

Marquette University

e-Publications@Marquette

Speech Pathology and Audiology Faculty
Research and Publications

Speech Pathology and Audiology, Department
of

9-2020

Differential Diagnosis of Apraxia of Speech in Children and Adults: A Scoping Review

Kristen M. Allison

Claire Cordella

Jenya Iuzzini-Seigel

Jordan R. Green

Follow this and additional works at: https://epublications.marquette.edu/spaud_fac

Marquette University

e-Publications@Marquette

Speech Pathology and Audiology Faculty Research and Publications/College of Health Sciences

This paper is NOT THE PUBLISHED VERSION.

Access the published version via the link in the citation below.

Journal of Speech, Language and Hearing Research, Vol. 63, No. 9 (September 2020): 2952-2994. [DOI](#). This article is © American Speech-Language-Hearing Association and permission has been granted for this version to appear in [e-Publications@Marquette](#). American Speech-Language-Hearing Association does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from American Speech-Language-Hearing Association.

Differential Diagnosis of Apraxia of Speech in Children and Adults: A Scoping Review

Kristen M. Allison

Department of Communication Sciences and Disorders, Northeastern University, Boston, MA

Claire Cordella

Frontotemporal Disorders Unit, Department of Neurology, Massachusetts General Hospital, Boston, MA

Jenya Iuzzini-Seigel

Department of Speech Pathology and Audiology, Marquette University, Milwaukee, WI

Jordan R. Green

Department of Communication Sciences and Disorders, MGH Institute of Health Professions, Boston, MA

Program in Speech and Hearing Bioscience and Technology, Harvard University, Boston, MA

Abstract

Purpose: Despite having distinct etiologies, acquired apraxia of speech (AOS) and childhood apraxia of speech (CAS) share the same central diagnostic challenge (i.e., isolating markers specific to an

impairment in speech motor planning/ programming). The purpose of this review was to evaluate and compare the state of the evidence on approaches to differential diagnosis for AOS and CAS and to identify gaps in each literature that could provide directions for future research aimed to improve clinical diagnosis of these disorders. **Method:** We conducted a scoping review of literature published between 1997 and 2019, following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews guidelines. For both AOS and CAS, literature was charted and summarized around four main methodological approaches to diagnosis: speech symptoms, quantitative speech measures, impaired linguistic- motor processes, and neuroimaging. **Results:** Results showed that similar methodological approaches have been used to study differential diagnosis of apraxia of speech in adults and children; however, the specific measures that have received the most research attention differ between AOS and CAS. Several promising candidate markers for AOS and CAS have been identified; however, few studies report metrics that can be used to assess their diagnostic accuracy. **Conclusions:** Over the past two decades, there has been a proliferation of research identifying potential diagnostic markers of AOS and CAS. In order to improve clinical diagnosis of AOS and CAS, there is a need for studies testing the diagnostic accuracy of multiple candidate markers, better control over language impairment comorbidity, more inclusion of speech-disordered control groups, and an increased focus on translational work moving toward clinical implementation of promising measures.

Introduction

Differential diagnosis of apraxia of speech in adults and children continues to be a major clinical and research challenge, despite decades of research. Although acquired apraxia of speech (AOS) and childhood apraxia of speech (CAS) have distinct etiologies, both disorders are presumed to be defined by difficulties with motor planning and programming of speech movements. In the absence of biomarkers for AOS/CAS, behavioral phenotyping based on clinical symptomatology remains the "gold standard" for the diagnosis of both disorders. Although disruptions in articulation and prosody are among the most commonly cited speech symptoms associated with both AOS and CAS, consensus is lacking about the core speech symptoms, most sensitive diagnostic criteria, and best assessment protocols. The absence of pathognomonic speech features has led to multiple clinical and research challenges, including inaccurate and delayed diagnosis (Basilakos, 2018; Forrest, 2003; McNeil et al., 2004; Mumby et al., 2007), leading to difficulties identifying the most appropriate treatment approach. The resulting challenges with differential diagnosis have posed obstacles to research focused on identifying speech apraxia biomarkers and the biological mechanisms of apraxia (e.g., genetic, neurological, physiological).

Why Is Differential Diagnosis of AOS So Challenging?

Differential diagnosis of speech apraxia in both adult and pediatric populations relies on the identification of symptoms that are sensitive and specific to apraxia and can, therefore, separate apraxia from both a higher level language impairment (i.e., aphasia in adults, phonological disorders in children) and a lower level impairment in motor execution (i.e., dysarthria). Achieving agreement on operationally defined criteria for diagnosing AOS and CAS has been an ongoing focus of research and source of disagreement. Several factors have contributed to the difficulty with establishing diagnostic

criteria, including debates surrounding theoretical models of AOS, overlap in symptomatology among speech disorders, and the frequency of comorbidities that also influence speech patterns.

Theoretical Models of Apraxia of Speech

Several theoretical frameworks have been proposed to explain the deficient neural processes that give rise to apraxia of speech. Some models of acquired AOS attribute the disorder to a breakdown in translating encoded phonological representations to articulated speech, which is considered the planning/programming stage of speech production. Linguistic models, such as the one proposed by Levelt (1992), conceptualize this breakdown as part of a serial processing model, specifically affecting the construction of an accurate phonetic plan (i.e., phonetic encoding). Although in theory differentiable, these model stages are not easily clinically observed (Maassen, 2002). For this reason, some researchers have argued for a conceptualization of apraxia of speech that focuses more on dynamic interactions of linguistic and motor speech processes (Ziegler et al., 2012). In fact, recent computational models have emerged that integrate linguistic and motor speech processes (Guenther et al., 2006; Levelt et al., 1999; Tourville & Guenther, 2011) and serve as a basis for making specific predictions about neuroanatomic correlates to speech production processes. The Directions Into Velocities of Articulators (DIVA) is one such example of neuroanatomically specific computational account of speech production (Guenther, 2016; Guenther et al., 2006; Tourville & Guenther, 2011). The DIVA model emphasizes the importance of integrated feedforward and feedback commands in speech production and theorizes that apraxia of speech can result from weak feedforward commands, resulting in overreliance on feedback. The DIVA model and other computational models, including the State Feedback Control (Houde & Nagarajan, 2011) and Hierarchical State Feedback Control (Hickok, 2012) models, have served as the theoretical framework for several behavioral paradigms aimed at testing hypotheses of feedforward versus feedback deficits (Iuzzini-Seigel et al., 2015; Maas et al., 2015; Parrell et al., 2017). A recent review of behavioral, computational, and imaging studies of AOS concluded that the integration of evidence across these different levels of analysis is critical for understanding underlying neural mechanisms and how they manifest as clinical symptoms (Ballard, Tourville, & Robin, 2014).

Isolating an impairment in motor planning/programming is even more challenging in children with a congenital speech disorder, as the presence of the motor speech disorder influences children's development of phonological representations (Stackhouse & Wells, 1997). Thus, linguistic models of apraxia of speech are further underspecified for children (Maassen, 2002). The DIVA model was developed in a way that accounts for development, and it has been used to model the symptoms of CAS (Terband et al., 2009). Similar to AOS, results of this model suggest that CAS symptoms can result from weak feedforward commands during development, which authors hypothesize could be due to reduced somatosensory information or increased neural noise (Terband et al., 2009, 2014). In practice, however, these hypothesized breakdowns in processing require careful experimental design to test and the clinical implications of this work will need to be explored in future translational work that focuses on assessment and intervention.

Overlap in Speech Disorder Phenotypes

Another primary challenge to generating clear diagnostic criteria is that many speech characteristics associated with apraxia also occur in other speech disorders. Although core diagnostic features of both

CAS and AOS involve disruptions in prosody, speaking rate, and segmental accuracy, many of these features are not unique to apraxia and can also occur in dysarthria and/or phonological disorders. Slow rate, atypical prosody, and sound distortions, including vowel errors, are common characteristics of apraxia and dysarthria in both child and adult populations (American Speech-Language-Hearing Association [ASHA], 2007; Duffy, 2013; Haley et al., 2017; McNeil et al., 2009; Strand et al., 2014; Wambaugh et al., 2006; Wertz et al., 1984). Segmental errors, including substitutions and omissions, are also considered core features of AOS and CAS (ASHA, 2007; McNeil et al., 2009; Strand et al., 2014; Wambaugh et al., 2006; Wertz et al., 1984) but can present very similarly to phonemic paraphasias associated with aphasia in adults or phonological speech sound errors in children. Determining whether segmental errors are phonological versus apraxic in origin has been considered more clinically challenging than distinguishing between apraxia of speech and dysarthria. Dysarthria often involves impairments in respiration, phonation, and/or resonance in addition to articulation, which result in global distortions of the acoustic signal that are not typically present in apraxia or phonological disorders. Overall, the overlap in speech disorder phenotypes suggests that diagnostic features are likely to be sensitive but not specific.

Furthermore, differential diagnosis relies on the assumption that AOS/CAS is either present or absent; however, the specific speech characteristics exhibited by individuals are widely variable. Current clinical diagnosis is based on a speaker presenting with some but not necessarily all possible symptoms of AOS/CAS. This variability in individual speech presentations also adds to the challenges with relying on specific symptoms or speech features for reliable diagnosis.

Comorbidity

Another significant challenge to developing objective diagnostic criteria for AOS and CAS has been the high frequency of comorbidities associated with both disorders. Aside from neurodegenerative cases of pure progressive AOS, AOS most commonly occurs alongside concomitant aphasic deficits following a left hemisphere stroke (Duffy, 2013; Graff-Radford et al., 2014). Likewise, CAS frequently occurs in conjunction with language impairment (Murray et al., 2019; Shriberg et al., 1999) and fine/gross motor deficits (Iuzzini-Seigel, 2019; Knežević, 2019; Teverovsky et al., 2009; Tükel et al., 2015). Therefore, finding individuals with AOS or CAS who do not have concomitant impairments is challenging and further contributes to difficulties isolating diagnostic features specific to apraxia. Given the difficulties with relying on behavioral phenotypes to diagnose apraxia of speech, there is a need for identifying diagnostic markers that can be used to increase accuracy and reliability of diagnosis. The purpose of this review was to explore and describe the evidence related to diagnostic markers of AOS and CAS.

