Health Professions, 36 1st Avenue, Boston, MA 02129, United States. Tel.: +1 617 945 3624.
E-mail address: jgreen2@mghihp.edu (J.R. Green).

http://dx.doi.org/10.1016/j.jcomdis.2015.01.002

these feedforward programs are established, sensory feedback is used only intermittently for (1) making online corrections in speech and (2) adjusting articulator movements in response to changing speaking conditions (Houde & Jordan, 1998; Jones & Munhall, 2000; Lane & Tranel, 1971). This shift from feedback dependency to feedforward-dominant control presumably permits talkers to efficiently produce speech at an average rapid rate of 160 words per minute (Picheny, Durlach, & Braida, 1986).

In contrast to typically developing children, the reliance on auditory feedback control may persist into later stages of development in children with CAS, a pattern that could be an adaptive response to weak feedforward programs. Terband and Maassen (2009, 2010) proposed that, in children with CAS, the formation of robust feedforward programs may be hindered by excessive neural noise or by reduced oral somatosensation. Poorly formed feedforward programs would provide a parsimonious explanation for many features that are associated with CAS such as speech sound inconsistency, and vowel and voicing distortions (ASHA, 2007). The poor treatment outcomes and limited generalization exhibited by children with CAS (e.g., Forrest, 2003) could also be explained by a difficulty establishing robust feedforward programs. In contrast, it is unlikely that a feedforward control deficit would underlie speech impairment in children with speech delay (SD), whose errors are characterized by consistent substitutions and the use of phonological processes, and who typically show evidence of learning and generalization following treatment (e.g., Gierut, 1998).

In the absence of robust feedforward programs, children with CAS may adapt by continuously monitoring their speech to preempt, minimize, or repair speech errors (Terband, Maassen, Guenther, & Brumberg, 2009; Terband & Maassen, 2010). The continuous reliance on auditory feedback is, however, inefficient and is predicted to have several negative consequences on speech. For example, the online processing of one's speech is purported to take up to 100 ms (Houde & Nagarajan, 2011), which is probably too long to allow for rapid and accurate speech. Moreover, feedback has been evoked as a causal mechanism for stuttering. Disfluencies have been attributed to an over-reliance on feedback among stutterers, and interventions that delay or attenuate feedback often increase fluency in this population (e.g., Van Borsel, Reunes, & Van den Bergh, 2003).

Lastly, Terband, van Brenk, & van Doornik-van der Zee (2014) investigated effects of real-time formant-frequency perturbation in 17 typically developing children, and 11 children with speech sound disorders including 5 with CAS. Previous studies showed that most talkers produce a compensatory adaptation to the perturbation (e.g., Cai, Ghosh, Gunether, & Perkell, 2010; Houde & Jordan, 2002). Terband et al. found that typically developing controls tended to compensate for the perturbation where children with speech sound disorders tended to follow and exaggerate the frequency shift. These authors suggested that children with speech sound disorders perceived the formant-shift but did not compensate appropriately, which they attributed to impaired feedforward and feedback models.

In this study, we test the integrity of speech feedforward programs in children with CAS using an auditory masking paradigm. In this paradigm, calibrated levels of noise are used to disrupt the talkers' access to auditory feedback. It is hypothesized that well-established programs are relatively resistant to the perturbing effects of masking noise whereas weak programs will become disrupted as indicated by acoustic changes to speech. To test the weak feedforward program hypothesis, we addressed the following experimental questions: (1) Do school-aged children with CAS evidence comparable VOTs, vowel durations, and vowel space areas relative to those with SD and typical development (TD) in unmasked and masked speech, and (2) Do children with CAS, SD, and TD evidence an effect of noise-masking on VOT, vowel durations, vowel space areas, and speech intensity? Whereas previous research (Iuzzini & Forrest, 2008; Iuzzini, 2012) showed that preschool-aged children with CAS evidenced shorter VOTs and smaller vowel space areas relative to those with TD and SD, it is unknown whether school-aged children with CAS will also perform differently than their peers on these measures. It is posited that children with SD and TD will have intact feedforward control systems and will show an adaptive response to noise masking. In contrast, we hypothesize that children with CAS have a feedforward control deficit and therefore, will not adapt to noise masking.

VOT is a speech target of interest because Iuzzini (2012) showed that children with CAS are delayed in acquisition of the voicing contrast. Specifically, compared to children with SD, children with CAS produced shorter VOTs for voiceless plosives (e.g., /t/), which often overlapped with the VOTs for voiced cognates (e.g., /d/). The overlapping voicing categories may explain why children with CAS are often perceived to produce voicing errors or voicing distortions (Lewis et al., 2004). In contrast, typically developing children produce adult-like voicing categories by 3 years of age (Macken & Barton, 1980). Based on these findings, we predicted that school-aged children with CAS would have a vulnerability for this contrast, and therefore might evidence a regression to shorter VOTs in the presence of a perturbation—in this case, masking.

Examining masking effects on vowel production was also of interest because children with CAS have difficulty producing clear vowel contrasts (Lewis et al., 2004), which may reflect weak feedforward programs for these phonemes. Previous research by Iuzzini and Forrest (2008) showed that preschool-aged children with CAS evidenced smaller vowel space area relative to age-matched children with SD and TD, consistent with the percept of vowel neutralization for children with CAS. We posit that children with CAS will evidence smaller vowel space areas relative to those with SD and TD, and that these differences will be larger in the masked speech condition.

The current research also tests the effect of masking on vowel durations and speech intensity of vowels. Previous research showed that masking can induce an increase in speech intensity, which can in-turn lengthen vowel durations and alter vowel formant frequencies (Maas, Mailend, & Guenther, 2013; Rogers, Eyraud, Strand, & Storkel, 1996; Van Summers, Pisoni, Bernacki, Pedlow, & Stokes, 1988). This phenomenon, known as the Lombard effect, is considered an adaptive response for maximizing speech intelligibility in the presence of noise. We hypothesize that children with CAS may have difficulty making adaptive changes to speech in response to masking due to impaired control over segmental and suprasegmental aspects of speech.

