Published in final edited form as:
Dev Neuropsychol. 2014 July ; 39(5): 385–403. doi:10.1080/87565641.2014.939181.

Neurophysiology of Speech Differences in Childhood Apraxia of Speech

Jonathan L. Preston,
Haskins Laboratories, New Haven, CT & Southern Connecticut State University, New Haven, CT

Peter J. Molfese,
Haskins Laboratories, New Haven, CT

Nina Gumkowski,
Haskins Laboratories, New Haven, CT

Andrea Sorcinelli,
New York University, New York NY

Vanessa Harwood,
Haskins Laboratories, New Haven, CT & University of Connecticut, Storrs, CT

Julia Irwin, and
Haskins Laboratories, New Haven, CT & Southern Connecticut State University, New Haven, CT

Nicole Landi
Haskins Laboratories, New Haven, CT & University of Connecticut, Storrs, CT

Abstract

Event-related potentials (ERPs) were recorded during a picture naming task of simple and complex words in children with typical speech and with childhood apraxia of speech (CAS). Results reveal reduced amplitude prior to speaking complex (multisyllabic) words relative to simple (monosyllabic) words for the CAS group over the right hemisphere during a time window thought to reflect phonological encoding of word forms. Group differences were also observed prior to production of spoken tokens regardless of word complexity during a time window just prior to speech onset (thought to reflect motor planning/programming). Results suggest differences in pre-speech neurolinguistic processes.

Childhood apraxia of speech (CAS) is a subtype of speech sound disorder that impacts speech intelligibility through impaired precision and consistency of speech sounds. CAS is characterized by disrupted transitions between sounds and impairments in prosodic elements of speech such as lexical stress (ASHA, 2007). One manifestation of CAS is significant breakdown in the precision of speech as words become increasingly complex (e.g., increased number of phonemes and syllables). Although CAS is believed to have a neurobiological basis that results in an impaired ability to plan and/or program the movements for speech

Corresponding Author: Jonathan L. Preston, 300 George St., Suite 900, Haskins Laboratories, New Haven, CT, 06511 preston@haskins.yale.edu 203-865-6163 x 247.

(ASHA, 2007), there is currently little understanding of the neurobiological foundation of the core deficits in producing complex word forms and planning and programming movements. CAS is often identified in early childhood but symptoms can remain well into school age or even adulthood, and there may be lasting impacts on social, academic, and communicative success. Treatment outcomes remain less than optimal, reflecting a lack of understanding of the underlying neurolinguistic mechanisms for CAS. Much behavioral research has been conducted on CAS, but few studies have addressed neurobiological foundations of the disorder. The present study aims to characterize the neurobiological basis of differences in generating simple and complex word forms and preparing articulatory movements in school-age children with CAS.

One clinical characteristic of CAS is a breakdown in speech accuracy as words become increasingly complex, and this construct is often evaluated as part of diagnostic protocols in research and in clinical practice (Forrest, 2003; Lewis et al., 2004; Shriberg et al., 2010; Strand, McCauley, Weigand, Stoeckel, & Baas, 2013). Complex word forms can be difficult for children with CAS to produce, especially when challenging syllable structures such as consonant sequences are attempted (Crary, 1984; Maassen et al., 2001). In particular, multisyllabic utterances are commonly in error among children with CAS, with errors encompassing omissions, substitutions, additions and distortions of sounds, lexical stress errors, and sound sequencing errors (Crary, 1984; Lewis et al., 2004; Nijland et al., 2003; Peter, Button, Stoel-Gammon, Chapman, & Raskind, 2013; Shriberg et al., 2003; Velleman & Shriberg, 1999). Hence, errors in phonological and phonetic output become increasingly apparent as children with CAS generate longer phoneme sequences and plan the associated movements.

Numerous processes are involved prior to speech production, and the descriptions of those processes vary among researchers. In general, after a word is retrieved from the lexicon, phonological stages occur prior to phonetic stages. Phonological stages involve selection and retrieval of phonological representations as well as assembling/ordering phonological elements within a frame, and specifying lexical stress. Therefore, phonological stages encode the length, complexity, and prosody of the utterance, and these stages may be vulnerable to disruption in CAS (Shriberg et al., 2012). Phonological stages are followed by phonetic processes of motor planning and programming (or “transcoding,” cf. Shriberg et al., 2012). Motor planning involves generating the movement goals for a selected word or utterance (van der Merwe, 2008), and it requires specifying the general spatial and temporal parameters associated with targeted articulatory movements. After the motor plan has been constructed, motor programming ensues by specifying how the movement goal will be achieved; motor programming involves selecting and communicating with motor units that can accomplish the plan. Children with CAS may have difficulty with either or both of these phonetic processes (Crary, 1984; Maassen, Nijland, & Van Der Meulen, 2001; Nijland & Maassen, 2003; Nijland et al., 2002; Shriberg, Lohmeier, Strand, & Jakielski, 2012). However, differences related to phonological and phonetic processes are difficult to infer simply by listening to children’s speech. The present study seeks to identify objective, brain-based markers of these differences in children with CAS.

Picture naming is a common paradigm used during assessment and treatment of CAS. Levelt, Indefrey and colleagues (Indefrey, 2011; Indefrey & Levelt, 2004; Levelt, 1999; Levelt, Praamstra, Meyer, Helenius, & Salmelin, 1998; Levelt, Roelofs, & Meyer, 1999) have outlined a Word Production Model (WPM) that specifies the general time course of the psycholinguistic processes (and, to some degree, the associated neurolinguistic processes) revealed by picture naming studies of typically speaking adults. Although exact ranges of time windows may vary somewhat depending on the nature of the lexical items, participant characteristics, etc., some general principles have been identified in the WPM. The first approximately 175 ms after a picture is presented are generally associated with visual processing of the picture. This stage is followed by selection of the lexical item, which may occur from approximately 175–275 ms post picture onset. These early processes are not hypothesized to be impaired in CAS. However, subsequent processes in the phonological stages are thought to involve phonological retrieval (approximately 275–350 ms post picture onset) followed by syllabification (approximately 350–475 ms post picture onset), which involves combining phonemes into syllables and assigning stress patterns. This syllabification process is likely to begin soon after the first sounds are retrieved, with syllabification of initial sounds and syllables perhaps occurring before the final sounds are retrieved, and thus syllabification of early sounds may overlap with the phonological retrieval of later sounds in words (Indefrey, 2011). The syllabification process generally involves constructing the frame from which the motor plan is generated, which will differ depending on the complexity of the utterance.

The WPM identifies phonetic encoding as the final stage just prior to articulatory movement (approximately 475–600 ms post picture onset) in which messages are transmitted to muscle groups; this phase could also be characterized as motor planning and programming. Motor planning may overlap in time with syllabification, as the plan for early parts of a word may be initiated while later parts of the word are still being assembled (Meyer et al., 2003). Although the WPM may over-simplify the complex processes involved in lexical access and speech production (e.g., it fails to identify feedback processes), several decades of both behavioral and neuroimaging studies of adults support the general ordering of processes and the temporal windows associated with picture naming (Indefrey, 2011; Indefrey & Levelt, 2004; Levelt, 1999; Levelt et al., 1999). We assume that children would follow a similar ordering of processes, although children’s naming latencies may be slightly delayed relative to adults. The neural differences in phonological and phonetic processes in children with CAS would therefore be expected to emerge in a picture naming task around 350 ms after the picture is presented but before speech production occurs (i.e., during syllabification processes in the phonological stage, and during planning and programming in the phonetic stages).