What Makes a Good Diagnostic Marker? Look to New Standards for Diagnostic Test Accuracy

The accuracy of a differential diagnostic marker is the degree to which the measure accurately discriminates between individuals with the target disorder (AOS or CAS) and either normal controls or another disorder that is often confused with the target disorder (e.g., dysarthria for AOS or speech sound disorder [SSD] for CAS). Although guidelines for evaluating and reporting diagnostic accuracy are now well established, few research studies on speech apraxia have adhered to these standards (e.g., Bossuyt et al., 2003; Moher et al., 2015; Whiting et al., 2003). These standards have been advanced to

accelerate the pathway for establishing the levels of evidence needed to validate a candidate diagnostic marker.

The successful clinical integration of a speech apraxia marker will require evidence of its "analytical" validity (including tests of its discriminative accuracy, reproducibility, and reliability) and its "clinical" validity and utility (i.e., practical, reduces costs, and provides better analytic validity than current best practices for speech diagnostics). Analytical validity is established by testing the discriminative accuracy of a candidate marker, also called the index test, relative to that of a reference standard. The reference standard is the best available method for establishing the presence or absence of the target condition, which, for speech apraxia, is clinician-based expert diagnosis. Discriminative accuracy of a marker can be assessed using a variety of metrics, including sensitivity and specificity, likelihood ratios, positive and negative predictive values, diagnostic odds ratio, area under the receiver operating characteristic curve, and Youden's index (Šimundić, 2009). These metrics are commonly evaluated by comparing the sensitivity and specificity of a diagnostic index test to that of an established clinical reference standard in the same patient cohorts. Within a single study, confidence intervals around estimates of accuracy can be calculated to quantify the statistical precision of the measurements. Rigorous evaluation needs to include detailed information about the clinical context and the cohort because the accuracy of an index test is not constant but varies across different clinical contexts, disease spectrums, and even patient subgroups (Bossuyt et al., 2015).

Objectives of the Current Study

The primary goal of this review was to evaluate the state of the evidence on approaches that have been studied to improve differential diagnosis of apraxia in both adults with AOS and children with CAS. We chose to include both AOS and CAS in the review because the central diagnostic challenge is the same for both populations (i.e., to isolate markers specific to an impairment in speech motor planning/programming), and we hoped that a direct comparison of these literatures would help identify gaps in each and provide directions for future research. Our approach to this review was guided by the following questions: (a) What experimental approaches have been used in the literature to improve differential diagnosis of AOS in children and adults, and what is the state of the evidence for different approaches? and (b) What are the similarities and differences between the AOS and CAS literatures in terms of the state of the evidence for approaches to differential diagnosis?

We chose to conduct a scoping review because its format best matched our primary objectives, "to evaluate the extent, range, and nature" of evidence and to "identify research gaps in the existing literature" on the topic of differential diagnosis of AOS (Arksey & O'Malley, 2005, p. 21). Scoping reviews, first described by Arksey and O'Malley (2005), differ from systematic reviews in that they are designed to address a broadly focused research question, rather than a specific research question as is typically the aim of systematic reviews (Arksey & O'Malley, 2005; Levac et al., 2010).

Method

For this review, we followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews guidelines developed by the Enhancing the Quality and Transparency of Health Research Network (Tricco et al., 2018).

Eligibility Criteria

To be included in the review, articles had to focus on diagnosis of CAS or AOS and specifically on isolating diagnostic characteristics of apraxia of speech in either of these populations. Peer-reviewed articles were considered for inclusion if they (a) were published in the past 22 years (between 1997 and 2019), (b) were written in English, and (c) used a group design and included a group of participants with CAS or AOS. We focused on studies in the past two decades because definitions of CAS and AOS have evolved considerably over that time, and technology/quantitative methods to aid in diagnosis have also changed considerably. To narrow down the scope of our search, we focused specifically on group studies that related to apraxia diagnosis; thus, we excluded articles if they were (a) treatment studies; (b) case studies; (c) qualitative studies; (d) commentaries, opinion, or review articles; (e) animal studies; (f) not focused on CAS/AOS (e.g., nonverbal apraxia, syntax/semantics, cognitivecommunication); (g) focused on participation outcomes or longitudinal outcomes; or (h) focused on CAS associated with specific genetic, metabolic, or neurodevelopmental conditions (e.g., galactosemia, autism, cri du chat). We decided to exclude articles specifically focused on these complex neurodevelopmental disorders in order to maximize comparability between the child and adult literature.

Search and Selection of Sources of Evidence

To locate potential articles for inclusion, we searched several major databases: Harvard University Library's HOLLIS+ database (includes PubMed, PsycINFO, ERIC, Web of Science, Google Scholar, ScienceDirect), ASHAWire, and PubMed Central. Per specified eligibility criteria, we searched for peer-reviewed journal articles in English published between January 1, 1997, and November 11, 2019, using the search terms "diagnosis" AND ("apraxia of speech" OR "childhood apraxia of speech" OR "developmental apraxia of speech"). Besides database queries, additional sources of evidence included reference lists of included articles (see Figure 1).

The screening process to determine eligibility of returned articles was completed using a three-step sequential approach. The first step was a "title screen" by the first and second authors (K. A., C. C.), completed jointly and thus yielding a nondiscrepant list of included/excluded titles. Articles were excluded at this stage if the title indicated fulfillment of any exclusion criteria (e.g., treatment studies, qualitative studies); articles were retained if the title did not make it possible to evaluate whether inclusion/exclusion criteria were met. All articles surviving the "title screen" were subject to an "abstract screen," which was conducted independently by authors K. A. and C. C., with any discrepancies resolved through consensus. At this stage, articles were excluded based on the above-described exclusion criteria; in addition, articles were excluded if the abstract indicated that the article was a nongroup design study and/ or did not include a control group, or did not pertain specifically to diagnosis or differential diagnosis of AOS/CAS or experimentally valid distinctive features. Articles that remained following the "abstract screen" underwent a "fulltext screen" to ensure that they did, in fact, satisfy all inclusion/exclusion criteria. Articles surviving the "full-text screen" constituted the included sources of evidence that were subsequently charted and summarized.

Data Charting Process and Synthesis of Results

Data charting spreadsheets were jointly developed by the first and second authors to determine which information to extract from each included article. Charting was completed by research assistants and

revised by the first and second authors for accuracy. Any points of confusion were resolved through discussion and consensus. For each included source of evidence, we extracted and charted the following information: (a) methods used (independently coded by authors K. A. and C. C., with discrepancies resolved through consensus), (b) participant groups included, (c) primary dependent variables of interest, (d) main results, and (e) whether sensitivity/specificity was reported.

For a subset of included articles, we charted an additional item, inclusion criteria for CAS/AOS participant group, which we operationally defined as the clinical features or signs met by individual participants (and reported by the authors) in order to be considered part of the study's CAS/AOS diagnostic cohort. We defined the subset of articles as those published since 2007, the year that ASHA published its position statement on the diagnosis of CAS (ASHA, 2007). In addition to the core CAS features listed in the ASHA position statement, a more recent comprehensive list of diagnostic features proposed by Shriberg et al. (2011) and a list of features proposed by Davis and Velleman (2000) were cited in multiple studies and included in charting of these studies. Although no comparable position statement has been issued by ASHA for AOS, Wambaugh et al. (2006) proposed a candidate list of AOS diagnostic features. This effort was important for the ongoing efforts to build consensus among researchers and clinicians on core diagnostic AOS features. Since 2006, other feature lists have been proposed by McNeil et al. (2009) and most recently by Strand et al. (2014) as part of the Apraxia of Speech Rating Scale (ASRS). Historically, inconsistency in inclusion criteria has been a weakness of both the child and adult apraxia literature; however, with increasing consensus on diagnostic behavioral features in the past decade, consistency of inclusion criteria has improved. The charting and analysis of this additional data item are meant to provide a way of quantifying consistency in use of inclusion criteria across multiple studies, as well as to provide additional context for comparing the literature on adult and child populations.

Results

Selection and Characteristics of Sources of Evidence

Across all three databases, 1,254 nonduplicate citations met search criteria and were subsequently reviewed (by authors K. A. and C. C.) based on title only. One thousand seven articles were excluded based on the title-only screen. Relevant review articles were separated out at this stage and excluded for purposes of charting, although a subset of these reviews were used for general background information. A total of 247 nonreview articles passed the title-only screening stage, and this list was used for the subsequent abstract review stage. Following abstract review, an additional 157 articles were excluded. Interrater agreement on article inclusion/exclusion was 89% for the abstract review stage, and all disagreements were resolved by consensus. The remaining articles ($n = 90$) underwent full-text review, and all were determined to satisfy inclusion/exclusion criteria, meaning that a total of 90 articles were charted and summarized. The full process of selecting sources of evidence is detailed in Figure 1.

Articles were grouped into categories based on whether they focused on CAS ($n = 37$) or AOS ($n = 53$). In addition, we subcategorized articles into one of four main content categories based on the methodological approach used for diagnosis: (a) speech symptoms ($n = 27$), (b) quantitative speech measures ($n = 27$), (c) impaired linguistic-motor processes ($n = 17$), and (d) neuroimaging ($n =$

19), described in detail in Table 1. Results are presented in accordance with these content categories as a way to summarize the literature associated with each methodological approach. If more than one methodological approach was used in a single study, a primary content category was nonetheless assigned by consensus of the first and second authors, based on the stated aims and goals of the study. Tables 2 through 9 present data for each article according to these groups and are also summarized in narrative form. For each article, we charted the five primary data items, described above in the Method section (i.e., methods used, participant groups included, primary dependent variables of interest; main results; sensitivity/ specificity). For CAS articles, we also charted the age ranges studied. To compare the inclusion criteria used in AOS and CAS studies since 2007, a comprehensive list of inclusion criteria was generated, and the criteria used for each study were charted (see Figure 2 and Appendixes A and B).