2. Methods

2.1. Participants

A total of 30 children, between the ages of 6;1 and 17;6 (years; months), $M = 10;2$, were recruited for this study. Participants were recruited from multiple sources including existing research databases, speech-language pathologist referrals, and flyers distributed to the local public school system. Each child's parent provided informed consent statement and completed an intake form that provided demographic information. All subjects listed English as their primary language.

2.2. Group assignment

Each participant was assessed over the course of four, two-hour sessions. All participants passed a bilateral pure-tone hearing screening at 20 dB for the octave frequencies between 500 and 4000 Hz (ASHA, 1997). The evaluation included standardized and customized speech, language, and cognitive assessments. See Table 1 for a summary of demographics and test scores by group. All participants were required to have normal cognition based on a standard score of 75 or higher on the *Reynolds Intellectual Assessment Scales* (*RIAS*; Reynolds & Kamphaus, 2003). Normal language, as indicated by a standard score of 85 or higher on the *Clinical Evaluation of Language Fundamentals-Fourth Edition* (*CELF*; Semel, Wiig, & Secord, 2003), was required for participation in all groups.

Participants were assigned to the typically developing group (TD; $n = 11$) if they met the criteria listed above, scored higher than the 16th percentile on the *Goldman Fristoe Test of Articulation-2nn Edition* (*GFTA-2*; Goldman & Fristoe, 2000), and reported no history of speech or language treatment; this last criterion was specified to exclude children with atypical speech and language development.

All children with speech disorders had a history of speech treatment. Of these, 8 were referred with a CAS diagnosis by a speech-language pathologist with expertise in CAS. An additional child was reported to have a CAS diagnosis per parent report. Although these children were referred with a CAS diagnosis, in order to be included in the CAS group in this experiment, participants were required to evidence at least 5 out of 11 CAS characteristics (adapted from Shriberg, Potter, & Strand, 2011) during completion of the *GFTA-2*. The CAS characteristics assessed included vowel distortions, difficulty achieving initial articulatory configurations or transitionary movement gestures, equal stress or lexical stress errors, distorted substitutions, syllable segregation, groping, intrusive schwa, voicing errors, slow rate, increased difficulty with multisyllabic words, and disturbed resonance. Operational definitions were created for these characteristics and two CAS experts blind-rated each child's *GFTA-2* productions. Appendix A presents the operational definitions that were used for feature ratings.

Children were assigned to the SD group ($n = 10$) if they produced 4 or fewer CAS characteristics. In addition, children were excluded from the SD group if they had a history of treatment for CAS; this criterion was included because the profile of CAS is known to change with treatment and we did not want to assign children with partially resolved CAS to the SD group.

Although the cutoff for assignment to the CAS group was 5 or more features, children who were assigned to the CAS group exhibited 8 features on average (standard deviation = 3) and children assigned to the SD group had only 3 features (standard deviation = 1). One of the common pitfalls of CAS research is circularity between factors that contribute to group assignment and those evaluated as dependent variables. The current study required that participants evidenced a minimum of 5 features (out of 11) to be assigned to the CAS group, but none of these features were considered mandatory. This was done to avoid potential circularity between the group assignment criteria 'vowel distortions' and 'voicing errors' and our dependent measures 'vowel space area' and 'voice onset time.'

Inconsistency of speech sound errors is commonly associated with CAS, however, it has been shown to decrease following speech sound treatment (Iuzzini & Forrest, 2010). In addition, a standardized, norm-based procedure for assessing inconsistency in school-aged children with CAS and SD does not exist and therefore, specific cutoffs for determining if a child exhibits inconsistency commensurate with CAS versus SD are unavailable. Additionally, it is unknown whether inconsistency is differentially diagnostic in older children with CAS who have a history of speech treatment. Although the ASHA technical statement on CAS (2007) reported that inconsistency of speech errors had gained consensus in the field as

Table 1
Participant characteristics by group.

Measure	TD ($n = 11$)	CAS ($n = 9$)	SD ($n = 10$)
Age in months	134 (46)	126 (42)	102 (14)
GFTA-2 percentile	30 (13)	3 (3)	4 (3)
CELF-4 core SS	111 (12)	109 (13)	106 (15)
RIAS SS	111 (17)	112 (10)	113 (17)
# CAS features[a]	1 (1)	8 (2)	3 (1)

Note. Standard deviations are in parentheses. *GFTA-2* = Goldman Fristoe Test of Articulation-2nd Edition (Goldman & Fristoe, 2000); *CELF-4* = Clinical Evaluation of Language Fundamentals-4th Edition (Semel et al., 2003); RIAS = Reynolds Intellectual Assessment Scales (Reynolds & Kamphaus, 2003); SS = standard score.

[a] Feature list and assessment procedure adapted from Shriberg et al. (2011).

a diagnostic criterion to differentiate CAS and SD, the statement also reported that this was not yet a validated criterion, and that it was neither necessary nor sufficient for a CAS diagnosis (ASHA, 2007, pp. 9–10). It should be noted that although many children in our CAS group did evidence inconsistency of sounds and words across multiple repetitions on responses from a customized speech battery not discussed in the current report, this feature did not contribute to group assignment.

There were no significant group differences on (a) age, (b) nonverbal intelligence, or (c) language. As expected, the groups differed on *GFTA-2* percentile score, a measure of speech severity, $F(2,27) = 37.904$, $p < .001$. The CAS and SD groups were equivalent on this measure, $p = .958$, and both were significantly lower than the TD group, $p < .001$. A summary of participant characteristics is included in Table 1.

2.3. Stimuli

The stimuli items of interest for the current experiment were CVC pseudowords, /pɑb/, /pæb/, /pib/, /pub/, which sampled the corner vowels. Pseudowords were elicited within the carrier phrase "Say ▬▬▬ now" in the context of a larger pseudoword repetition task that contained eight CVC /pVb/ pseudowords. Words were produced five times each in random order, in each condition. All pseudowords were recorded by a male with a Midwestern dialect.