The functional neurobiology of speech planning and programming differences in CAS has not been well explored. Functional magnetic resonance imaging (fMRI) data have been reported on a speech production task in one family with a genetic mutation causing CAS, and these data suggest widespread differences in cortical and subcortical speech production networks. Liégeois et al. (2003) reported that, during a speech production task, family members with CAS showed reduced activation in inferior frontal gyrus compared to family members without CAS. However, the temporal resolution of fMRI is quite limited and does

not allow for an analysis of the specific processes underlying speech production. To evaluate the specific neurobiological characteristics of speech motor planning and programming, more temporally sensitive neurobiological tools such as EEG/ERP are needed.

Temporally sensitive measures

Electroencephalography (EEG) records electrical activity of the brain at the scalp. One advantage of EEG is its fine-grained temporal resolution, which allows for analysis of neural activity on the order of milliseconds. Time-locked analysis of the EEG signal in response to a stimulus is known as Event-Related Potentials (ERP), and ERPs have been used for many decades to assess neural signatures associated with a variety of speech and language processes in normal and disordered populations. To date, there appears to be only one ERP study of CAS. Froud and Khamis-Dakwar (2012) used ERPs to study the mismatch negativity (MMN) responses in five children with CAS and five age-matched controls. They observed that children with typical speech showed a standard MMN response to phonemically contrastive tokens (i.e.,/ba/vs./pa/), but children with CAS showed a less mature response. However, children with CAS did show a MMN response when presented with allophonic contrasts (i.e.,/pa/vs./p^ha/), which was not observed in the children with typical speech. The result was interpreted as atypical processing of phonetic and phonological information in children with CAS with overly specified representations for at least some phonological information. This study provides some indication that there may be underlying phonological deficits in addition to problems planning/programming speech-related movements. However, the prevailing theory of the fundamental difference in children with CAS is in planning/programming the movements for speech (ASHA, 2007), rather than in phonological processing or auditory perception per se. Thus, it is possible that a MMN paradigm might not be sensitive to the *causal* mechanisms underlying impairments in planning speech.

ERPs have been used to study neurophysiological processes involved in preparing to speak, although to a much lesser extent than studies of auditory perception. Using Levelt and Indefry's WPM as a guide, Eulitz, Hauk and Cohen (2000) used ERPs to examine covert relative to overt speech produced by adults in response to pictures. They found clear differences between these conditions in a window from 275–400 ms after picture onset. This difference was interpreted to reflect the component associated with phonological and phonetic stages necessary for producing the word. Additionally, they observed stronger activation over the left hemisphere than the right. Other imaging studies have observed physiological responses at time windows that generally agree with the behavioral studies of picture naming (Indefrey, 2011; Indefrey & Levelt, 2004; Levelt et al., 1998).

One additional ERP paradigm relevant to the current investigation involves identification of pre-motor potentials. A number of studies have observed slow negative ERP signals present before the onset of limb movements, oral movements, and speech (Mcardle, Mari, Pursley, Schulz, & Braun, 2009; Sochůrková, et al., 2006; Tremblay, Shiller & Gracco, 2008; Wheaton, Shibasaki & Hallett, 2005; Wohlert, 1993). In such studies, the ERP analyses are time-locked to the onset of movement. Thus, a slow negative wave prior to the onset of speech indexes motoric processes involved in planning or programming the movement. We

therefore expect to observe differences in this signal prior to speech onset in children with CAS.

Purpose and hypotheses

There is presently little neurophysiological evidence of the core impairments in CAS that are believed to occur before the onset of speech. The present study employs a picture naming paradigm to evaluate whether the assumed difference in these underlying processes can be revealed by ERP. Although the long-term goal is to identify neurobiological markers of CAS, the current study takes an initial step to determine if a simple picture naming task can be used to identify differences that are in line with theoretical predictions associated with phonological complexity and motor planning and programming. Using the temporal guidelines predicted by Levelt and Indefrey's WPM, we examine the electrophysiological basis of processes associated with preparing to speak simple and complex words in children with and without CAS. Based on prior studies of adults performing picture naming tasks, it was hypothesized that the groups would differ in their neural organization for preparing to produce complex words and that these differences would arise primarily during late phonological stages of syllabification (approximately 350–475 ms post-picture onset); that is, the effects of stimulus complexity would be expected to be revealed at these stages in which assembling phonemes and imposing lexical stress is occurring. Additionally, later stages of phonetic encoding in which articulatory movements are planned and programmed (approximately 475 ms and beyond) would also be expected to differ in CAS. Thus, we hypothesized that differences between children with CAS and controls would be revealed during syllabification of articulatorily complex items and during planning and programming stages of all items. Finally, whereas previous functional MRI studies have revealed increased engagement of the right hemisphere for speech production tasks, we sought to explore whether ERP data would reveal differences in scalp topographical lateralization[1] (which can indicate underlying neural lateralization) during a speech production task in children with CAS. The time windows of interest were phonological stages that encompass assembly and ordering of phonemes (i.e., 350–475 ms after picture onset) as well as later phonetic stages associated with motor planning and programming prior to speech movements (i.e., 475–600 ms after picture onset). As an additional analysis, we sought to examine pre-speech potentials just prior to the onset of movement.

Method

Participants

Two groups of participants were recruited: children with typical speech (TS) and children with childhood apraxia of speech (CAS). All participants in both groups were between 9–15 years of age (mean 12 yr 4 mo, SD 22 mos), had English as their native language, and reported no history of head trauma. Additionally, all participants achieved a standard score above 80 on the Expressive Vocabulary Test-2 (Williams, 2007).

[1]Although scalp laterality differences do not map directly to specific underlying regions, when responses are more or less lateralized, global laterality differences can be hypothesized.

To be eligible for the typical speech (TS) group, children were required to have no history of speech, language, hearing, or learning difficulties. They also had normal speech sound production and prosody, as judged by a licensed speech-language pathologist (the first author) in a short conversation and as determined by a standard score above 95 on the Goldman-Fristoe Test of Articulation-2 (Goldman & Fristoe, 2000). Fourteen children were recruited and met eligibility criteria for the TS group.