Differential Diagnosis Based on Speech Symptoms

A substantial number of articles ($n = 27$) focused on using surface speech characteristics as a method for differential diagnosis of AOS and CAS. The focus of these articles was to better describe the phenotypical features of AOS and CAS, using procedures relying on perceptual or clinical judgment. Methods used in these studies included phonetic transcription, perceptual judgment of speech characteristics, and quantitative analysis of error patterns (e.g., place/ manner/voicing errors, token-to-token inconsistency). Studies largely focused on the identification of core surface features that reliably differentiated individuals with AOS or CAS from individuals without apraxia and could be used to improve differential diagnosis in clinical settings.

CAS

Fifteen of the included articles focused on use of surface speech characteristics in diagnosis of CAS (see Table 2). The majority of these 15 articles used phonetic transcription and various analyses of segmental accuracy error patterns to describe surface speech characteristics ($n = 8$). Coding of prosody or lexical stress ($n = 5$) and clinical ratings of speech features ($n = 5$) were also common. The majority of CAS studies in the speech symptoms category included a comparison group of children with SSDs ($n = 12$). Two studies additionally included a comparison group of children with language impairment. Two studies included only a comparison group of typically developing (TD) children, and none of the studies included a dysarthria comparison group. Most of the studies focused on preschool or schoolage children ($n = 11$), but four studies focused on early speech characteristics of children later diagnosed with CAS (Highman et al., 2008; Overby, Belardi, & Schreiber, 2019; Overby & Caspari, 2015; Overby, Caspari, & Schreiber, 2019).

Overall, most studies in this category reported reduced segmental accuracy and/or greater error inconsistency in children with CAS compared to control groups (Aziz et al., 2010; Iuzzini-Seigel et al., 2017; Keske-Soares et al., 2018; Murray et al., 2015; Thoonen et al., 1997; Velleman & Shriberg, 1999). Prosodic deficits or lexical stress errors were also reported to differentiate children with CAS from control groups in several studies (Aziz et al., 2010; Murray et al., 2015; Shriberg et al., 1997a, 1997b); however, one study reported that lexical stress errors were similar between children with suspected CAS and children with other SSDs (Velleman & Shriberg, 1999). Task complexity was found to influence group differences on transcription-based measures (Iuzzini-Seigel et al., 2017; Murray et al., 2015; Strand et al., 2013; Thoonen et al., 1997). The four studies examining early speech features of children

later diagnosed with CAS showed that possible early signs of CAS include reduced babbling, smaller phonetic inventory, limited syllable structure, and fewer resonant sounds (Highman et al., 2008; Overby, Belardi, & Schreiber, 2019; Overby & Caspari, 2015; Overby, Caspari, & Schreiber, 2019). Four of the included studies in the speech symptoms category reported diagnostic accuracy statistics related to outcome measures (Iuzzini-Seigel et al., 2017; Murray et al., 2015; Shriberg et al., 1997a; Strand et al., 2013): Iuzzini-Seigel et al. (2017) reported high sensitivity (70%) and specificity (80%) of token-to-token inconsistency for differentiating children with CAS from children with other SSDs and those with language impairment, particularly in monosyllabic words or at the phrase level (i.e., repeated production of "buy Bobby a puppy"). Murray et al. (2015) reported that a statistical model, including four perceptually derived speech measures (i.e., syllable segregation, lexical stress matches, percentage phonemes correct in polysyllabic words, and articulatory accuracy during /pataka/), had high diagnostic accuracy (91%) for differentiating CAS from other SSDs. A validation study of the Dynamic Evaluation of Motor Speech Skill (Strand et al., 2013) demonstrated high specificity (97%) and moderate sensitivity (65%) for diagnosis of CAS.

AOS

Twelve of the included articles focused on use of surface speech characteristics in diagnosis of AOS, either in poststroke ($n = 8$) or progressive aphasia ($n = 4$) populations (see Table 3). The AOS group of interest had comorbid aphasic deficits in all but one study (Strand et al., 2014). Ten of the 12 articles in this category included an aphasia-only disease control group. Four articles reported results from neurologically healthy, age-matched controls. Only two studies (Jonkers et al., 2017; Ziegler, 2002) included a dysarthria comparison group; an additional five studies reported on the incidence of comorbid dysarthria in the AOS group of interest.

The majority of articles in this category relied on phonetic transcription to derive error counts and to characterize types of errors ($n = 7$). Clinician rating of errors was also common ($n = 4$), with errors characterized in terms of overall count, type (e.g., distortion vs. substitution), and consistency. A limited number of studies included secondary acoustic ($n = 2$) or imaging evidence ($n = 2$).

Overall, results from this category of studies indicate that individuals with aphasia and AOS make a greater number of production errors compared to aphasia-only populations and healthy controls (Ash et al., 2010; Bislick et al., 2017; Croot et al., 2012; Cunningham et al., 2016; Haley et al., 2012, 2017). The majority of articles characterized the observed production errors as predominantly phonetic (cf. phonemic); however, one article reported results contrary to this trend, with phonemic errors being more common than phonetic errors in the AOS group (Ash et al., 2010). Characterization and/or description of suprasegmental speech features (e.g., sentence scanning index) was less common, and for articles reporting such measures, no significant differences emerged between AOS and control groups (Haley et al., 2012). When compared to a dysarthria control group, however, individuals with AOS were reported to have greater syllable isochrony (Ziegler, 2002). Likewise, articles reporting on error consistency generally found no significant between-groups differences on such measures (Bislick et al., 2017; Haley et al., 2013, 2012).

A subgroup of articles ($n = 8$) in this category reported on the reliability of either speech-language pathology perceptual ratings or speech-language pathology phonetic transcriptions. Two primary types of speech-language pathology perceptual ratings were reported: (a) gestalt clinician ratings (i.e., no

operationalized speech features to guide clinician ratings) and (b) operationalized ratings, whereby clinicians were asked to rate specific aspects of speech (e.g., sound distortions, rate). Results were mixed in terms of whether gestalt clinician ratings yielded reliable diagnoses of AOS, with three studies indicating high reliability across raters (Bislick et al., 2017; Duncan et al., 2019; Mumby et al., 2007) and another indicating low overall reliability of gestalt ratings (Haley et al., 2012). Studies reporting on operationalized metrics, as opposed to/in addition to gestalt impressions, generally reported high levels of interrater agreement on apraxic features (Haley et al., 2012; Jonkers et al., 2017; Strand et al., 2014). Of particular note, Strand et al. (2014) outlined 16 diagnostic features of AOS with good to excellent interrater reliability that together comprise the ASRS, a partially standardized assessment of AOS.

Only one study (Croot et al., 2012) reported on the diagnostic accuracy of specific measures for identifying AOS. In this study, Croot et al. (2012) demonstrated that apraxic-type errors (i.e., phonetic distortions, syllable segregation, equal/excess stress) observed during a polysyllable word repetition task had high sensitivity (89%) for identifying individuals with progressive AOS and successfully differentiated these individuals from an aphasia-only group.

Differential Diagnosis Based on Quantitative Speech Measures

An equally large number of studies ($n = 27$) focused on quantifying surface features that have been associated with apraxia of speech through objective acoustic or kinematic measurements. These studies focused on identifying quantitative markers of CAS or AOS that may be more sensitive and reliable than perceptual measures and have the potential to establish more empirical criteria for apraxia diagnosis.

CAS

Twelve of the included studies examined the use of quantitative speech measures for aiding in diagnosis of CAS (see Table 4). Quantitative methods used included acoustic measures ($n = 11$) and articulatory kinematic measures ($n = 5$). The majority of studies in this category included a control group of children with other SSDs ($n = 10$), although several only included a TD control group ($n = 4$). Only one study included a comparison group of children with dysarthria. The majority of studies in this category focused on children between the ages of 3 and 10 years ($n = 10$); however, a series of studies by Shriberg and colleagues included a large sample of speakers with CAS ranging from 3 to 23 years (Shriberg et al., 2017a, 2017b, 2017c, 2017d).

Studies in this category demonstrated that several quantitative measures used to index core speech symptoms (i.e., coarticulation, motor variability, and prosody) differed between children with and without CAS. Three acoustic studies demonstrated that formant measures indexing anticipatory coarticulation differ between children with CAS and those with typical development (Maassen et al., 2001; Nijland et al., 2002; Nijland, Maassen, Van der Meulen, Gabreëls, et al., 2003). Two kinematic studies demonstrated that variability of lip and jaw movement signals across repeated productions of words and syllables was greater in children with CAS compared to children with typical development and SSD groups (Case & Grigos, 2016; Grigos et al., 2015). Two acoustic measures of lexical stress also differentiated children with CAS from children with typical development and other SSDs (Munson et al., 2003; Shriberg et al., 2003). One recent study (Kopera & Grigos, 2019) did not find acoustic differences in lexical stress in children with CAS compared to control groups but did find differences in

jaw kinematics reflecting reduced marking of lexical stress in the CAS group. Of the 12 studies, only two reported diagnostic accuracy statistics. Shriberg and colleagues demonstrated strong sensitivity (86.8%) and specificity (100%) of the Pause Marker, an acoustic-aided measure of appropriate pausing, for differentiating children with CAS from other SSDs (Shriberg et al., 2017a, 2017b). The only study to include a comparison group of speakers with dysarthria reported high sensitivity and specificity (ranging from 89% to 100%) of maximum performance tasks (i.e., maximum phonation duration, fricative duration, and diadochokinesis) for differentiating between dysarthria, CAS, and SSD (Thoonen et al., 1999); however, the inclusion criteria used for the CAS group in this study did not include prosodic errors or difficulty with articulatory transitions, which are now accepted core features of CAS (ASHA, 2007).