2.4. Procedures

Participants' acoustic speech productions were digitally recorded at a sampling rate of 44.1 kHz using a stationary head-mounted microphone (Shure, BETA 54) and solid-state audio recorder (Marantz, PMD660). Headphones were used to bilaterally present pseudoword stimuli and masking noise.

Two versions of the protocol were used to counterbalance order effects of condition (masking). For the unmasked condition, each item was presented to the participant through headphones, using the previously stated carrier phrase. Once the stimulus item was presented, the participant repeated it in a normal speaking voice. This procedure was repeated until all stimuli had been presented to the participant, five times each, in random order, for a total of 10 productions of our stimuli of interest.

For the masked condition, white-noise was presented at 65 dB following the presentation of each stimulus item, such that the participant's response was produced concurrently with the masking noise. This noise level is quieter than the 90+ dB typically used in masking studies with adults, and was selected to reduce the risk of inducing a temporary or permanent hearing loss in our young participants. As described in Section 1, masking was used to attenuate auditory feedback during this condition. The masking noise level was calibrated prior to each testing session.

Following the initial assessment, the audio files were analyzed using the acoustic analysis software, PRAAT (Boersma & Weenink, 2013). Each nonword trial was identified in the full length audio string, and labeled (i.e., annotated in a text grid). Waveforms and broadband spectrograms were used to measure VOT, vowel durations, and formant frequencies. Praat software was also used to measure relative speech intensity (dB) of masked and unmasked productions of the vowel /ɑ/ from the stimulus item /pɑb/.

VOT was measured as the duration between the release of the stop burst and the onset of glottal pulsing for the vowel (Lisker & Abramson, 1964). Mean VOTs were calculated for each participant, for each condition, and subjected to statistical comparisons. In addition, VOT durations were dichotomized as either correct (1) or incorrect (0), based on a cutoff of 35 ms (Lisker & Abramson, 1964, 1967). VOTs for /p/ that are shorter than 35 ms constitute either a phonetic or phonemic error, and would therefore be perceived as a distortion or as a /b/, respectively (Blumstein, Cooper, Goodglass, Statlender, & Gottlieb, 1980). Percentage of accurate responses was recorded for each participant for each condition.

Vowel durations were measured as the duration between the onset of glottal pulsing for the vowel /ɑ/ and the offset of periodicity of the waveform, which converged with the final striation of the first and second formants on the spectrogram. Mean vowel durations were calculated for trials from each condition.

The effect of speech intensity on masking were limited to within-group comparisons because audio recordings were not calibrated across sessions.

The first two formants (Hz) for each of the four corner vowels were tracked using Praat. As such, the average frequency for each formant was calculated across the entire duration of each vowel. Hand-editing was used where needed to confirm accuracy of tracking. Means were calculated for the first and second formants for each vowel, across the five trials, for each condition. Vowel space area was calculated using the following formula (Vorperian & Kent, 2007):

$$\text{Vowel space area } (\text{Hz}^2) = .5 \times [(/\text{i}/\ \text{F2} \times /\text{æ}/\ \text{F1} + /\text{æ}/\ \text{F2} \times /\text{ɑ}/\ \text{F1} + /\text{ɑ}/\ \text{F2} \times /\text{u}/\ \text{F1} + /\text{u}/\ \text{F2} \times /\text{i}/\ \text{F1}) - (/\text{i}/\ \text{F1} \times /\text{æ}/\ \text{F2} + /\text{æ}/\ \text{F1} \times /\text{ɑ}/\ \text{F2} + /\text{ɑ}/\ \text{F1} \times /\text{u}/\ \text{F2} + /\text{u}/\ \text{F1} \times /\text{i}/\ \text{F2})]$$

2.5. Statistical analyses

All statistical tests were conducted using IBM SPSS version 21 (Chicago, IL). To detect differences within and among the groups, mixed ANOVAs were conducted with the least significant difference (LSD) test as the post hoc (Carmer & Swanson, 1973) if assumptions were met. If assumptions were violated, the Kruskal–Wallis with Mann–Whitney *U* as follow-up test was used to detect between-group differences and Wilcoxon Signed Rank test was used to detect differences within groups.

Data were examined for the presence of outliers. Six out of 300 trials were recognized as outliers (i.e., greater than 3 SD above or below the mean) and removed from the vowel duration data; these points were spread across groups (CAS: $n = 1$; SD: $n = 3$; TD: $n = 2$). No other outliers were identified.

We posited that children with CAS would evidence shorter VOTs, fewer accurate productions of /p/, and smaller vowel space areas than the other groups and that these differences would be larger in the masked condition. We also hypothesized that all groups would evidence equivalent vowel durations in the unmasked condition (Odell & Shriberg, 2001), but that in the masked condition children with CAS would evidence shorter vowel durations relative to the other groups. Lastly, we expected that children with CAS would evidence a masking effect on all measures, and that the TD and SD groups would show a masking effect on speech intensity only.

2.6. Reliability

Ten percent of the nonwords from each condition were remeasured to calculate inter-rater reliability. The intra-rater correlation coefficient (ICC) with absolute error in parenthesis was .99 (2.0 ms), .98 (4.3 ms), .99 (3.8 ms), .99 (13.92 Hz), and .99 (23.69 Hz) for unmasked VOTs, masked VOTs, unmasked vowel durations, masked vowel durations, unmasked formant frequencies and masked formant frequencies respectively. The ICC for the CAS feature ratings was .93 (.6 features), showing a high level of agreement for perceptual feature rating using the operational definitions that are included in the supplemental section.

3. Results

The current study investigated the role of auditory feedback on speech production in children with CAS, SD, and typical speech development. Specifically, we sought to determine (1) if school-aged children with CAS would produce VOTs, vowel space areas and vowel durations in masked and unmasked conditions that are comparable to peers with SD and TD; and (2) if children with CAS, SD, and TD would evidence a masking effect on these measures. Table 2 and Fig. 1a–e present mean group data for masked and unmasked VOTs, percent VOTs correct, vowel durations, vowel space areas, and intensity.