Children with Childhood Apraxia of Speech (CAS) were diagnosed by local clinicians as having CAS. All had been in speech-language therapy since the age of 2 or 3 and were still receiving services. For the present study, two speech-language pathologists familiar with CAS (including the first author) confirmed the diagnosis through analysis of children's performance on several tasks (cf. Preston, Brick & Landi, 2013). To confirm that children had speech sound production problems, children in the CAS group were required to score below 80 on the GFTA-2 (mean 57.8, SD 15.8). Although all participants with CAS had errors on rhotic sounds/r, ɝ, ɚ/during the GFTA-2 and conversational speech, children with CAS produced sound substitutions, omissions, and distortions on other sounds as well (including consonants as well as vowels). In addition, errors of migration (i.e., moving sounds out of order) and/or metathesis (switching the order of sounds in words) were also observed in 3–4 syllable words. Errors in lexical and phrasal stress were also observed in multisyllabic words and in conversation. Hence, their errors were not just phoneme-specific distortions as might be expected of children with residual speech sound errors who do not have CAS. Additionally, to verify speech motor sequencing difficulties, participants with CAS had to score below 85% on the Sequencing subtest of the Verbal Motor Production Assessment for Children (Hayden & Square, 1999). Qualitatively, children with CAS produced slow and inaccurate sound sequences on diadochokinetic tasks (e.g.,/p^t^k^/,/a-m-u/, etc.). Finally, the Recalling Sentences subtest of the CELF-4 was used both to evaluate language skills (language memory and complex morpho-syntax) and to judge articulatory precision, sound sequencing, and prosody at the sentence level. Errors on this task in production of vowels, consonant sequences, and lexical and phrasal stress were used to confirm the diagnosis of CAS. Because there are currently no specified values for diagnosing these features in CAS, these symptoms were used to identify CAS but specific thresholds were not employed. Of the 14 children who were referred by clinicians as having CAS, nine met eligibility criteria for the CAS group.

One child from each group did not provide sufficiently analyzable ERP data for the present study (see ERP Data Collection and Processing below). Therefore, the ERP data included 13 children in the TS group (9 male, 4 female) and 8 in the CAS group (7 male, 1 female).

ERP Task and Stimuli

A picture naming task was used to compare the two groups on items that require relatively low or relatively high demands on speech motor planning (i.e., Simple and Complex items). A stimulus list (see Appendix) was developed for each of the two conditions, with 12 items in each condition. All items in both the Simple and Complex conditions began with voiced plosives/b/or/d/. Items in the *Simple* condition were designed to place relatively low demands on speech planning. Simple items were monosyllabic, contained no consonant

clusters, and had the phonological forms CV (e.g., boy) or CVC (e.g., bone). Complex items were designed to place relatively high demands on speech planning. Complex items were 2–3 syllables in length and were allowed to contain consonant clusters (e.g., butterfly, blueberry). All items in the Complex condition began with a stressed syllable. As participants were being prepared for the EEG cap, they were familiarized with the pictures to ensure they used the correct word (e.g., “boy” rather than “child”); to remind the participant of the target word, any picture in which the participant used the wrong label was repeated two more times before the ERP experiment began. Pictures were 300 × 300 pixels (4.2 × 4.2 inches) and were displayed on an 18 inch monitor with the participant sitting in a chair approximately 40 inches from the monitor. Participants were instructed to “name the picture when it appears on the screen.”

Six blocks were presented, with each block containing all 12 Simple and all 12 Complex words in a random sequence. Hence, 144 individual trials were available for analysis for each participant (72 Simple, 72 Complex). The experimental task took approximately 12 minutes. Pictures were on the screen for a jittered interval of 3250–3550 ms, then a black screen was presented for 2000 ms after the picture disappeared. A microphone was positioned 4 inches from the children to track response time based on the onset of the acoustic signal.

ERP Data Collection

Stimulus presentation was controlled using a recent model Dell PC with E-PRIME 2.0 Professional (PST, Inc.) running Windows XP Service Pack 3. EEG data were collected using a 128 electrode net with a Net Amps 300 amplifier (EGI, Inc.) connected to an Apple Mac Pro desktop computer running Mac OS X 10.6.8. Data were sampled at 500 Hz with reference to Cz. Accurate timing was ensured through routine measurements of visual onset delay as measured by a photocell. Impedances measured before and after each recording session and were kept below 40 kΩ.

Stimulus-locked data analysis

Following data acquisition, ERP data were filtered using a bandpass of 0.1 Hz to 30 Hz. Files were segmented into 700 ms epochs, which included 100 ms before picture onset and 600 ms after. To ensure only trials with valid responses were included, trials were marked for exclusion if response time between picture onset and the acoustic response was less than 350 ms or greater than 1400 ms post stimulus onset (6% of trials). Additional trials to be discarded from the analysis had been identified by the ERP lab manager, an analysis specialist who conducted the experiment, and confirmed by the senior author. These included trials in which the participant did not name the picture, said something other than the target word, or revised their response. However, trials in which speech sound errors occurred (e.g., sound distortions, substitutions or omissions) were not removed from the analysis; these errors are characteristic of the disorder and thus were kept in for purposes of ecological validity.

Next, trials containing artifacts (eye movements, eye blinks, and bad channels) were identified using a semi-automated method. Thresholds for identification of eye blinks and

eye movements were those that exceeded a max-min absolute threshold of 150μV measured using electrodes directly above and below the eyes; channels horizontal to the eyes that exceeded a threshold of 100μV were identified as eye movements. Other channels in the net were marked unusable if they exceeded a threshold of 200μV. Channels found to be unusable for more than 40% of the segments were excluded for all segments. Segments were removed if they contained more than 10 unusable channels, eye blinks, or eye movements. On remaining trials, bad channels were replaced using spherical spline interpolation (Perrin, Pernier, Bertrand, & Echallier, 1989) as recommended by Picton et al. (2000). All remaining epochs were then re-referenced to the average reference to remove channel bias (Dien, 1998), and baseline corrected based on the 100 ms prior to stimulus onset. Finally data were averaged by epoch type (Simple vs. Complex) for each participant separately.

Subjects providing fewer than 12 good trials per condition were removed from statistical analyses; this criterion resulted in one CAS participant being excluded. The average number of trials for each group is presented in Table 1. All remaining subjects were screened for excessive noise using an averaging procedure which flips the polarity of every other trial. The total noise perturbation from 0 was measured over the entire epoch and averaged across all channels for each participant using code written in R by the second author. One subject in the TS group was found to be an outlier based on a standard normal distribution; this participant was excluded from the analysis.

ERP averaged data were processed using the ERP PCA Toolkit (Dien, 2010). Data for all participants were down sampled to 250 Hz to reduce the impacts of noise and temporal correlation in the data. Next, a principal components analysis (PCA) with promax (oblique) rotation was run to identify temporal factors of interest; the purpose of the PCA was to identify systematic variance (without direct reference to group or to stimulus condition). Seven temporal factors were extracted from the PCA using a scree test (Cattell, 1966). Although PCA temporal factors are active over the course of the entire ERP average, we used a loading criterion of 0.6 to identify time windows when the factors were most active (Dien, 2010). We then retained only temporal factors that accounted for at least 5% of the variance. There were three temporal factors (TFs) that met these criteria. Two TFs encompassed time windows of theoretical interest (TF2, TF3). TF1 (accounting for 29% of the variance) encompassed a broad time window from 88–372 ms post stimulus onset, and included time windows canonically associated primarily with visual processing, lexical retrieval and the early stages of phonological retrieval. TF1 was not further explored, although the Supplemental Table lists the outcomes of the statistical comparisons.

The two TFs of theoretical interest identified by the PCA were TF2 and TF3. The temporal factor with the second highest eigenvalue (TF2) loaded at 484–600 ms post stimulus onset, encompassing a late time window. This corresponds relatively well to the time window relevant to motor planning and programming. TF2 accounted for 21% of the temporal variance. The PCA also identified a third temporal factor (TF3) at 372–460 ms post stimulus onset, which accounted for 17% of the temporal variance. This time window appeared to generally correspond to a window associated with syllabification and early stages of motor planning as described above. However, based on visual inspection of the grand average data (see Figure 1), the waveform in this time window appeared to begin earlier than 372 ms,

thus our analysis window was selected to encompass the entire window of the waveform, which was 300–460 ms post stimulus onset.