AOS

Fifteen of the included articles used quantitative speech features to aid in the diagnosis of AOS. A majority of these 15 articles focused on a poststroke population ($n = 12$), while a smaller number ($n = 3$) studied individuals with a progressive etiology (see Table 5). The AOS group had comorbid aphasic deficits in the vast majority of studies ($n = 13$) in this category. Aphasic status was unknown in one study (Patel et al., 2013), and only one study reported results from a pure (progressive) AOS group (Duffy et al., 2017). The majority of articles ($n = 10$) in this category included an aphasia-only disease control group. Most articles also reported results from additional control groups, including healthy individuals ($n = 11$) or other disease control groups (e.g., individuals with stroke but no aphasia, behavioral variant frontotemporal dementia; $n = 2$). Two studies included a dysarthria comparison group. A limited number of studies ($n = 2$) included secondary imaging evidence.

Studies in this category overwhelmingly used acoustic measures ($n = 14$) to quantify differences between groups in speech rate, lexical stress, and phonemic accuracy. The most commonly investigated acoustic measure was pairwise variability index for vowel duration and/or intensity; eight studies provided robust support for the utility of this measure to differentiate AOS groups from aphasia-only groups in both poststroke and progressive populations (Ballard et al., 2016; Ballard, Savage, et al., 2014; Basilakos et al., 2017; Courson et al., 2012; Duffy et al., 2017; Haley & Jacks, 2019; Scholl et al., 2018; Vergis et al., 2014). Specifically, results overwhelmingly indicated a reduced pairwise variability index for AOS groups. Several studies also found a reduced rate of speech—either in spontaneous speech or on maximum performance tasks (e.g., diadochokinetic rate)—among individuals with AOS as compared to individuals with aphasia only (Duffy et al., 2017; Melle & Gallego, 2012; Wilson et al., 2010); however, studies that also incorporated a dysarthria control reported results to indicate that rate alone may not distinguish AOS from dysarthria. Melle and Gallego (2012), for instance, report the alternating motion rate alone failed to distinguish AOS and dysarthria groups whereas the sequential motion rate did, thereby suggesting the importance of task in eliciting group differences.

Nontemporal acoustic variables were also evaluated across several studies, many of which focused on phonemic accuracy of vowels (Jacks et al., 2010) and consonants (Haley, 2002). In general, this group of studies found no evidence to support systematic differences in phonemic accuracy that could be uniquely attributed to AOS; however, results from Basilakos et al. (2017) report significant differences between AOS and disease control groups in consonantal production, as measured using the high-

frequency band of an envelope modulation spectrum. Other studies investigating variability measures including error variability (Scholl et al., 2018), voice onset time variability (Basilakos et al., 2017), and formant variability (Jacks et al., 2010; Melle & Gallego, 2012)-showed equivocal results: Studies generally reported greater error variability for AOS groups, but no between-groups differences were found between AOS and aphasia-only groups for measures such as voice onset time or formant variability (Basilakos et al., 2017; Jacks et al., 2010).

A single study (Bartle-Meyer et al., 2009) used kinematic, as opposed to acoustic, measures to report on articulatory coupling (i.e., the degree of coordination in movement between various articulators). Study results showed that coupling was greater for a majority of individuals with AOS compared to healthy controls; importantly though, this study did not include an aphasia-only control group.

Of the 14 articles in this category, five reported on diagnostic accuracy for several of the quantitative measures of interest (Ballard et al., 2016; Ballard, Savage, et al., 2014; Basilakos et al., 2017; Duffy et al., 2017; Scholl et al., 2018). Mirroring the group-level results, the pairwise variability index measure was shown to have good predictive value for AOS across several studies (Ballard et al., 2016; Ballard, Savage, et al., 2014; Duffy et al., 2017). In one of these studies (Ballard, Savage, et al., 2014), the authors assessed comparative diagnostic accuracy of several different acoustic measures and demonstrated a greater predictive value for the pairwise variability index for vowel duration as compared to the pairwise variability index for intensity, as well as measures of silence duration/variability in silence duration. Another of these studies (Ballard et al., 2016) compared the diagnostic accuracy of the pairwise variability index for vowel duration for different types of multisyllabic stimuli, namely, trisyllabic words with a weak-strong (e.g., "banana") versus strong-weak (e.g., "butterfly") stress pattern. The authors found that diagnostic accuracy was greater when the pairwise variability index was measured for multisyllabic words with a weak-strong stress pattern. Basilakos et al. (2017) reported very high classification accuracy for a comprehensive set of acoustic features, with measures of consonantal production (envelope modulation spectrum) accounting for the greatest single-variable contribution to overall accuracy. Two articles highlighted the importance of task-specifically the inclusion of longer multisyllabic words-in inducing errors that in turn demonstrate good diagnostic accuracy for AOS (Ballard et al., 2016; Duffy et al., 2017). Duffy et al. (2017), for example, demonstrate that diagnostic accuracy increases for trisyllabic word stimuli, such as "catastrophe" or "stethoscope", as compared to monosyllabic word stimuli.

Differential Diagnosis Based on Impaired Linguistic-Motor Processes

The third group of studies ($n = 17$) focused on using experimental paradigms to isolate deficits in planning/ programming of speech in order to differentiate individuals with AOS/CAS from other speech diagnoses. These paradigms are based on theoretical models that posit a planning/programming level in the speech production process, which may be separated from both higher level language processes and more downstream motor execution processes (Guenther et al., 2006; Hickok, 2012; Houde & Nagarajan, 2011; Levelt et al., 1999; Tourville & Guenther, 2011). Mechanistic studies of apraxia of speech experimentally manipulate aspects of the typical speech production process in an attempt to isolate impairments at this planning/ programming level.

CAS

Nine studies used experimental paradigms to try to isolate the level of processing breakdown associated with CAS (see Table 6). Experimental protocols included perturbation paradigms ($n = 2$; i.e., using a bite block [Nijland, Maassen, & van der Meulen, 2003] or auditory masking [Iuzzini-Seigel et al., 2015]), electroencephalography (EEG; $n = 2$; Froud & Khamis-Dakwar, 2012; Preston et al., 2014), and behavioral measures ($n = 5$; i.e., phonemic error patterns [Shriberg et al., 2012, 2017c], rhythm imitation [Peter & Stoel-Gammon, 2008], and speech perception tasks [Ingram et al., 2019; Zuk et al., 2018]) to examine processing deficits in CAS. The majority of studies in this category only included a control group of TD speakers ($n = 5$), but four studies included an SSD comparison group. No studies included a dysarthria comparison group. Age ranges varied widely across studies, but all focused on children with CAS over 4 years of age.

Both perturbation studies demonstrated different adaptation responses in children with CAS compared to children with typical development or other SSDs, supporting theoretical deficits in feedforward commands in children with CAS (Iuzzini-Seigel et al., 2015; Nijland, Maassen, & van der Meulen, 2003). The EEG studies identified differences in perception of phonological and phonetic detail (Froud & Khamis-Dakwar, 2012) as well as phonological encoding during word production (Preston et al., 2014) in children with CAS compared to TD children. Behavioral studies indicated general timing deficits (Peter & Stoel-Gammon, 2008) and transcoding deficits (i.e., speech sound additions in a nonword syllable repetition task; Shriberg et al., 2012) in children with CAS compared to controls. Speech perception studies yielded mixed findings; one suggested speech perception deficits in children with CAS (Ingram et al., 2019), and the other suggested that speech perception deficits are not a core characteristic of CAS, but instead related to concomitant language impairment (Zuk et al., 2018). Sensitivity and specificity were not reported for any studies in this category.

AOS

Eight studies in the AOS literature used experimental paradigms to identify the mechanism of impairment and thereby differentiate individuals with AOS from individuals with aphasia only and healthy control individuals (see Table 7). Experimental paradigms used altered/masked auditory feedback ($n = 4$), visuomotor tracking ($n = 2$), bite-block perturbation ($n = 1$), and an auditory discrimination task ($n = 1$). The majority of studies in this category included both a healthy control and an aphasia-only comparison group ($n = 5$); four studies included only a healthy control comparison group.

Results from two out of four altered/masked auditory feedback paradigm studies indicated a decrement in performance—measured in terms of reaction time (Mailend & Maas, 2013), vowel duration, and/or vowel contrast (Maas et al., 2015)—for AOS groups in altered/masked auditory conditions, suggesting impaired feedforward control of speech in AOS. A third study employing a similar auditory feedback paradigm reported the opposite effect (i.e., improved performance on multiple measures of speech fluency) but nonetheless interpreted results in favor of an intact, overrelied upon feedback system, coupled with impaired feedforward control (Jacks & Haley, 2015). The final study involving an altered/masked auditory feedback paradigm investigated patterns of compensation and adaptation rather than more objective performance metrics and found evidence for a greater adaptation among individuals with AOS; the authors suggest that this may be due to a more

malleable motor control system and the modification of feedforward commands therein (Ballard et al., 2018). A bite-block perturbation study (Jacks, 2008) also reported results in line with the hypothesis of feedforward control deficits in AOS, as did both studies using a visuomotor tracking paradigm (Ballard & Robin, 2007; Robin et al., 2008). Ballard and Robin (2007) additionally reported evidence for inefficient integration of feedback leading to suboptimal refinement of feedforward programs. No studies in this category reported on metrics of diagnostic accuracy.