3.1. VOT

The mixed ANOVA indicated a group effect $F(2,26) = 3.55$, $p = .04$, partial $\eta^2 = .22$, power = .60. The LSD post hoc test reported children with CAS produced significantly shorter VOTs than the TD group ($p = .019$), but were not different from the SD group ($p = .49$). No significant differences were observed between the SD and TD groups ($p = 065$). There were neither main effects for condition, $F(1,26) = .77$, $p = .39$, $\eta^2 = .03$, power = .14, nor for the group X condition interaction, $F(2,26) = .61$, $p = .55$, $\eta^2 = .05$, power = .14. See Fig. 1a for masked and unmasked VOTs by group.

For the dichotomized VOT data, the percentage of correct VOT productions was calculated for masked and unmasked productions of /pɑb/. The results of the Kruskal–Wallis for unmasked VOTs indicated no statistically significant group difference ($p = .19$); however, there was a significant group effect for masked productions ($p = .009$). Follow-up Mann–Whitney U tests detected that the CAS group had a significantly lower percentage of optimal /p/ exemplars than the TD ($p = .009$) and SD groups ($p = .049$). There were no significant differences between the SD and TD groups, ($p = .42$). The Related-Samples Wilcoxon Signed Rank Test showed a masking effect for only the CAS group ($p = .04$); the other groups did not show an effect of masking on accurate exemplars of /p/. See Fig. 1b for masked and unmasked percent VOTs correct by group.

For unmasked productions the CAS group produced accurate VOTs in 84% of trials, relative to 70% in masked productions. The SD group produced accurate productions in 92% of unmasked and 91% of masked trials and the TD group produced accurate productions in 98% of both masked and unmasked trials.

Table 2
Mean VOTs, percentage VOTs correct, vowel durations, vowel space areas, and intensity by group, by condition. Standard error is in hypotheses.

Measure	TD ($n = 11$)		CAS ($n = 9$)		SD ($n = 10$)	
	Un	M	Un	M	Un	M
Mean VOT /p/ (ms)	64 (5)	78 (6)	62 (7)	54 (6)	78 (6)	64 (7)
VOTs correct (%)	98 (2)	98 (2)	84 (8)	70 (11)	92 (3)	91 (6)
Mean vowel duration (ms)	247 (7)	268 (7)	207 (14)	222 (15)	269 (8)	267 (9)
Intensity (dB)	73 (3)	77 (2)	80 (2)	82 (3)	83 (.6)	83 (.8)
Vowel space area (kHz^2)	543 (75)	577 (92)	470 (40)	386 (52)	458 (31)	481 (67)

VOTs correct: VOTs for /p/ were dichotomized as correct/incorrect based on a cutoff of 35 ms (Lisker & Abramson, 1964; 1967); Un = unmasked; M = masked.

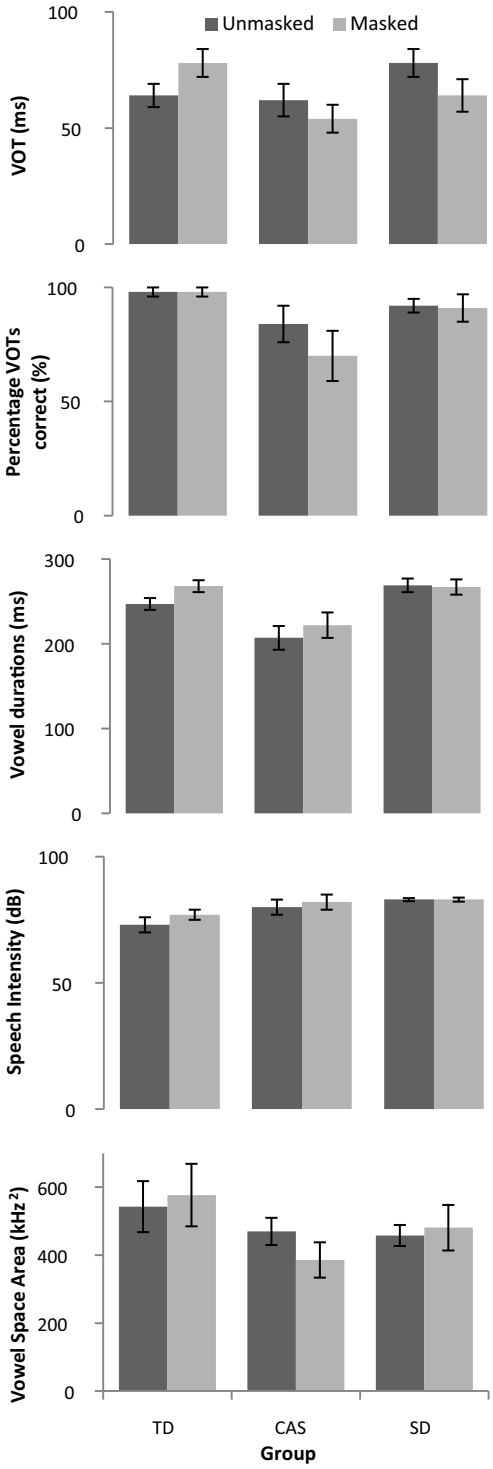

Fig. 1. (a) Mean VOTs, (b) percentage VOTs correct, (c) vowel durations, (d) speech intensity, and (e) vowel space areas by group, by condition. Error bars show standard error.

3.2. Vowel durations

Results of the mixed ANOVA indicated a significant group effect, $F(2,26) = .42$, $p = .001$, partial $\eta^2 = .42$, power = .96. The CAS group produced shorter vowel durations than the SD ($p < .001$) and TD groups ($p = .002$). Although the masking effect

was not significant, $F(1,26) = 2.94$, $p = .09$, there was a moderate effect, partial $\eta^2 = .10$, power = .38, and a trend observed where masked vowel durations were, on average, longer than the unmasked durations.