Following the temporal PCA, a spatial PCA with infomax rotation was then run on each temporal factor to identify channels that loaded strongly within each time window. The mean amplitudes from the channels in the first spatial factor were extracted for both left and right hemispheres and were submitted for statistical analysis for each temporal-spatial factor pairing. When the spatial PCA identified clusters involving both anterior and posterior electrodes, differences in polarity preclude averaging across all electrodes. We selected the cluster that included the most electrodes (posterior electrodes in the case of the early time window, anterior electrodes in the case of the later time window).

To summarize the stimulus-locked analysis, an early time window (300–460 ms post stimulus onset) and a late time window (484–600 ms post stimulus onset) identified by the temporal PCA were used for group comparison. These time windows are in line with those described by the Levelt and Indefrey WPM as later stages of phonological encoding (syllabification) and phonetic processes of motor planning and programming. Within each time window, a spatial PCA was run to identify clusters of electrodes with the maximal variance in the time window. The groups were compared using separate 2 (group: TS, CAS) x 2 (condition: simple, complex) x 2 (hemisphere: left, right) ANOVAs for the regions associated with the early temporal factor and the later temporal factor.

Response-locked analysis

In addition to the aforementioned analysis that was time-locked to the onset of the picture, we also employed a supplemental response-locked analysis. This analysis examined the 600 ms prior to the acoustic response at the onset of the word. Procedures were similar to the stimulus-locked analysis except that the segmention included 700 ms epochs prior to the acoustic response, including a 100 ms baseline. Six participants in the CAS group and nine in the TS group provided usable data for this analysis. A temporal-spatial PCA was again conducted using the response-locked segmentation. First, a temporal PCA with promax rotation was conducted. Three TFs were identified that had temporal loadings greater than 0.6 and that accounted for at least 5% of the variance. The first temporal factor (with the highest loading, accounting for 47% of the variance) ranged from −188 ms prior to the acoustic response to 0ms (the acoustic response). This factor was further explored as it was presumed to reflect the processes of planning/programming movements just prior to the onset of speech. A spatial PCA was run with infomax rotation to identify electrodes associated with the first spatial factor. Both anterior and posterior electrodes loaded on this factor; there were more posterior electrodes so these were averaged and included in the statistical analysis. The groups were compared in this time window and in these posterior electrodes using a 2 (group: TS, CAS) x 2 (condition: simple, complex) x 2 (hemisphere: left, right) ANOVA. The remaining two TFs identified by the temporal PCA were not further explored, although statistical comparisons are presented in Supplemental Table 2.

Results

Descriptive statistics for behavioral and demographic data are presented in Table 1. The groups did not differ in age (t[19]=.049, p=0.63) or EVT-2 scores (t[19]=2.1, p=0.061) but as expected, they did differ in GFTA-2 scores (t[1, 19], p<0.0001). Additionally, with respect to response time (from picture onset to the acoustic response), the CAS group was significantly slower than the TS group (F[1, 38]=27.4, p<0.0001); there was no significant main effect for complexity (F[1, 38]=0.18, p=0.67) and no group x complexity interaction (F[1, 38]=0.08, p=0.78). A 2 × 2 ANOVA confirmed that there was no significant difference in the number of usable ERP trials between the groups (F[1, 38]=0.3, p=0.58) or conditions (F[1, 38]=0.03, p=0.86) and no group x condition interaction (F[1, 38]=0.03, p=0.86) in the stimulus-locked analysis.

Analysis of the early stimulus-locked temporal window

Visual inspection of the waveforms revealed patterns similar to those described in other picture naming studies (e.g., Eulitz et al., 2000), providing confirmation that the task was eliciting the expected responses. Figure 1 shows selected electrodes from regions identified by the spatial PCA along with representative waveforms associated with the earlier "planning" time window from 300–460 ms (TF3). As can be seen in the figure, the CAS group shows separation of the simple and complex conditions in the posterior right hemisphere, whereas no such separation exists for the TS group in either hemisphere. The results of the ANOVA indicated a significant three-way interaction of group x hemisphere x complexity (F[1, 19]=6.65, p =0.018, partial η^2=0.259) in this early time window. The three-way interaction was decomposed by running separate group x complexity ANOVAs in the left and right hemispheres. There was no significant group x complexity interaction over the left hemisphere (F[1, 19]=0.92, p=0.348, partial η^2=0.046) but there was a significant group x complexity interaction over the right hemisphere (F[1,19]=4.78, p=0.041, partial η^2=0.201). Thus, the magnitude of the difference between simple and complex items was greater for the CAS group than for the TS group in the right hemisphere. When decomposed by group, the CAS group showed a significant complexity x hemisphere interaction (F[1, 7]=7.09, p=0.032, partial η^2=.503) but the TS group did not (F[1, 12]=0.44, p=0.518, partial η^2=0.036). Thus, the pattern of response for simple and complex items was similar across hemispheres for the TS group, but the electrophysiological response for simple and complex items was not the same across hemispheres for the CAS group. As can be seen in Figure 1, the difference between simple and complex items over the right hemisphere is larger for the CAS group than for the TS group (with reduced amplitude over the right hemisphere for complex items in the CAS group); the left hemisphere waveforms appear similar by group and by condition.

Analysis of the late stimulus-locked temporal window

Figure 2 shows the electrodes selected for analysis with representative waveforms for the later time window from 484–600 ms (TF2). Mean amplitudes were analyzed from this time window identified by the PCA. As can be seen in Figure 2, the waveforms diverge for the TS and CAS groups in this window over the right hemisphere during this late time window. The three-way interaction of group x hemisphere x complexity was not statistically

significant (F[1, 19] = 1.17, p=0.293, partial η^2= 0.06). The main effect of complexity was not significant (F[1, 19]=0.89, p=0.357, partial η^2=0.05) and complexity did not interact with group (F[1, 19]=0.81, p=0.379, partial η^2=0.04). However, there was a significant group x hemisphere interaction (F[1, 19]=6.80, p=0.017, partial η^2=0.26) indicating that, in this later time window, the groups engaged the left and right hemispheres differently. There was a significant complexity x hemisphere interaction (F[1, 19]=4.79, p=0.041], partial η^2=0.20) suggesting that complexity was processed differently over the two hemispheres. As can be seen in Figure 2, the TS and CAS groups show similar patterns over the left hemisphere during this late time window, but the two groups' waveforms diverge over the right hemisphere for both simple and complex words.

Response-locked analysis

Because the response times differed between the groups, we sought to confirm that this late time-window was truly reflecting similar planning/programming processes for both groups. Thus, an additional analysis was undertaken by using the acoustic response as an estimate of the onset of movement. We therefore analyzed the first TF identified by the PCA at −188 to 0 ms prior to the acoustic response.