Differential Diagnosis Based on Neuroimaging

The fourth group of studies ($n = 19$) focused on use of neuroimaging biomarkers as a basis for identification of speech apraxia. These studies used imaging modalities that include structural magnetic resonance imaging (MRI) to assess gray and white matter integrity, diffusion tensor imaging (DTI) to assess white matter tract integrity, and positron emission tomography (PET) imaging to identify patterns of hypometabolism (i.e., areas characterized by decreased glucose consumption, a proxy for functional brain activity). Importantly, most studies in this category have not used imaging markers as the basis for differential diagnosis of CAS/AOS but rather have focused on the preliminary step of identifying specific patterns of atrophy or hypometabolism that are characteristic of CAS/AOS and that may, in the future, aid in differential diagnosis.

CAS

Only one study meeting our inclusion criteria was found for examination of neuroimaging biomarkers in children with CAS (Fiori et al., 2016; see Table 8). This study used diffusion-weighted MRI to examine differences in white matter microstructure between children with CAS and TD children over the age of 4 years. Results indicated weakened connectivity of speech-language networks in children with CAS.

AOS

Eighteen of the included studies that used neuroimaging techniques have attempted to identify neuroanatomic correlates to AOS (see Table 9). In contrast to other study categories (i.e., symptoms, quantitative features, processes), the AOS neuroimaging literature is heavily focused on individuals with progressive forms of AOS ($n = 14$) as opposed to poststroke acquired AOS ($n = 4$). Because isolated AOS is more common in cases of progressive, neurodegenerative etiologies (cf. pure poststroke AOS), a large percentage (71%) of studies in this category focused on progressive AOS included a pure AOS group; one of the poststroke studies also included a pure AOS group, although it was relatively small. Regardless of etiology, the majority of studies in this category include an aphasia-only comparison group ($n = 11$) and/or a healthy control group ($n = 9$). A single study in this category included a dysarthria control group.

In terms of imaging modality, the vast majority of studies in this category included structural MRI ($n = 16$). A sizable subset also included PET imaging ($n = 7$), typically fluorodeoxyglucose-PET or tau-PET, to look at patterns of brain hypometabolism and tau uptake, respectively. Six studies also use DTI to evaluate white matter tract integrity. Two studies used functional MRI to look at resting-state connectivity. Two studies included postmortem pathology findings alongside *in vivo* imaging results. One study (Utianski et al., 2019) investigated EEG recording profiles.

Results from the imaging studies indicate that there exist unique patterns of atrophy; reduced connectivity; and, to a lesser extent, hypometabolism in AOS that can be at least partially dissociated

from aphasia-associated atrophy patterns. Multiple studies found a relationship between AOS and atrophy, hypometabolism and/or reduced restingstate connectivity in the precentral gyrus/primary motor area (Basilakos et al., 2015; Botha et al., 2018; Itabashi et al., 2016; Josephs et al., 2014), premotor area (Botha et al., 2015, 2018; Josephs et al., 2014, 2013, 2012, 2006; New et al., 2015; Whitwell, Duffy, Strand, Machulda, et al., 2013), and supplementary motor area (Botha et al., 2015, 2018; Josephs et al., 2012, 2006; Whitwell, Duffy, Strand, Machulda, et al., 2013). Greater left than right atrophy/ reduced connectivity/hypometabolism was reported in each of these regions. These same regions were also implicated across several studies investigating tau uptake using tau-PET scans; these studies demonstrated increased tau uptake in these speech-related regions of interest and, moreover, showed that this uptake pattern was at least partially unique to AOS-only or AOS-predominant (cf. aphasia) groups (Utianski, Whitwell, Schwarz, Duffy, et al., 2018; Utianski, Whitwell, Schwarz, Senjem, et al., 2018). At least one study found a relationship between AOS and atrophy and/or hypometabolism in the midbrain (Josephs et al., 2014, 2013), basal ganglia (Josephs et al., 2014), and somatosensory areas (Basilakos et al., 2015). Results relating atrophy of Broca's area and the insular region were equivocal across studies: Two studies endorsed a relationship between AOS and atrophy in either Broca's area or the insula (Botha et al., 2015; Trupe et al., 2013). However, other studies found that atrophy in these regions was associated with agrammatism and not AOS per se (Josephs et al., 2013; Whitwell, Duffy, Strand, Xia, et al., 2013). DTI results demonstrated white matter damage in left intrafrontal tracts to be correlated with AOS, particularly the left posterior premotorsupplementary motor area pathway (Josephs et al., 2014, 2013, 2012; Mandelli et al., 2014). Studies that looked at underlying pathology through use of postmortem autopsy findings reported strong associations with AOS-predominant syndromes and underlying tau pathology (Caso et al., 2014; Josephs et al., 2006). None of the included neuroimaging studies reported on sensitivity/specificity of neuroimaging biomarkers.

Inclusion Criteria Used for AOS and CAS

The inclusion criteria used by authors to validate diagnoses of CAS or AOS for participants in each reviewed study since 2007 were charted (see Appendixes A and B). Specific features were counted as inclusion criteria if the authors listed the feature as a criterion for diagnosis of CAS/AOS or if they made explicit reference to a criteria set (e.g., ASRS) that includes that feature. The percentage of articles using each speech feature as part of the inclusion criteria was calculated separately for AOS studies and CAS studies (see Figure 2). Figure 3 displays the comparison between the frequency of different inclusion characteristics used for each population.

Discussion

Results of this review found that a wide variety of methods have been used to study differential diagnosis of apraxia of speech in both adult and child populations. The state of the evidence for different approaches to differential diagnosis and remaining barriers to their clinical implementation are discussed below.

State of the Evidence for Different Approaches to Differential Diagnosis

Diagnosis Through Speech Symptoms

Collectively, evidence supports the clinical use of speech symptoms for diagnosis of CAS and AOS. Evidence from studies of CAS indicates good sensitivity and specificity of a few auditory-perceptual measures (or combinations of measures) for distinguishing CAS from other SSDs. This suggests promise for development of assessment batteries based on measures of perceptual speech symptoms that could improve consistency in clinical diagnosis of CAS. In AOS as well, there has been progress toward the development of more standardized assessment batteries to improve the diagnosis of AOS. The ASRS is the best known and most widely used of these assessments, and its authors have also reported on the reliability of each of its component metrics (Strand et al., 2014). In both CAS and AOS, there is potential for improved reliability and diagnostic accuracy of perceptual feature sets as more research is done to identify optimal feature subsets and to determine the utility of clinician training for increasing reliability of perceptual approaches.

Diagnosis Through Quantitative Speech Features

Evidence supports the potential diagnostic utility of quantitative speech measures for improving the reliability of apraxia of speech diagnosis in adults and children. For CAS, one quantitative measure of pausing (i.e., Pause Marker; Shriberg et al., 2017a, 2017b, 2017c) has the strongest evidence supporting its utility as a diagnostic marker for CAS, while other measures may have potential clinical utility in the future. For AOS, the measure with the most robust literature support is the pairwise variability index, an acoustic measure of relative stress in multisyllabic words (Ballard et al., 2016; Ballard, Savage, et al., 2014; Basilakos et al., 2017; Courson et al., 2012; Duffy et al., 2017; Scholl et al., 2018; Vergis et al., 2014). There is also good evidence for the use of rate measures-especially maximum rate measures-to differentiate AOS from phonological or other language impairments but not from dysarthria (Melle & Gallego, 2012; Wilson et al., 2010). Overall, the AOS literature indicates that temporally based quantitative measures likely have better clinical utility as diagnostic markers as compared to measures of phonemic accuracy or production variability.

Diagnosis Through Identifying Impaired Linguistic-Motor Processes

The experimental paradigms varied widely across studies included in this category, limiting our ability to make conclusions about the utility of particular paradigms for differential diagnosis of AOS or CAS. Pediatric studies yielded mixed findings regarding whether the level of processing breakdown in CAS is isolated to just motor planning/ programming or if deficits in phonological encoding, speech perception, and more general deficits in rhythm/memory are also involved. Few studies controlled for comorbid language impairment, suggesting the need for additional validation of findings considering this common comorbidity. In the AOS literature, there seems to be an emerging consensus that AOS reflects a deficit in planning/programming differentiable from phonological impairment on the one hand and motor execution on the other. Despite different experimental paradigms across studies, results tended to support the specific hypothesis of feedforward control deficits as the underlying mechanism of AOS and also a deficit in CAS.

Diagnosis Through Neuroimaging

Neuroimaging evidence related to CAS is extremely limited, and currently, there are no neural markers that inform clinical diagnosis of CAS. Though beyond the scope of this review, genetic biomarkers have

been an emerging area of interest in CAS (Centanni et al., 2015; Laffin et al., 2012; Worthey et al., 2013). We did not find any genetic studies that met our criteria for inclusion in this review. In contrast to CAS, there is a robust and growing body of literature using neuroimaging techniques to aid in the understanding and diagnosis of AOS. The neuroimaging literature on AOS is particularly focused on progressive etiologies, because this population offers a unique opportunity to study AOS in the absence of comorbid language deficits. Neuroimaging evidence demonstrates that AOS is associated with distinct patterns of atrophy (left > right) and other neuroanatomic abnormalities (e.g., hypometabolism, reduced functional connectivity). The most commonly cited regions purported to underlay apraxic speech deficits include the premotor area (Botha et al., 2015, 2018; Josephs et al., 2014, 2013, 2012, 2006; New et al., 2015; Whitwell, Duffy, Strand, Machulda, et al., 2013), precentral gyrus/primary motor area (Basilakos et al., 2015; Botha et al., 2018; Itabashi et al., 2016; Josephs et al., 2014), and supplementary motor area (Botha et al., 2015, 2018; Josephs et al., 2012, 2006; Whitwell, Duffy, Strand, Machulda, et al., 2013). Although neuroimaging evidence has greatly advanced the understanding of the mechanisms of impairment in AOS, the literature is limited with regard to its clinical utility as a diagnostic marker.