Although the results failed to detect a statistically significant group X condition interaction ($p = .29$); both the CAS and TD groups displayed a trend toward longer vowel durations for masked productions relative to unmasked productions (TD = 20 ms difference; CAS = 12 ms difference). The SD group showed no measurable difference between conditions. See Fig. 1c for masked and unmasked vowel durations by group.

3.3. Speech intensity

Wilcoxon Signed Ranks Tests were conducted to determine within-group effects of masking on intensity. Children with CAS and TD showed a masking effect on speech intensity where masked speech was louder than unmasked speech (CAS: $p = .017$; TD: $p = .004$). There were no significant differences in loudness across conditions in the SD group ($p = .80$). See Fig. 1d for masked and unmasked speech intensity by group.

3.4. Vowel space area

The Kruskal–Wallis tests indicated no between-group differences in vowel space area for masked ($p = .310$) or unmasked ($p = .656$) productions. Related Samples Wilcoxon Signed Rank Tests reported the CAS group showed a masking effect on vowel space area ($p = .05$) where masked vowel space areas were smaller than unmasked ones. The SD ($p = .953$) and TD ($p = .386$) groups did not show an effect of masking on vowel space area. See Fig. 1e for masked and unmasked vowel space area by group.

4. Discussion

4.1. Increased reliance on auditory feedback suggests weak feedforward representations in children with CAS

Children with CAS produced fewer optimal VOTs for /p/ targets, and reduced vowel space area when auditory feedback was attenuated; in contrast, the SD and TD groups showed no effect of masking on any measures. This finding is consistent with the suggestion that children with CAS rely on auditory feedback to compensate for poorly formed feedforward programs (Terband & Maassen, 2010). In the course of typical development, auditory feedback is used to establish increasingly precise feedforward representations, which ultimately promote rapid, efficient, and consistently correct speech production (Guenther, 2006; Pickering & Garrod, 2013). The current data suggest that children with CAS may not achieve this level of feedforward control and consequently, rely more heavily on information provided by auditory feedback. The suggestion that children with CAS are unable to successfully update and improve the accuracy of their feedforward representations points to deficits in implicit learning that may account for why children with CAS often exhibit poor generalization in response to treatment. Shriberg, Lohmeier, Strand, and Jakielski (2012) have similarly implicated a role for memory and learning deficits in CAS.

4.2. Ongoing reliance on feedback by children with CAS

In our sample, children with CAS produced fewer optimal VOTs for /p/ targets and smaller vowel space areas when auditory feedback was attenuated. These shortened VOTs are consistent with prior observations of voicing distortions and short VOTs in younger children with CAS (Iuzzini & Forrest, 2008; Iuzzini, 2012; Lewis et al., 2004). Acquiring the voicing contrast seems to be an area of difficulty for children with CAS (Iuzzini, 2012) and as such children with CAS may rely heavily on auditory feedback to produce this contrast correctly. Also, noise masking is known to promote enhanced speech clarity in typical talkers, and the current study showed that speech intensity did correlate with vowel space area in talkers with SD and TD in the masked condition. In contrast, the CAS group did not show an association between speech intensity and vowel space area, suggesting that even though this group did show increased loudness in the masked condition, this change did not result in more differentiated vowels. This finding is in agreement with findings from a recent feedback perturbation study in children with speech sound disorders including CAS (Terband et al., 2014). Results revealed that TD controls tended to evidence a compensatory response to the frequency shift, where children with speech sound disorders tended to follow and amplify the shift—although individuals in each group also showed the opposite response. Terband and colleagues suggested that participants with speech sound disorders detected the feedback perturbation but did not have a sufficient feedforward and feedback internal model to compensate appropriately.

The current findings, which support an increased reliance on feedback by children with CAS, are in contrast to the negative effects that auditory feedback has been hypothesized to have on speech in people who stutter. For instance, disfluencies have been attributed to an over-reliance on feedback among stutterers and interventions that delay or attenuate feedback often increase fluency in this population (e.g., Van Borsel, Reunes, & Van den Bergh, 2003). In children with CAS, increased reliance on feedback may help to compensate for a deficient feedforward system, but in persons who stutter, the feedback system may be detrimentally overly engaged.

Our findings motivate additional research on the possible beneficial effects of feedback (i.e., auditory, somatosensory, and visual) on the speech of children with CAS. Visual feedback has been shown to facilitate accurate productions in children

with CAS (Preston, Brick, & Landi, 2013) and adults with AOS (Katz et al., 2007; Katz, McNeil, & Garst, 2010). Feedback may be particularly effective for promoting the consolidation of robust feedforward programs through, for example, errorless learning (e.g., Fillingham, Hodgson, Sage, & Lambon Ralph, 2003).

4.3. Auditory attenuation affects motor programs, not phonological representations

Children with SD were less affected by attenuated auditory feedback than were children with CAS. Because the groups were matched on a variety of factors, it is unlikely that variables such as age or speech severity explain their different responses to masking. This study provides additional support to suggest different underlying deficits in these groups where the errors in our SD group relate to an impairment at the level of the abstract phonologic representation and errors in children with CAS relate to an impairment of motor planning (Levelt, 1989; Munson, Bjorum, & Windsor, 2003; Stackhouse, 1992). Within this framework, the current findings suggest that attenuated auditory feedback may be more effective at disrupting the motor planning process associated with feedforward programs than abstract, linguistic-based phonological representations.

4.4. Increased intensity associated with shorter durations and reduced spectral contrasts in children with CAS

In typical talkers, masking noise elicits increased speech intensity, longer vowel durations, and exaggerated vowel formant frequencies (Junqua, 1993; Maas et al., 2013; Rogers et al., 1996; Van Summers et al., 1988). Presumably, these adaptations are intended to maximize speech clarity in the presence of noise. In the current study, all groups evidenced equivalent VOTs and vowel space areas in the unmasked condition showing that when school-aged children with CAS have access to feedback, they perform similar to their peers on these measures. When feedback was attenuated, however, children with CAS evidenced increased speech intensity but reduced spectral contrasts and shorter VOTs, suggesting that louder speech was not necessarily clearer in these participants. These findings suggest that children with CAS were not capable of implementing adaptive articulatory modifications in response to masking noise.