Figure 2 shows waveforms from the same right hemisphere electrode identified by the second TF in the stimulus-locked analysis (Panel B) as well as in the response-locked analysis (Panel C). This is presented to visually compare the same electrode under both the stimulus-locked and response-locked analysis methods. It can be seen that a relatively similar pattern emerges (revealing group differences), enhancing our confidence that the late temporal window in the stimulus-locked condition reflects a waveform associated with pre-movement potentials.

Figure 3 depicts waveforms from a left and right hemisphere electrode identified by the spatial PCA. A slow negative wave is evident immediately prior to the acoustic response, with clear group distinctions in the −188 to 0 ms time window identified by the first temporal factor in the response-locked PCA. As can be seen from Figure 3, this response-locked analysis reveals a pattern of group differences (regardless of condition). The mean amplitudes from the response-locked waveform from −188 to 0 ms in all posterior electrodes identified by the spatial PCA were submitted to a 2 (group) x 2 (complexity) x 2 (hemisphere) ANOVA. Results indicated no significant interactions: no significant three-way interaction (F[1, 13] = 2.43, p = 0.143, partial η^2= 0.158), no significant complexity x group interaction (F[1, 13] = .495, p = 0.494, partial η^2=0 .037), no significant group x hemisphere interaction (F[1, 13] = 0.198, p = 0.664, partial η^2= 0.015), and no significant hemisphere x complexity interaction (F[1, 13] = 0.860, p = 0.371, partial η^2= 0.062). Additionally, there was no significant main effect of hemisphere (F[1, 13] = 1.02, p = 0.330, partial η^2=0.073), or complexity (F[1, 13] = 3.03, p =0.105, partial η^2= 0.189). However, there was a significant main effect of group (F[1,13] = 11.94, p = 0.004, partial η^2=0 .479), suggesting the amplitude of response prior to the onset of movement was different for the CAS and TS groups.

Discussion

The aim of this study was to identify specific neural characteristics associated with speech planning and programming differences in CAS. Because of the novel nature of the study, a combination of data-driven analysis (PCA) and theoretically-motivated selection of time windows was used to identify the components of interest. The PCA identified temporal windows that correspond reasonably well with those defined by Levelt and Indefrey's WPM of the processes involved during picture naming. The ERP responses observed in these time windows provide the first confirmation of the hypothesis that neurophysiological differences might exist in time windows associated with phonological and phonetic processing of complex words in children with CAS. The results provide a neurophysiological complement to prior behavioral studies which have used speech output to make inferences about disrupted processes that occur prior to speech output (Maas & Mailend, 2012; Maas, Robin, Wright, & Ballard, 2008; Maassen et al., 2001; Nijland & Maassen, 2003; Shriberg et al., 2012).

Response times observed here are generally similar to those observed in prior studies. The TS group mean response times were as expected for this type of task (i.e., approximately 675 ms), though they were approximately 10% slower than adults on this same task (see Table 1 footnotes). Given the presumed impairments in motor planning and programming, slower response times for the CAS group are not unexpected. The lack of complexity effect as measured by response times is in line with prior research indicating that longer/more complex words are not necessarily named slower than simpler items when items are presented in a randomized fashion, as they were here (see Meyer et al., 2003).

The group-by-complexity interaction over the right hemisphere in the early time window (300–460 ms post picture onset) is in line with clinical observations that differences arise in the generation of longer word forms in children with CAS, as they typically show significantly greater difficulty producing complex words (e.g., butterfly) than simple words (e.g., bone) (e.g., Peter et al., 2013). Neurophysiologically, the difference with processing complex words was associated in this study with reductions in amplitude of activity over the right hemisphere for complex words (relative to simple words) in the CAS group. There was no complexity effect for the TS group over either hemisphere but complex words were processed differently across hemispheres for the CAS group. The observed differences may reflect a reduction in neural activity during the generation of complex phonological word forms. One possible psycholinguistic explanation is that the complex multisyllabic utterances require the imposition of lexical stress, which is known to differ in children with CAS (Ballard, Djaja, Arciuli, James, & van Doorn, 2012; Ballard, Robin, McCabe, & McDonald, 2010; Shriberg et al., 2003). Alternatively, it may be the challenge of planning complex sequences of sounds and syllables that underlies this difference during the planning window (Peter et al., 2013). It should also be noted that, based on the WPM, the time window encompasses both late stages of phonological retrieval as well as syllabification; because there is only one single, slow wave in the ERP signal at this time point, the paradigm used here may not be able to fully distinguish between retrieval and syllabification (which, as noted above, may temporally overlap). Thus a phonological component to CAS may exist, particularly as it relates to constructing complex phonological word forms.

The differences in the later time window were also evident only over the right hemisphere. These group differences were not dependent upon complexity but provided evidence that children with CAS have detectably different electrophysiological activity over the right hemisphere in the later stages of speech preparation. If this indeed reflects phonetic stages of motor planning or programming, the implication may be that children with CAS are recruiting different neuronal populations to control the articulators for speech production. In the response-locked analysis, a large number of posterior electrodes were identified and the group difference was observed regardless of hemisphere. Thus, the magnitude of group differences may vary somewhat depending on the nature of the analysis (stimulus-locked or response-locked) or the location of the recordings across the scalp, but the differences between groups are detectable at the group level even with relatively small samples. Because group differences were observed using both data analytic methods, this enhances the reliability of the findings of group differences just prior to speech production.

Although the study could be viewed as primarily descriptive (i.e., confirming that neurobiological differences exist in preparing for speech), the identification of objective, brain-based measures associated with phonological processes (word complexity) and phonetic processes (issuing commands prior to speech onset) now offers guidance for studies comparing neurophysiological responses in CAS and other subtypes of speech and language disorders. For example, children with residual articulation errors who do not have symptoms of CAS would not be expected to differ from children with TS in this type of ERP paradigm. With further refinement of the paradigm, it is possible that an ERP task such as the one described here could eventually serve as an objective, theoretically motivated measure for differential diagnosis of CAS versus other subtypes of speech sound disorders.

Caveats and Limitations

As is the case for most studies of low-incidence populations such as CAS, the sample size is limited. The significant interactions observed here are therefore likely to be robust statistical effects. Future studies should explore modifications to the existing paradigm that could enhance the effect further to be sufficiently reliable at both the group level and the individual level. For example, the complex words included here all began with a stressed syllable, but words beginning with weak-strong stress patterns (e.g., *computer*) may be more challenging for children to learn (Ballard et al., 2012). The magnitude of the complexity effect might therefore be enhanced by including items that begin with weak-strong stress patterns. Additional manipulations could be explored to further stress the speech production system and to enhance the group distinction (e.g., eliciting only 3–4 syllable words in the complex condition). Finally, the diagnostic utility of this paradigm could be explored, as modifications to the paradigm may help to reveal differences between children with CAS and other subtypes of developmental speech and language disorders.