Barriers to Clinical Implementation

Methodologies used in the reviewed literature lie on a continuum from behavioral research to neuroimaging research, with varying strengths and limitations to their clinical applicability. Behavioral measures (i.e., observation of surface speech features) have the advantage of being more ecologically valid, more directly informing treatment, and being easy to implement in a clinical setting; however, these measures have historically been inadequate to clearly differentially diagnose apraxia of speech because of the degree of overlap in clinical features between different speech diagnoses, the amount of individual variability among people with motor speech disorders, and challenges with reliable measurement and quantification of behavioral speech features. In contrast, quantitative, experimental, and neuroimaging approaches to differential diagnosis have the advantage of being more objective and reliable, more sensitive to subtle differences, more diagnostically specific, and potentially informative about the underlying etiology. However, these techniques rely on specialized equipment or detailed and time-consuming analysis techniques that are not typically feasible in most clinical settings. While several of these quantitative measures appear promising for assisting with differential diagnosis, research efforts are needed to translate them into clinically feasible tools.

An additional limitation of existing literature is that most studies of both CAS and AOS assume that individuals included in the studies were accurately identified by expert clinical judges based on a defined set of criteria. Using expert clinical judgment as the diagnostic "gold standard" inherently leads to circular logic in research studies; results showing a difference between a priori defined speech apraxia and control groups on quantitative measures provide information about how the groups differ but do not validate the initial accuracy of the clinical diagnosis for included participants. To our knowledge, the reliability of expert clinical diagnosis of CAS and AOS has not been tested, and given the inconsistency in inclusion criteria used across studies, it is likely there may be discrepancies across expert clinicians and research groups regarding diagnosis. This suggests the need for increased consensus on a clinical diagnostic standard and research on the reliability of clinicians' ratings of diagnostic features.

There are also remaining gaps in the research literature that currently limit the clinical utility of some promising potential diagnostic measures and are important areas for future research efforts. First, a major gap in both the child and adult literature is the lack of inclusion of dysarthria comparison groups. Although the majority of studies included a phonological comparison group (i.e., SSD group in child studies, aphasia group in adult studies), only one CAS study and three AOS studies included a dysarthria comparison group. Given the frequency of prosodic and rate disturbances in speakers with dysarthria, the lack of data on these measures from speakers with dysarthria is a critical limitation to discriminating between CAS/AOS and dysarthria. Second, a small proportion of the reviewed studies reported diagnostic accuracy statistics. Sensitivity, specificity, and positive/negative predictive values of potential diagnostic measures are essential for individual-level prediction, which is what is needed in clinical settings to be an effective diagnostic marker. Third, comorbidity with language impairment is a major issue in both child and adult populations. In children, CAS commonly occurs in the presence of comorbid language impairment, but few studies controlled for language impairment in their analyses. In the adult literature, the problem of comorbidity has nothing to do with the inclusion of an aphasia-only control group- which the vast majority of studies include-but rather to do with the fact that pure (poststroke) AOS is rare and most groupings of individuals with AOS have concomitant language impairments, often of a different type than the language impairments seen in the aphasia-only control groups (e.g., nonfluent vs. anomic aphasia). This confound is avoided in studies of primary progressive AOS and highlights the unique contribution of this body of literature (Duffy & Josephs, 2012).

For children, another consideration is age and changes with development and treatment. Features that have been identified as potentially helpful for differential diagnosis have primarily been studied in children over 4 years of age. Current evidence is limited regarding diagnostic features in younger children, although this appears to be an active area of emerging research. Continued future research in this area is needed to improve early identification of children with CAS.

Comparison Between AOS and CAS Literature

Both the AOS and CAS literature show continuing inconsistencies in the criteria used to validate the diagnosis in research participants. Analysis of criteria used in studies since 2007 to qualify individuals for inclusion in speech apraxia groups revealed a greater degree of consensus regarding specific diagnostic features in CAS as compared to AOS. Eight of the 20 total CAS inclusion criteria were used in a majority (> 50%) of studies (i.e., dysprosody, nonspeech groping, increased errors with complexity, distortions, disrupted coarticulation, vowel errors, voicing errors, and inconsistent errors), two of which (i.e., dysprosody and nonspeech groping) were used in more than 80% of studies. In contrast, only three of 16 total AOS inclusion criteria- sound distortions, slow rate, and distorted substitutions- were used in a majority of studies (67%, 67%, and 60% of total studies, respectively), and no features garnered consensus above 70%. It is worth noting, however, that consensus regarding diagnostic criteria has improved markedly since the 2014 publication of the ASRS, which suggests that, for both AOS and CAS, consistency in diagnostic inclusion criteria has benefited from the introduction of formalized guidelines. This emerging consensus in diagnostic criteria is essential for ensuring that findings from research studies are comparable to each other and for their applicability to clinical practice.

Comparisons between the AOS and CAS diagnostic criteria also highlighted the substantial differences in clinical presentations associated with CAS and AOS. Diagnostic criteria used in CAS studies had a relatively greater focus on specific segmental features compared to AOS. Six of the top eight most cited CAS features were segmental, compared to only four of the top eight AOS features. Moreover, CAS segmental features included several that were not used for diagnosis in any AOS studies, including disrupted coarticulation, vowel errors, and voicing errors. This difference in diagnostic inclusion criteria highlights important differences in the clinical presentations associated with AOS and CAS despite the shared theoretical breakdown in speech motor planning/programming. Specifically, this comparison showed more similarity in suprasegmental characteristics between CAS and AOS than in segmental characteristics. The shared suprasegmental characteristics identified in Figure 3 may be particularly valuable for identifying points of overlap where the CAS and AOS bodies of research may best help inform each other.

This review also identified important similarities and differences in methodologies used in AOS and CAS studies that may provide valuable directions for future research. Although similar methodological approaches have been used in both AOS and CAS populations, there are differences in the specific measures that have been most frequently studied. To the degree that symptoms and processing deficits overlap between CAS and AOS, some quantitative features and experimental paradigms that have shown strong evidence in one population may be promising to translate to the other. For example, measures of motor variability (spatiotemporal index) have been primarily studied in CAS but may be useful in AOS studies as well. Because slow rate is a common feature of CAS and AOS, Shriberg and colleagues' Pause Marker (Shriberg et al., 2017a, 2017b, 2017c, 2017d), which has shown good diagnostic accuracy for CAS, may also be useful to study in regard to differential diagnosis of AOS. Conversely, some acoustic measures that have shown promise for aiding in diagnosis of AOS, such as the pairwise variability index, have rarely been studied in CAS and may be useful to examine in future research. Neurogenetic biomarkers are likely to be specific to AOS or CAS, given their distinct etiologies. Thus, although more neuroimaging work is needed to understand the neuroanatomic basis of CAS, it is less likely that knowledge from AOS literature would inform CAS research in this area.

Clinical Implications

Despite the remaining challenges associated with diagnosing AOS and CAS, findings from this review suggest some important implications for practicing clinicians. This review makes clear that, at least among researchers, consensus is building around use of operationalized feature sets, in particular the Mayo 10 criteria (Shriberg et al., 2011) and the ASHA position statement criteria (ASHA, 2007) for CAS, and the ASRS (Strand et al., 2014) for AOS. Thus, clinicians should consider using these criteria sets in their clinical practice to improve consistency in diagnosis and to have greater confidence that findings from the research literature are applicable to the clients on their caseloads.

Second, the literature demonstrates the importance of task considerations in eliciting speech features relevant to differential diagnosis. With regard to CAS, task complexity was shown to be an important factor in differential diagnosis across studies, suggesting the importance of including multiple tasks at varying levels of complexity as part of a clinical evaluation (e.g., single-syllable words, multisyllable words, connected speech samples, diadochokinesis). Many diagnostic features with the strongest support in the literature (e.g., lexical stress or prosodic errors, increased articulatory errors with

increased complexity) are likely to be better elicited through more complex speech tasks; however, inconsistency in errors may best differentiate children with CAS from those with other SSDs in simpler speech tasks. For younger children or those with more severe speech impairment, the Dynamic Evaluation of Motor Speech Skill (Strand et al., 2013) is a published assessment tool with good evidence for its utility in differential diagnosis. In the AOS literature, several of the diagnostic features with broad support (e.g., syllable segmentation, increased errors with increased rate or complexity) require the use of multisyllabic stimuli as part of the assessment battery; moreover, there is evidence that the use of longer multisyllabic stimuli leads to greater diagnostic accuracy for identifying AOS (Duffy et al., 2017). Within the category of multisyllabic words, stimuli with contrastive stress patterns are particularly useful for deriving measures of relative vowel duration.

Third, results of this review show evidence for the potential utility of quantitative measures to support clinical diagnosis. For example, the Pause Marker (Shriberg et al., 2017a, 2017b, 2017c) could be used to increase confidence in making a CAS diagnosis, and pairwise variability indices could inform clinical judgment about equal/excess stress patterns for AOS. Clinical neuroimaging that shows canonical lesion/atrophy patterns (e.g., left-lateralized premotor, primary, and/or supplementary motor areas) also might be cited in support of a clinical diagnosis of AOS. As discussed previously, an important direction for future research is to translate these promising quantitative measures into clinically feasible tools.

Conclusions and Future Directions

The objectives of this scoping review were to (a) summarize the experimental approaches that have been used in the literature to improve differential diagnosis of apraxia of speech in children and adults and to examine the state of the evidence for different approaches and (b) examine the similarities and differences between the AOS and CAS literatures in terms of the state of the evidence for approaches to differential diagnosis. Overall, we found a large body of research that has used speech symptoms, quantitative speech features, experimental paradigms focused on determining impaired linguistic-motor processes, and neuroimaging approaches to address the challenge of differential diagnosis of apraxia of speech in adults and children. Although several promising measures have been identified for improving differential diagnosis of AOS and CAS, few have been tested for their analytical validity, clinical validity, and utility. Clearly, the field is in the early stages with different labs exploring different approaches. Although these efforts, collectively, represent a broad strategy for improving our understanding of apraxia of speech, the findings are not easily harmonized and consolidated, making it difficult to appraise the existing evidence and ultimately achieve scientific consensus. More data are likely to result in more uncertainty unless efforts are made to (a) establish standards that enable researchers to use consistent protocols and data across the research community (e.g., common data elements, standardized assessor instructions, rater training protocols) and (b) promote best practices for testing and reporting diagnostic accuracy (Bossuyt et al., 2003; Moher et al., 2015; Whiting et al., 2011).