4.5. Clinical implications

Unlike the children with SD and TD, those with CAS performed differently between the masked and unmasked conditions for optimal productions of /p/ and vowel space area. For instance, 77% of participants with CAS evidenced smaller vowel space in masked relative to unmasked productions, whereas only 36% of TD and 50% of SD showed the effect. In other words, the masking effect was over 1.5 times more common in children with CAS relative to the other groups. These findings may have a number of implications for the assessment and treatment of children with CAS. For example, additional research is needed to determine if speech in these children tends to break down in noisy environments such as a school lunch room. Given the wide range of communication environments that talkers are subjected to, there may be a benefit of assessing children in varying and ecologically valid communication contexts rather than testing a child only in quiet clinical or research environments. Also, treatments that incorporate feedback-manipulations could possibly be useful in eliciting optimized productions or in training stable and robust feedforward programs in this population.

4.6. Limitations

The issue of accurate differential diagnosis in CAS continues to be a challenge in the field. Because a validated list of diagnostic criteria does not exist, clinicians and researchers must rely on a constellation of factors to ensure accurate diagnosis. In this study, we used a two-pronged approach to differential diagnosis for CAS and SD. First, we required that participants in the CAS group had a history of CAS diagnosis and treatment. Second, we had two blinded raters assess each participant on a list of features that are commonly associated with CAS, although not necessarily pathognomonic, and used a cutoff of $\geq$5/11 features to confirm a CAS diagnosis. On average participants with CAS evidenced 8 features whereas those with SD had 3. It is important to note that we rated these features based on whole-word responses on the *GFTA-2* articulation test. We reasoned that if children with a history of CAS—who did not have comorbid language impairment, cognitive deficit, or dysarthria—produced a high number of features on simple test items, we could be more certain in confirming the CAS diagnosis rather than a different disorder such as dyslexia, which could also yield copious errors on complex test items (Catts, 1989).

The current study reports the effects of attenuated auditory feedback on VOT, vowel durations, and intensity calculated on repeated production of the stimulus item /pɑb/. This study provides a proof of concept for greater reliance on auditory feedback by children with CAS relative to SD and TD. Findings should be considered with caution and replicated in other stimuli items in the future.

5. Conclusion

The current study investigated the role of auditory feedback in children with CAS, SD, and TD. Results showed that the speech productions made by children with CAS were negatively affected by attenuated auditory feedback, suggesting this was effective in disrupting the motor planning/programming process believed to be deficient in children with CAS. In contrast, the speech productions of typically developing children and those with SD were not perturbed by masking, which

suggests they have intact feedforward programs robust enough to withstand a temporary disruption of feedback. Future studies should investigate factors that mediate learning of feedforward programs in these populations. In addition, the role of auditory feedback in assessment and treatment of children with CAS should be explored.

Acknowledgements

This research was supported by the University of Nebraska Health Research Consortium (Co-PIs: Hogan & Green), the Childhood Apraxia of Speech Association of North America (PI: Iuzzini), the University of Nebraska Undergraduate Creative Activities and Research Experiences Student Fellowship Program, and the Barkley Memorial Trust. The authors wish to thank the following individuals: Allison Hanen, Kimber Green, Sara Benham, Dyann Rupp, Tacy Corson, Phoebe Chung, Anna Davidson, Natalie Vanderveen and Kristin Schneller. Portions of these data were presented at the American Speech, Language and Hearing Association's annual convention (Iuzzini, Hanen, Green, & Hogan, 2013), the Conference for Motor Speech (Iuzzini, Hogan, & Green, 2014) and the University of Nebraska Undergraduate Research Fair (Hanen, Iuzzini, Green, & Hogan, 2013).

Appendix A. Operational definitions for CAS characteristics

1. *Vowel error*: A vowel production error in which the vowel is substituted for another phoneme OR in which the vowel is recognizable as a specific phoneme but it is not produced exactly correctly (e.g., not a prototypical production, may sound like it is in between two vowels). It is not considered an error if the vowel is substituted with another phoneme that is consistent with an adult-like model (e.g., /hɑtdɑg/, /hɑtdɔg/).
2. *Consonant distortion*: A consonant production error in which a speech sound is recognizable as a specific phoneme but it is not produced exactly correctly (e.g., an /s/ that is produced with lateralization or dentalization).
3. *Stress errors*: An error in which the appropriate stress is not produced correctly. For example: conDUCT and CONduct have different stress patterns. It is considered an error if the stress is inappropriately equalized across syllables, or placed on the wrong syllable.
4. *Syllable segregation*: Brief or lengthy pause between syllables which is not appropriate.
5. *Groping*: Prevocalic (silent) articulatory searching prior to onset of phonation, possibly in an effort to improve the accuracy of the production. Video is needed to assess this feature.
6. *Intrusive schwa (e.g., in clusters)*: A schwa is added between consonants. For example, it may be inserted in between the consonants in a cluster (e.g., /blu/ becomes /bə̲lu/). This NOT considered a "vowel error".
7. *Voicing errors*: A sound is produced as its voicing cognate (e.g., a /p/ that is produced as a /b/). In addition, this could also describe productions which appear to be in between voicing categories (e.g., blurring of voicing boundaries).
8. *Slow rate*: Speech rate is not typical. It is slower during production of part (e.g., zzziiiiiiper/zipper) or the whole word (e.g., tooommmmaaatoooo/tomato).
9. *Increased difficulty with multisyllabic words*: The participant has a disproportionately increased number of errors as the number of syllables increases (as compared to words with fewer syllables).
10. *Resonance or nasality disturbance: Sounds either hyponasal*: not enough airflow out of nose/"stuffy" OR *hypernasal*: too much airflow out of nose for non-nasal phonemes (e.g., plosives).
11. *Difficulty achieving initial articulatory configurations or transitionary movement gestures*: Initiation of utterance or initial speech sound may be difficult for child to produce and may sound lengthened or uncoordinated. Also, child may evidence lengthened or disrupted coarticulatory gestures or movement transitions from one sound to the next.