At present, the regions responsible for group differences in syllabification, planning and programming remain speculative. However, the updated WPM model laid out by Indefrey (2011) may provide some insight related to the regions involved in specific processes. For example, the translation from linguistic to motor representations may involve communications between posterior superior temporal regions and inferior frontal regions,

which may rely on the arcuate fasciculus. Later stages of motor programming involve a number of cortical and subcortical regions including supplementary motor area, inferior frontal/precentral regions, and cerebellum. However, other neurobiological accounts of speech production may also be relevant to the present investigation. For example, the DIVA model has been used to characterize CAS as a disorder with reduced feed-forward control and increased reliance on feedback mechanisms (Terband, Maassen, Guenther & Brumberg, 2009; Terband & Maassen, 2010). The DIVA model presently does not specify the precise time-course of events in a manner similar to the WPM, making it difficult to compare predictions of the two models directly. Moreover, the differential reliance on feed-forward and feedback mechanisms in CAS could have implications for differential processing in the left and right hemispheres (Tourville & Guenther, 2011). However, it should be acknowledged that EEG data are recorded at the scalp and do not necessarily reflect neural activity directly below the electrodes. Whereas work by Liégeois et al (2003) observed reduced activation for family members with CAS compared to controls in the left inferior frontal gyrus, the primary differences observed in our study were over the right hemisphere. The present results are not necessarily incompatible with these results as the actual sources of the ERP signals are unclear. Future research could pair ERP data with technologies such as functional MRI to understand how spatially-sensitive hemodynamic data might inform (and be informed by) the temporally-sensitive ERP data. Moreover, follow up work that uses dipole modeling (e.g., LORETA) could also be utilized; however due to the lack of photogrammetry data and small sample size this was not utilized here.

Summary and Conclusions

This study adds to our understanding of CAS in several important ways. The data provide evidence of disrupted neurobiological responses in CAS during a speech production task. The study also reveals that the observed differences in CAS are a function of both task (i.e., the complexity of the items to be planned) and hemisphere (with reduced amplitudes over the right hemisphere but not the left). Finally, differences were observed in the time windows associated with both phonological processes (assembly of sound sequences and/or imposition of stress, which was associated with complexity of the word) and phonetic processes (generating a motor plan and/or transmission of motor commands). The evidence of differences in the brain's response during speech preparation sets the foundation for further pursuit of theoretically-driven research on neurophysiological markers of subtypes of speech impairment during speech production tasks.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding was provided by the Childhood Apraxia of Speech Association of North America (CASANA), the Language and Early Assessment Research Network (LEARN) at Haskins Laboratories, and NIH grant P01HD001994. Thanks to Emily Phillips and Erin Grohman for assistance with data collection.

References

ASHA. Childhood Apraxia of Speech [Technical Report]. 2007. Available from www.asha.org/policy

Ballard KJ, Djaja D, Arciuli J, James DGH, van Doorn J. Developmental trajectory for production of prosody: Lexical stress contrastivity in children ages 3 to 7 years and in adults. Journal of Speech, Language, and Hearing Research. 2012; 55(6):1822–1835.

Ballard KJ, Robin DA, McCabe P, McDonald J. A treatment for dysprosody in childhood apraxia of speech. Journal of Speech, Language, and Hearing Research. 2010; 53(5):1227–1245.10.1044/1092-4388(2010/09-0130)

Cattell RB. The scree test for the number of factors. Multivariate Behavioral Research. 1966; 1(2): 245–276.

Crary MA. Phonological characteristics of developmental verbal dyspraxia. Seminars in Speech and Language. 1984; 5(2):71–83.

Dien J. Issues in the application of the average reference: Review, critiques, and recommendations. Behavior Research Methods, Instruments, & Computers. 1998; 30(1):34–43.

Dien J. The ERP PCA Toolkit: An open source program for advanced statistical analysis of event-related potential data. Journal of Neuroscience Methods. 2010; 187(1):138–145. [PubMed: 20035787]

Eulitz C, Hauk O, Cohen R. Electroencephalographic activity over temporal brain areas during phonological encoding in picture naming. Clinical Neurophysiology. 2000; 111(11):2088–2097. [PubMed: 11068246]

Forrest K. Diagnostic Criteria of Developmental Apraxia of Speech Used by Clinical Speech-Language Pathologists. American Journal of Speech-Language Pathology. 2003; 12(3):376–380. [PubMed: 12971826]

Froud K, Khamis-Dakwar R. Mismatch negativity responses in children with a diagnosis of childhood apraxia of speech (CAS). American Journal of Speech-Language Pathology. 2012; 21(4):302–312. [PubMed: 22564903]

Goldman, R.; Fristoe, M. Goldman Fristoe Test of Articulation. 2. Circle Pines, MN: AGS; 2000.

Hayden, D.; Square, P. Verbal Motor Production Assessment for Children. San Antonio, TX: Psychological Corporation; 1999.

Indefrey P. The spatial and temporal signatures of word production components: a critical update. Frontiers in Psychology. 2011; 2:255. [PubMed: 22016740]

Indefrey P, Levelt WJM. The spatial and temporal signatures of word production components. Cognition. 2004; 92(1–2):101–144. [PubMed: 15037128]

Levelt WJ. Models of word production. Trends in Cognitive Sciences. 1999; 3(6):223–232. [PubMed: 10354575]

Levelt WJ, Praamstra P, Meyer AS, Helenius P, Salmelin R. An MEG study of picture naming. Journal of Cognitive Neuroscience. 1998; 10(5):553–567. [PubMed: 9802989]

Levelt WJ, Roelofs A, Meyer AS. A theory of lexical access in speech production. Behavioral & Brain Science. 1999; 22(1):1–38. discussion 38–75.

Lewis BA, Freebairn LA, Hansen AJ, Iyengar SK, Taylor HG. School-Age follow-up of children with childhood apraxia of speech. Language, Speech, and Hearing Services in Schools. 2004; 35:122–140.

Liégeois F, Baldeweg T, Connelly A, Gadian DG, Mishkin M, Vargha-Khadem F. Language fMRI abnormalities associated with FOXP2 gene mutation. Nature Neuroscience. 2003; 6(11):1230–1237.

Maas E, Mailend ML. Speech planning happens before speech execution: Online reaction time, ethods in the study of apraxia of speech. Journal of Speech, Language and Hearing Research. 2012; 55(5):S1523–1534.

Maas E, Robin DA, Wright DL, Ballard KJ. Motor programming in apraxia of speech. Brain and Language. 2008; 106(2):107–118. [PubMed: 18417200]

Maassen B, Nijland L, Van Der Meulen S. Coarticulation within and between syllables by children with developmental apraxia of speech. Clinical Linguistics & Phonetics. 2001; 15(1):145–150. [PubMed: 21269115]

McArdle JJ, Mari Z, Pursley RH, Schulz GM, Braun AR. Electrophysiological evidence of functional integration between the language and motor systems in the brain: A study of the speech bereitschaftspotential. Clinical Neurophysiology. 2009; 120(2):275–284. [PubMed: 19109058]

Meyer AS, Roelofs A, Levelt WJM. Word length effects in object naming: The role of a response criterion. Journal of Memory and Language. 2003; 48(1):131–147.

Nijland L, Maassen B. Evidence of motor programming deficits in children diagnosed with DAS. Journal of Speech, Language & Hearing Research. 2003; 46(2):437–450.