Similar methodological approaches have been used to study differential diagnosis of apraxia of speech in adults and children; however, the specific measures that have received the most research attention differ between AOS and CAS. Comparison of inclusion criteria revealed some differences in the speech symptoms associated with CAS and AOS, but also similarities, particularly in suprasegmental

characteristics. To the extent that speech symptomatology overlaps, measures that have shown promise for aiding in differential diagnosis in one population may be appropriate to explore in the other.

This review has also highlighted several areas common to both the CAS and AOS literature where future research is needed. For both child and adult populations, there is a need for comparative studies testing the diagnostic accuracy of multiple candidate markers, better control over language impairment comorbidity, and inclusion of dysarthria control groups. In addition, there is a critical need for translational work moving toward clinical implementation of promising measures. Although speech signs and symptoms can vary significantly from person to person, most studies on speech apraxia have reported on a small number of participants. This long-standing small-sample size problem is, however, now being addressed by (a) promising new advances in mobile recording devices and automated speech analytics (Berry et al., 2019; Connaghan et al., 2019; Rusz et al., 2018; Rutkove et al., 2019) and (b) the establishment of large, publicly available, wellcurated impaired speech databases (Kim et al., 2008; Rudzicz et al., 2012). At best, Big Data approaches will yield efficient and effective multivariate diagnostic models of speech apraxia, and at worst, they will be useful for generating novel hypotheses about differential diagnostic markers that may otherwise not be identified. Overall, the research efforts of the past two decades have resulted in major strides in understanding apraxia of speech in adults and children and made us well positioned for further improvement in objective and reliable clinical diagnosis of AOS and CAS.

Acknowledgments

This project was supported by National Institute on Deafness and Other Communication Disorders Grants K24DC016312 (awarded to Jordan Green), F31DC015703 (awarded to Claire Cordella), and F32DC016484-01 (awarded to Kristen Allison). We would like to thank the research assistants who helped us assemble and organize this literature.

Disclosure: The authors have declared that no competing interests existed at the time of publication.

References

- American Speech-Language-Hearing Association. (2007). Childhood apraxia of speech [Position statement]. <http://www.asha.org/policy>
- Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology*, 8(1), 19-32.
- Ash, S., McMillan, C., Gunawardena, D., Avants, B., Morgan, B., Khan, A., Moore, P., Gee, J., & Grossman, M. (2010). Speech errors in progressive non-fluent aphasia. *Brain and Language*, 113(1), 13-20.
- Aziz, A. A., Shohdi, S., Osman, D. M., & Habib, E. I. (2010). Childhood apraxia of speech and multiple phonological disorders in Cairo-Egyptian Arabic speaking children: Language, speech, and oro-motor differences. *International Journal of Pediatric Otorhinolaryngology*, 74(6), 578-585.
- Bahr, R. H. (2005). Differential diagnosis of severe speech disorders using speech gestures. *Topics in Language Disorders*, 25(3), 254-286.

- Ballard, K. J., Azizi, L., Duffy, J. R., McNeil, M. R., Halaki, M., O'Dwyer, N., Layfield, C., Scholl, D. I., Vogel, A. P., & Robin, D. A. (2016). A predictive model for diagnosing stroke-related apraxia of speech. *Neuropsychologia*, 81, 129-139.
- Ballard, K. J., Halaki, M., Sowman, P., Kha, A., Daliri, A., Robin, D. A., Tourville, J. A., & Guenther, F. H. (2018). An investigation of compensation and adaptation to auditory perturbations in individuals with acquired apraxia of speech. *Frontiers in Human Neuroscience*, 12, 510.
- Ballard, K. J., & Robin, D. A. (2007). Influence of continual biofeedback on jaw pursuit-tracking in healthy adults and in adults with apraxia plus aphasia. *Journal of Motor Behavior*, 39(1), 19-28.
- Ballard, K. J., Savage, S., Leyton, C. E., Vogel, A. P., Hornberger, M., & Hodges, J. R. (2014). Logopenic and nonfluent variants of primary progressive aphasia are differentiated by acoustic measures of speech production. *PLOS ONE*, 9(2), Article e89864.
- Ballard, K. J., Tourville, J. A., & Robin, D. A. (2014). Behavioral, computational, and neuroimaging studies of acquired apraxia of speech. *Frontiers in Human Neuroscience*, 8, 892. <https://doi.org/10.3389/fnhum.2014.00892>
- Bartle-Meyer, C. J., Goozée, J. V., Murdoch, B. E., & Green, J. R. (2009). Kinematic analysis of articulatory coupling in acquired apraxia of speech post-stroke. *Brain Injury*, 23(2), 133-145.
- Basilakos, A. (2018). Contemporary approaches to the management of post-stroke apraxia of speech. *Seminars in Speech and Language*, 39(1), 25-36.
- Basilakos, A., Rorden, C., Bonilha, L., Moser, D., & Fridriksson, J. (2015). Patterns of poststroke brain damage that predict speech production errors in apraxia of speech and aphasia dissociate. *Stroke*, 46(6), 1561-1566.
- Basilakos, A., Yourganov, G., den Ouden, D.-B., Fogerty, D., Rorden, C., Feenaughty, L., & Fridriksson, J. (2017). A multivariate analytic approach to the differential diagnosis of apraxia of speech. *Journal of Speech, Language, and Hearing Research*, 60(12), 3378-3392.
- Berry, J. D., Paganoni, S., Carlson, K., Burke, K., Weber, H., Staples, P., Salinas, J., Chan, J., Green, J. R., Connaghan, K., Barback, J., & Onnela, J. P. (2019). Design and results of a smartphone-based digital phenotyping study to quantify ALS progression. *Annals of Clinical and Translational Neurology*, 6(5), 873-881.
- Bislick, L., McNeil, M. R., Spencer, K. A., Yorkston, K. M., & Kendall, D. L. (2017). The nature of error consistency in individuals with acquired apraxia of speech and aphasia. *American Journal of Speech-Language Pathology*, 26(2S), 611-630.
- Bossuyt, P. M., Reitsma, J. B., Bruns, D. E., Gatsonis, C. A., Glasziou, P. P., Irwig, L. M., Lijmer, J. G., Moher, D., Rennie, D., & de Vet, H. C. W. (2003). Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD initiative. *Radiology*, 226(1), 24-28.
- Bossuyt, P. M., Reitsma, J. B., Bruns, D. E., Gatsonis, C. A., Glasziou, P. P., Irwig, L., Lijmer, J. G., Moher, D., Rennie, D., de Vet, H. C. W., Kressel, H. Y., Rifai, N., Golub, R. M., Altman, D. G., Hooft, L., Korevaar, D. A., & Cohen, J. F. (2015). STARD 2015: An updated list of essential items for reporting diagnostic accuracy studies. *Clinical Chemistry*, 61(12), 1446-1452.
- Botha, H., Duffy, J. R., Whitwell, J. L., Strand, E. A., Machulda, M. M., Schwarz, C. G., Reid, R. I., Spychalla, A. J., Senjem, M. L., Jones, D. T., Lowe, V., Jack, C. R., & Josephs, K. A. (2015). Classification and clinicoradiologic features of primary progressive aphasia (PPA) and apraxia of speech. *Cortex*, 69, 220-236.

- Botha, H., Utianski, R. L., Whitwell, J. L., Duffy, J. R., Clark, H. M., Strand, E. A., Machulda, M. M., Tosakulwong, N., Knopman, D. S., Petersen, R. C., Jack, C. R., Jr., Josephs, K. A., & Jones, D. T. (2018). Disrupted functional connectivity in primary progressive apraxia of speech. *NeuroImage: Clinical*, 18, 617-629.
- Case, J., & Grigos, M. I. (2016). Articulatory control in childhood apraxia of speech in a novel word-learning task. *Journal of Speech, Language, and Hearing Research*, 59(6), 1253-1268.
- Caso, F., Mandelli, M. L., Henry, M., Gesierich, B., Bettcher, B. M., Ogar, J., Filippi, M., Comi, G., Magnani, G., Sidhu, M., Trojanowski, J. Q., Huang, E. J., Grinberg, L. T., Miller, B. L., Dronkers, N., Seeley, W. W., & Gorno-Tempini, M. L. (2014). In vivo signatures of nonfluent/agrammatic primary progressive aphasia caused by FTLD pathology. *Neurology*, 82(3), 239-247.
- Centanni, T. M., Sanmann, J. N., Green, J. R., Iuzzini-Seigel, J., Bartlett, C., Sanger, W. G., & Hogan, T. P. (2015). The role of candidate-gene CNTNAP2 in childhood apraxia of speech and specific language impairment. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 168(7), 536-543.
- Cerami, C., Dodich, A., Greco, L., Iannaccone, S., Magnani, G., Marcone, A., Pelagallo, E., Santangelo, R., Cappa, S. F., & Perani, D. (2017). The role of single-subject brain metabolic patterns in the early differential diagnosis of primary progressive aphasias and in prediction of progression to dementia. *Journal of Alzheimer's Disease*, 55(1), 183-197.
- Connaghan, K. P., Green, J. R., Paganoni, S., Chan, J., Weber, H., Collins, E., Richburg, B., Eshghi, M., Onnela, J., & Berry, J. D. (2019). Use of Beiwe smartphone app to identify and track speech decline in amyotrophic lateral sclerosis. In *Proceedings of Interspeech*, pp. 4504-4508.
- Courson, M.-E. A., Ballard, K. J., Canault, M., Layfield, C. A., Scholl, D. I., & Gentil, C. (2012). Lexical stress production in healthy and apraxic speakers of Australian English or French. *Journal of Medical Speech-Language Pathology*, 20(4), 47-52.
- Croot, K., Ballard, K., Leyton, C. E., & Hodges, J. R. (2012). Apraxia of speech and phonological errors in the diagnosis of nonfluent/agrammatic and logopenic variants of primary progressive aphasia. *Journal of Speech, Language, and Hearing Research*, 55(5), 1562-1572.
- Cunningham, K. T., Haley, K. L., & Jacks, A. (2016). Speech sound distortions in aphasia and apraxia of speech: Reliability and diagnostic significance. *Aphasiology*, 30(4), 396-413.
- Davis, B. L., & Velleman, S. L. (2000). Differential diagnosis and treatment of developmental apraxia of speech in infants and toddlers. *Infant-Toddler Intervention: The Transdisciplinary Journal*, 10(3), 177-192.
- Duffy, J. R. (2013). *Motor speech disorders: Substrates, differential diagnosis, and management* (3rd ed.). Elsevier Health Sciences.
- Duffy, J. R., Hanley, H., Utianski, R., Clark, H., Strand, E., Josephs, K. A., & Whitwell, J. L. (2017). Temporal acoustic measures distinguish primary progressive apraxia of speech from primary progressive aphasia. *Brain and Language*, 168, 84-94.
- Duffy, J. R., & Josephs, K. A. (2012). The diagnosis and understanding of apraxia of speech: Why including neurodegenerative etiologies may be important. *Journal of Speech, Language, and Hearing Research*, 55(5), S1518-S1522.
- Duncan, E. S., Donovan, N. J., & Sajjadi, S. A. (2019). Clinical assessment of characteristics of apraxia of speech in primary progressive aphasia. *American Journal of Speech-Language Pathology*, 29(1S), 485-497.