Appendix B. Percentage of participants in each group who evidenced each characteristic at least once during administration of the *GFTA-2*

	Vowel error	Consonant distortion	Stress error	Syllable segregation	Groping	Intrusive schwa	Voicing error	Slow rate	Difficulty with multisyllabic words	Resonance or nasality disturbance	Difficulty initiating articulation or with coarticulatory transitions
CAS	100	89	78	89	11	89	89	44	33	56	56
SD	20	100	20	30	10	40	60	20	0	0	20
TD	18	82	18	27	9	18	9	19	0	9	9

Appendix C. Continuing education questions

1. Inconsistent speech errors may indicate
   a. that children with CAS have weak oral musculature
   b. that children with CAS have strong feedforward programs
   c. that children with CAS have weak feedforward programs
   d. a & b
2. Children with speech delay, presumably
   a. have weak feedforward programs
   b. have intact feedforward programs
   c. have difficulty integrating auditory feedback
   d. a & c
   e. b & c
3. Children with CAS evidenced _____ vowel space areas in masked relative to unmasked speech.
   a. smaller
   b. larger
   c. equal
   d. groping
4. Increased intensity was associated with _____ in children with CAS.
   a. increased vowel contrasts
   b. decreased vowel contrasts
   c. longer voice onset time durations
   d. disfluencies
5. Weak feedforward programs may be caused by
   a. neural noise
   b. decreased sensitivity of the tongue and palate
   c. groping
   d. a & b
   e. b & c

References

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

American Speech-Language-Hearing Association (1997). *Guidelines for audiologic screening [guidelines]*. Available from www.asha.org/policy

Blumstein, S. E., Cooper, W. E., Goodglass, H., Statlender, S., & Gottlieb, J. (1980). Production deficits in aphasia: A voice-onset time analysis. *Brain and Language, 9*(2), 153–170.

Boersma, P., & Weenink, D. (2013). *Praat: Doing phonetics by computer (version 5.3.43) [software]*. Retrieved from http://www.fon.hum.uva.nl/praat/

Browman, C. P., & Goldstein, L. (1989). Articulatory gestures as phonological units. *Phonology, 6*, 201–251.

Cai, S., Ghosh, S. S., Guenther, F. H., & Perkell, J. S. (2010). Adaptive auditory feedback control of the production of formant trajectories in the Mandarin triphthong /iau/ and its pattern of generalization. *The Journal of the Acoustical Society of America, 128*(4), 2033–2048.

Carmer, S. G., & Swanson, M. R. (1973). An evaluation of ten pairwise multiple comparison procedures by Monte Carlo methods. *Journal of the American Statistical Association, 68*(341), 66–74.

Catts, H. W. (1989). Speech production deficits in developmental dyslexia. *Journal of Speech and Hearing Disorders, 54*(3), 422–428.

Fillingham, J. K., Hodgson, C., Sage, K., & Lambon Ralph, M. A. (2003). The application of errorless learning to aphasic disorders: A review of theory and practice. *Neuropsychological Rehabilitation, 13*(3), 337–363.

Forrest, K. (2003). Diagnostic criteria of developmental apraxia of speech used by clinical speech-language pathologists. *American Journal of Speech-Language Pathology, 12*(3), 376–380.

Forssberg, H., Eliasson, A. C., Kinoshita, H., Johansson, R. S., & Westling, G. (1991). Development of human precision grip. I: Basic coordination of force. *Experimental Brain Research, 85*(2), 451–457.

Gierut, J. A. (1998). Treatment efficacy: Functional phonological disorders in children. *Journal of Speech, Language, and Hearing Research, 41*(1), S85–S100.

Goldman, R., & Fristoe, M. (2000). *Goldman–Fristoe Test of Articulation* (2nd ed.). Circle Pines, MN: Pearson.

Guenther, F. H. (2006). Cortical interactions underlying the production of speech sounds. *Journal of Communication Disorders, 39*(5), 350–365.

Haas, G., Diener, H. C., Rapp, H., & Dichgans, J. (1989). Development of feedback and feedforward control of upright stance. *Developmental Medicine & Child Neurology, 31*(4), 481–488.

Hanen, A., Iuzzini, J., Green, J. R., & Hogan, T. P. (2013, April). *Effect of auditory masking in children with CAS, speech delay, and typical speech development* Poster presentation at undergraduate research fair Lincoln, Nebraska: University of Nebraska-Lincoln.

Haruno, M., Wolpert, D. M., & Kawato, M. (2001). Mosaic model for sensorimotor learning and control. *Neural Computation, 13*(10), 2201–2220.

Houde, J. F., & Jordan, M. I. (1998). Sensorimotor adaptation in speech production. *Science, 279*(5354), 1213–1216.

Houde, J. F., & Jordan, M. I. (2002). Sensorimotor adaptation of speech. I: Compensation and adaptation. *Journal of Speech, Language, and Hearing Research, 45*(2), 295–310.

Houde, J. F., & Nagarajan, S. S. (2011). Speech production as state feedback control. *Frontiers in Human Neuroscience, 5*, 1–14.

Iuzzini, J. (2012). *Inconsistency of speech in children with childhood apraxia of speech, phonological disorders, and typical speech development* (Unpublished dissertation) Bloomington, IN: Indiana University.

Iuzzini, J., & Forrest, K. (2008, March). *Acoustic vowel area in phonologically disordered, CAS, and normal speech* Poster presented at the conference for motor speech, Monterey, CA .

Iuzzini, J., & Forrest, K. (2010). Evaluation of a combined treatment approach for childhood apraxia of speech. *Clinical Linguistics and Phonetics, 24*, 335–345.

Iuzzini, J., Hanen, A., Green, J. R., & Hogan, T. P. (2013, November). *Effect of auditory masking in children with CAS, speech delay, and typical speech development* Technical session at the annual convention of the American Speech, Language, and Hearing Association, Chicago, IL .