Nijland L, Maassen B, Van der Meulen S, Gabreels F, Kraaimaat FW, Schreuder R. Coarticulation patterns in children with developmental apraxia of speech. Clinical Linguistics & Phonetics. 2002; 16(6):461–483. [PubMed: 12469451]

Nijland L, Maassen B, van der Meulen S, Gabreels F, Kraaimaat FW, Schreuder R. Planning of syllables in children with developmental apraxia of speech. Clinical Linguistics & Phonetics. 2003; 17(1):1. [PubMed: 12737052]

Perrin F, Pernier J, Bertrand O, Echallier J. Spherical splines for scalp potential and current density mapping. Electroencephalography and Clinical Neurophysiology. 1989; 72(2):184–187. [PubMed: 2464490]

Peter B, Button L, Stoel-Gammon C, Chapman K, Raskind WH. Deficits in sequential processing manifest in motor and linguistic tasks in a multigenerational family with childhood apraxia of speech. Clinical Linguistics & Phonetics. 2013; 27(3):163–191. [PubMed: 23339324]

Picton T, Bentin S, Berg P, Donchin E, Hillyard S, Johnson R, Rugg M. Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. Psychophysiology. 2000; 37(2):127–152. [PubMed: 10731765]

Preston JL, Brick N, Landi N. Ultrasound biofeedback treatment for persisting childhood apraxia of speech. American Journal of Speech-Language Pathology. 2013; 22(4):627–643. [PubMed: 23813207]

Shriberg LD, Campbell TF, Karlsson HB, Brown RL, McSweeny JL, Nadler CJ. A diagnostic marker for childhood apraxia of speech: The lexical stress ratio. Clinical Linguistics & Phonetics. 2003; 17(7):549–574. [PubMed: 14608799]

Shriberg LD, Fourakis M, Hall SD, Karlsson HB, Lohmeier HL, McSweeny JL, Wilson DL. Extensions to the speech disorders classification system (SDCS). Clinical Linguistics & Phonetics. 2010; 24(10):795–824. [PubMed: 20831378]

Shriberg LD, Lohmeier HL, Strand EA, Jakielski KJ. Encoding, memory, and transcoding deficits in childhood apraxia of speech. Clinical Linguistics & Phonetics. 2012; 26(5):445–482. [PubMed: 22489736]

Sochůrková D, Rektor I, Jurák P, Stančák A. Intracerebral recording of cortical activity related to self-paced voluntary movements: A bereitschaftspotential and event-related desynchronization/synchronization. Seeg study. Experimental Brain Research. 2006; 173(4):637–649. [PubMed: 16544136]

Strand EA, McCauley RJ, Weigand SD, Stoeckel RE, Baas BS. A Motor speech assessment for children with severe speech disorders: Reliability and validity evidence. Journal of Speech, Language & Hearing Research. 2013; 56(2):505–520.

Terband H, Maassen B. Speech motor development in childhood apraxia of speech: Generating testable hypotheses by neurocomputational modeling. Folia Phoniatrica et Logopaedica. 2010; 62(3):134–142. [PubMed: 20424469]

Terband H, Maassen B, Guenther FH, Brumberg J. Computational neural modeling of speech motor control in childhood apraxia of speech (CAS). Journal of Speech, Language, and Hearing Research. 2009; 52(6):1595–1609.

Tourville JA, Guenther FH. The diva model: A neural theory of speech acquisition and production. Language and Cognitive Processes. 2011; 26(7):952–981. [PubMed: 23667281]

van der Merwe, A. A theoretical framework for characterization of pathological speech sensorimotor control. In: McNeil, M., editor. Clinical management of sensorimotor speech disorders. New York: Thieme; 2008.

Velleman SL, Shriberg LD. Metrical analysis of the speech of children with suspected developmental apraxia of speech. Journal of Speech, Language & Hearing Research. 1999; 42(6):1444–1460.

Williams, KT. Expressive Vocabulary Test. 2. Pearson; 2007.

Wohlert AB. Event-related brain potentials preceding speech and nonspeech oral movements of varying complexity. Journal of Speech and Hearing Research. 1993; 36(5):897–905. [PubMed: 8246478]

Appendix: Stimulus items

Simple items were CV or CVC, and complex items were at least two syllables and were allowed to contain consonant clusters. Each item was presented six times.

Simple	Complex
Boy	Basketball
Bee	Blueberry
Book	Band-Aid
Bone	Blanket
Bat	Battery
Bow	Butterfly
Dot	Dandelion
Dime	Dentist
D	W
Dog	Diamond
Duck	Donuts
Doll	Dinosaur

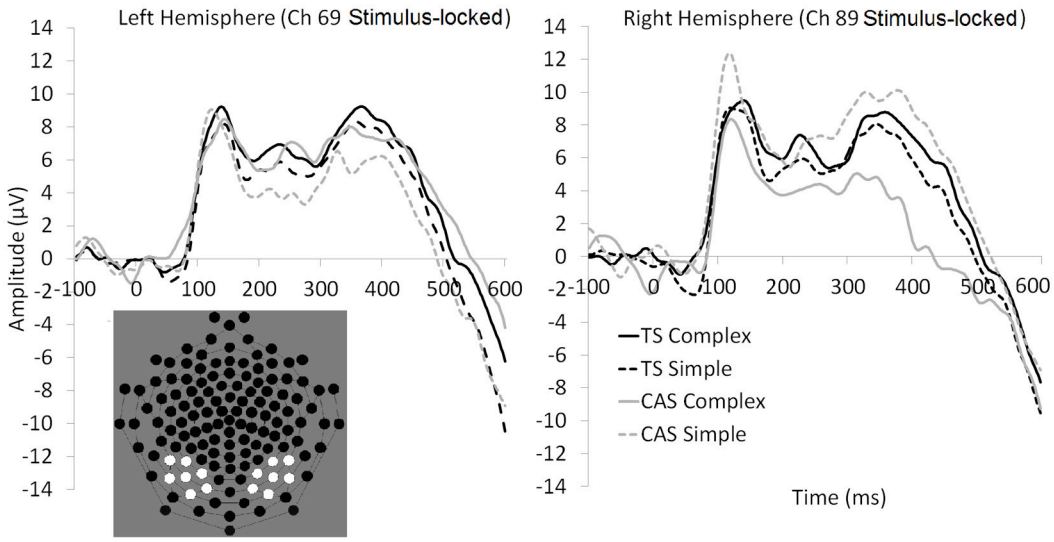

Figure 1.

Representative stimulus-locked ERP responses in posterior left and right hemispheres to simple and complex words for Typically Speaking (TS) and Childhood Apraxia of Speech (CAS) groups.

Note. Electrode map in bottom left corner shows selected electrodes (white) used in ANOVA testing the early time window (300–460 ms); these were chosen based on the third temporal factor from the principle components analysis. There was a significant group x hemisphere x condition interaction; this can be seen by a separation between simple and complex responses in the CAS group in the right hemisphere during this time window but very little separation for the TS group (see text).