Iuzzini, J., Hogan, T., & Green, J. (2014, February). Effect of masking on vowel space area in childhood apraxia of speech. *Presentation at the Conference for Motor Speech, Sarasota, Florida*.
Jones, J. A., & Munhall, K. G. (2000). Perceptual calibration of F0 production: Evidence from feedback perturbation. *The Journal of the Acoustical Society of America, 108*, 1246–1251.
Junqua, J. C. (1993). The Lombard reflex and its role on human listeners and automatic speech recognizers. *The Journal of the Acoustical Society of America, 93*(1), 510–524.
Katz, W. F., Garst, D. M., Carter, G. S., McNeil, M. R., Fossett, T. R., Doyle, P. J., et al. (2007). Treatment of an individual with aphasia and apraxia of speech using EMA visually-augmented feedback. *Brain and Language, 103*(1), 213–214.
Katz, W. F., McNeil, M. R., & Garst, D. M. (2010). Treating apraxia of speech (AOS) with EMA-supplied visual augmented feedback. *Aphasiology, 24*(6-8), 826–837.
Lane, H., & Tranel, B. (1971). The Lombard sign and the role of hearing in speech. *Journal of Speech, Language, and Hearing Research, 14*(4), 677–709.
Levelt, W. J. (1989). *Speaking. From intention to articulation*. Cambridge, MA: MIT Press.
Lewis, B. A., Freebairn, L. A., Hansen, A. J., Iyengar, S. K., & Taylor, H. G. (2004). School-age follow-up of children with childhood apraxia of speech. *Language, Speech, and Hearing Services in Schools, 35*(2), 122–140.
Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word, 20*, 384.
Lisker, L., & Abramson, A. S. (1967). Some effects of context on voice onset time in English stops. *Language and Speech, 10*(1), 1–28.
Maas, E., Mailend, M. L., & Guenther, F. (2013). Feedback and feedforward control in speech production in apraxia of speech and aphasia. *Clinical aphasiology conference*.
Maassen, B., Nijland, L., & Van Der Meulen, S. (2001). Coarticulation within and between syllables by children with developmental apraxia of speech. *Clinical Linguistics & Phonetics, 15*(1–2), 145–150.
Macken, M. A., & Barton, D. (1980). The acquisition of the voicing contrast in English: A study of voice onset time in word-initial stop consonants. *Journal of Child Language, 7*(01), 41–74.
Munson, B., Bjorum, E. M., & Windsor, J. (2003). Acoustic and perceptual correlates of stress in nonwords produced by children with suspected developmental apraxia of speech and children with phonological disorder. *Journal of Speech, Language, and Hearing Research, 46*(1), 189–202.
Nijland, L., Maassen, B., Meulen, S. V. D., Gabreëls, F., Kraaimaat, F. W., & Schreuder, R. (2002). Coarticulation patterns in children with developmental apraxia of speech. *Clinical Linguistics & Phonetics, 16*(6), 461–483.
Odell, K. H., & Shriberg, L. D. (2001). Prosody-voice characteristics of children and adults with apraxia of speech. *Clinical Linguistics & Phonetics, 15*(4), 275–307.
Picheny, M. A., Durlach, N. I., & Braida, L. D. (1986). Speaking clearly for the hard of hearing. II: Acoustic characteristics of clear and conversational speech. *Journal of Speech, Language, and Hearing Research, 29*(4), 434–446.
Pickering, M. J., & Garrod, S. (2013). Forward models and their implications for production, comprehension, and dialogue. *Behavioral and Brain Sciences, 36*(4), 377–392.
Preston, J. L., Brick, N., & Landi, N. (2013). Ultrasound biofeedback treatment for persisting childhood apraxia of speech. *American Journal of Speech-Language Pathology, 22*(4), 627–643.
Reynolds, C. R., & Kamphaus, R. W. (2003). *Reynolds Intellectual Assessment Scales*. Lutz, FL: PAR Inc.
Rogers, M. A., Eyraud, R., Strand, E. A., & Storkel, H. (1996). The effects of noise masking on vowel duration in three patients with apraxia of speech and a concomitant aphasia. *Clinical Aphasiology, 24*, 83–96.
Semel, E., Wiig, E. H., & Secord, W. A. (2003). *Clinical evaluation of language fundamental* (4th ed.). San Antonio, TX: Pearson.
Shriberg, L. D., Aram, D. M., & Kwiatkowski, J. (1997). Developmental apraxia of speech. I. Descriptive and theoretical perspectives. *Journal of Speech, Language, and Hearing Research, 40*(2), 273–285.
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.
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.
Stackhouse, J. (1992). Developmental verbal dyspraxia. I: A review and critique. *International Journal of Language & Communication Disorders, 27*(1), 19–34.
Terband, H., & Maassen, B. (2010). Speech motor development in childhood apraxia of speech: Generating testable hypotheses by neurocomputational modeling. *Folia Phoniatrica et Logopaedica, 62*(3), 134–142.
Terband, H., Maassen, B., Guenther, F. H., & Brumberg, J. (2009). Computational neural modeling of speech motor control in childhood apraxia of speech (CAS). *Journal of Speech, Language, and Hearing Research, 52*, 1595–1609.
Terband, H., van Brenk, F., & van Doornik-van der Zee, A. (2014). Auditory feedback perturbation in children with developmental speech sound disorders. *Journal of Communication Disorders, 51*, 64–77.
Van Borsel, J., Reunes, G., & Van den Bergh, N. (2003). Delayed auditory feedback in the treatment of stuttering: Clients as consumers. *International Journal of Language & Communication Disorders, 38*(2), 119–129.
Van Summers, W., Pisoni, D. B., Bernacki, R. H., Pedlow, R. I., & Stokes, M. A. (1988). Effects of noise on speech production: Acoustic and perceptual analyses. *The Journal of the Acoustical Society of America, 84*(3), 917–928.
Vorperian, H. K., & Kent, R. D. (2007). Vowel acoustic space development in children: A synthesis of acoustic and anatomic data. *Journal of Speech, Language, and Hearing Research, 50*(6), 1510–1545.