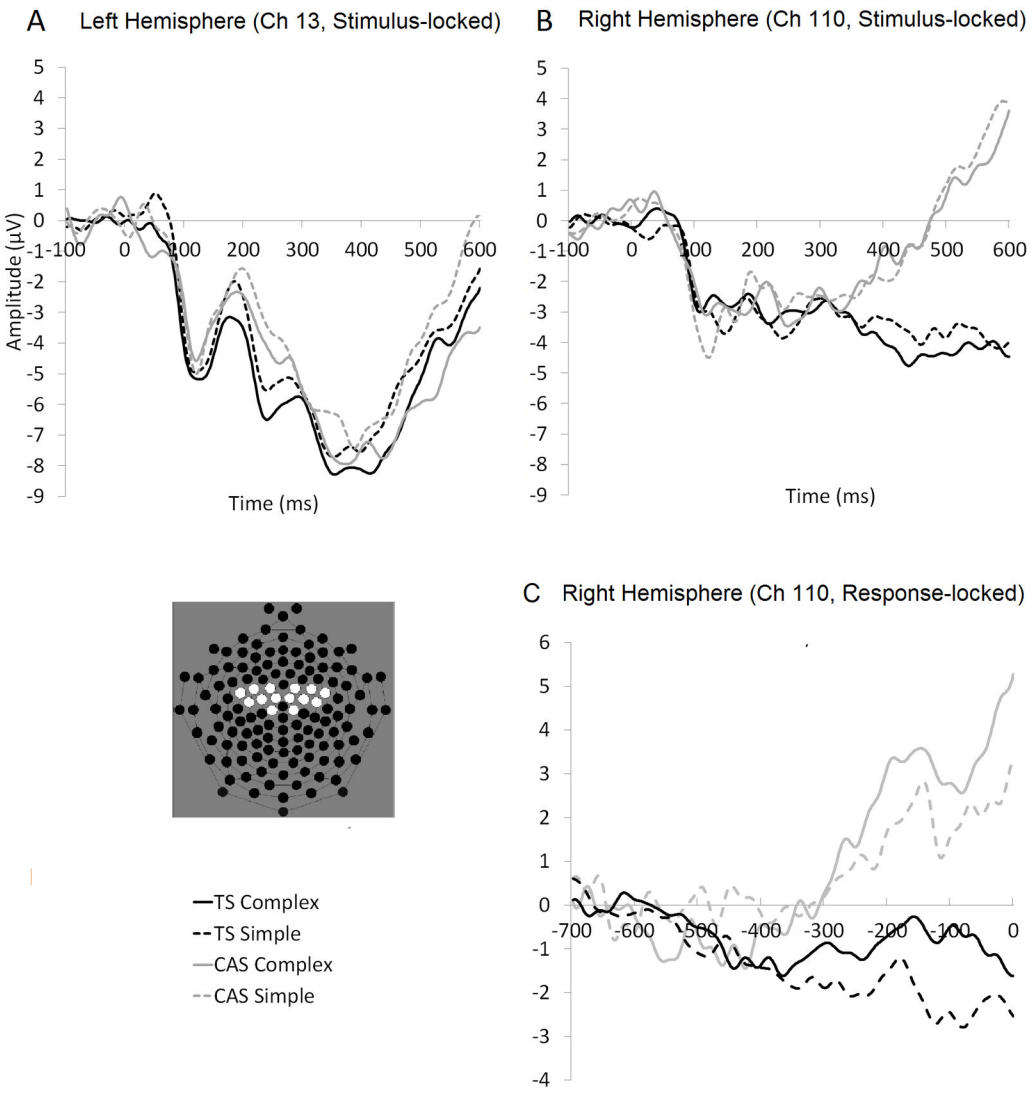

Figure 2.

Representative stimulus-locked ERP responses in anterior left and right hemispheres to simple and complex words for Typically Speaking (TS) and Childhood Apraxia of Speech (CAS) groups.

Note. Electrode map in top left shows selected electrodes (white) used in ANOVA testing the later time window (484–600 ms); these were chosen based on the second temporal factor from the principle components analysis. There was a significant group x hemisphere interaction in this late time window, which is depicted by the separation of responses between the groups in the right hemisphere but not in the left (see text).

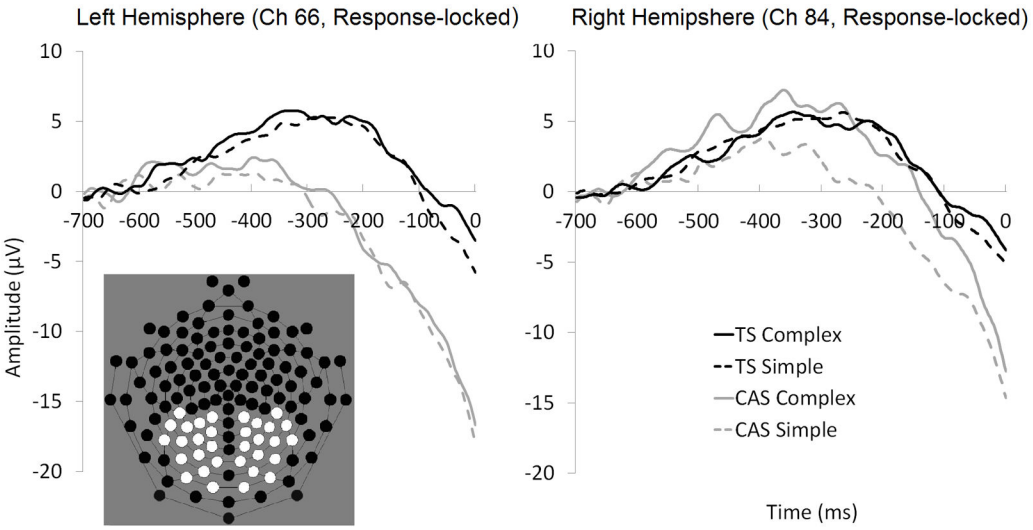

Figure 3.

Representative response-locked ERP responses in posterior left and right hemispheres to simple and complex words for Typically Speaking (TS) and Childhood Apraxia of Speech (CAS) groups

Note. Electrode map in bottom left shows selected electrodes (white) used in ANOVA testing the time window −188 to 0 ms prior to the acoustic response; these were chosen based on the first temporal factor from the principle components analysis. There was a significant main effect of group and no significant interactions in these electrodes (see text)

Table 1

Descriptive statistics for demographic and behavioral measures and ERP responses

	TS group mean (SD)	CAS group mean (SD)
Age	12 y 2 mo (23 mo)	12 y 7 mo (28 mo)
EVT-2 Standard Score	110.3 (8.5)	95.5 (18.3)
GFTA-2 Standard Score	101.7 (2.2)	57.9 (15.6)
Number of usable ERP trials (stimulus-locked analysis):		
Simple	33.1 (13.0)	30.3 (10.0)
Complex	31.9 (13.6)	30.4 (12.4)
Number of usable ERP trials (stimulus-locked analysis):		
Simple	43.2 (8.7)	28.7 (11.8)
Complex	40.8 (7.0)	31.0 (12.5)
Response time (ms)		
Simple	675.6 (102.2)	818.7 (80.6)
Complex	679.9 (92.8)	839.1 (75.6)

Notes: TS: Typical speech. CAS: Childhood apraxia of speech. ERP: Event-related potentials. EVT-2: Expressive Vocabulary Test-2 (Williams, 2007). GFTA-2: Goldman-Fristoe Test of Articulation-2 (Goldman & Fristoe, 2000). Not that, in a pilot study of 15 typically speaking adults ages 18–35, the average response time was 608 ms (SD 58) for simple items and 617 ms (SD 55) for complex items, suggesting that the TS group of children was approximately 65 ms (or 10%) slower than typically speaking adults.