- Fiori, S., Guzzetta, A., Mitra, J., Pannek, K., Pasquariello, R., Cipriani, P., Tosetti, M., Cioni, G., Rose, S. E., & Chilosi, A. (2016). Neuroanatomical correlates of childhood apraxia of speech: A connectomic approach. *NeuroImage: Clinical*, 12, 894-901.
- 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.
- Froud, K., & Khamis-Dakwar, R. (2012). Mismatch negativity responses in children with a diagnosis of childhood apraxia of speech (CAS). *American Journal of Speech-Language Pathology*, 21(4), 302-312.
- Graff-Radford, J., Jones, D. T., Strand, E. A., Rabinstein, A. A., Duffy, J. R., & Josephs, K. A. (2014). The neuroanatomy of pure apraxia of speech in stroke. *Brain and Language*, 129(1), 43-46.
- Grigos, M. I., Moss, A., & Lu, Y. (2015). Oral articulatory control in childhood apraxia of speech. *Journal of Speech, Language, and Hearing Research*, 58(4), 1103-1118.
- Guenther, F. H. (2016). *Neural control of speech*. MIT Press.
- Guenther, F. H., Ghosh, S. S., & Tourville, J. A. (2006). Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain and Language*, 96(3), 280-301.
- Haley, K. L. (2002). Temporal and spectral properties of voiceless fricatives in aphasia and apraxia of speech. *Aphasiology*, 16(4-6), 595-607.
- Haley, K. L., & Jacks, A. (2019). Word-level prosodic measures and the differential diagnosis of apraxia of speech. *Clinical Linguistics & Phonetics*, 33(5), 479-495.
- Haley, K. L., Jacks, A., & Cunninghama, K. T. (2013). Error variability and the differentiation between apraxia of speech and aphasia with phonemic paraphasia. *Journal of Speech, Language, and Hearing Research*, 56(3), 891-905.
- Haley, K. L., Jacks, A., de Riesthal, M., Abou-Khalil, R., & Roth, H. L. (2012). Toward a quantitative basis for assessment and diagnosis of apraxia of speech. *Journal of Speech, Language, and Hearing Research*, 55(5), S1502-S1517.
- Haley, K. L., Jacks, A., Richardson, J. D., & Wambaugh, J. L. (2017). Perceptually salient sound distortions and apraxia of speech: A performance continuum. *American Journal of Speech-Language Pathology*, 26(2S), 631-640.
- Haley, K. L., & Overton, H. B. (2001). Word length and vowel duration in apraxia of speech: The use of relative measures. *Brain and Language*, 79(3), 397-406.
- Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature Reviews Neuroscience*, 13(2), 135-145.
- Highman, C., Hennessey, N., Sherwood, M., & Leitão, S. (2008). Retrospective parent report of early vocal behaviours in children with suspected childhood apraxia of speech (sCAS). *Child Language Teaching and Therapy*, 24(3), 285-306.
- Houde, J. F., & Nagarajan, S. S. (2011). Speech production as state feedback control. *Frontiers in Human Neuroscience*, 5, 82.
- Ingram, S. B., Reed, V. A., & Powell, T. W. (2019). Vowel duration discrimination of children with childhood apraxia of speech: A preliminary study. *American Journal of Speech-Language Pathology*, 28(2S), 857-874.
- Itabashi, R., Nishio, Y., Kataoka, Y., Yazawa, Y., Furui, E., Matsuda, M., & Mori, E. (2016). Damage to the left precentral gyrus is associated with apraxia of speech in acute stroke. *Stroke*, 47(1), 31-36.

- Iuzzini-Seigel, J. (2019). Motor performance in children with childhood apraxia of speech and speech sound disorders. *Journal of Speech, Language, and Hearing Research*, 62(9), 3220-3233.
- Iuzzini-Seigel, J., Hogan, T. P., & Green, J. R. (2017). Speech inconsistency in children with childhood apraxia of speech, language impairment, and speech delay: Depends on the stimuli. *Journal of Speech, Language, and Hearing Research*, 60(5), 1194-1210.
- Iuzzini-Seigel, J., Hogan, T. P., Guarino, A. J., & Green, J. R. (2015). Reliance on auditory feedback in children with childhood apraxia of speech. *Journal of Communication Disorders*, 54, 32-42.
- Jacks, A. (2008). Bite block vowel production in apraxia of speech. *Journal of Speech, Language, and Hearing Research*, 51(4), 898-913.
- Jacks, A., & Haley, K. L. (2015). Auditory masking effects on speech fluency in apraxia of speech and aphasia: Comparison to altered auditory feedback. *Journal of Speech, Language, and Hearing Research*, 58(6), 1670-1686.
- Jacks, A., Mathes, K. A., & Marquardt, T. P. (2010). Vowel acoustics in adults with apraxia of speech. *Journal of Speech, Language, and Hearing Research*, 53(1), 61-74.
- Jonkers, R., Feiken, J., & Stuive, I. (2017). Diagnosing apraxia of speech on the basis of eight distinctive signs. *Canadian Journal of Speech-Language Pathology and Audiology*, 41(3), 303-319.
- Josephs, K. A., Duffy, J. R., Strand, E. A., Machulda, M. M., Senjem, M. L., Gunter, J. L., Schwarz, C. G., Reid, R. I., Spsychalla, A. J., Lowe, V. J., Jack, C. R., Jr., & Whitwell, J. L. (2014). The evolution of primary progressive apraxia of speech. *Brain*, 137(10), 2783-2795.
- Josephs, K. A., Duffy, J. R., Strand, E. A., Machulda, M. M., Senjem, M. L., Lowe, V. J., Jack, C. R., Jr., & Whitwell, J. L. (2013). Syndromes dominated by apraxia of speech show distinct characteristics from agrammatic PPA. *Neurology*, 81(4), 337-345.
- Josephs, K. A., Duffy, J. R., Strand, E. A., Machulda, M. M., Senjem, M. L., Master, A. V., Lowe, V. J., Jack, C. R., Jr., & Whitwell, J. L. (2012). Characterizing a neurodegenerative syndrome: Primary progressive apraxia of speech. *Brain*, 135(5), 1522-1536.
- Josephs, K. A., Duffy, J. R., Strand, E. A., Whitwell, J. L., Layton, K. F., Parisi, J. E., Hauser, M. F., Witte, R. J., Boeve, B. F., Knopman, D. S., Dickson, D. W., Jack, C. R., Jr., & Petersen, R.C. (2006). Clinicopathological and imaging correlates of progressive aphasia and apraxia of speech. *Brain*, 129(Pt. 6), 1385-1398.
- Keske-Soares, M., Uberti, L. B., Gubiani, M. B., Gubiani, M. B., Ceron, M. I., & Pagliarin, K. C. (2018). Performance of children with speech sound disorders in the Dynamic Evaluation of Motor Speech Skills. *Codas*, 30(2), Article e20170037.
- Kim, H., Hasegawa-Johnson, M., Perlman, A., Gunderson, J., Huang, T. S., Watkin, K., & Frame, S. (2008). Dysarthric speech database for universal access research. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, Brisbane, Queensland, Australia.
- Knežević, D. (2019). Are children with childhood apraxia of speech a subgroup of children with developmental coordination disorders? *Logopedija*, 9(1), 9-13.
- Kopera, H. C., & Grigos, M. I. (2019). Lexical stress in childhood apraxia of speech: Acoustic and kinematic findings. *International Journal of Speech-Language Pathology*, 22(1), 12-23.
- Laffin, J. J. S., Raca, G., Jackson, C. A., Strand, E. A., Jakielski, K. J., & Shriberg, L. D. (2012). Novel candidate genes and regions for childhood apraxia of speech identified by array comparative genomic hybridization. *Genetics in Medicine*, 14(11), 928-936.

- Levac, D., Colquhoun, H., & O'Brien, K. K. (2010). Scoping studies: Advancing the methodology. *Implementation Science*, 5, 69.
- Levelt, W. J. M. (1992). Accessing words in speech production: Stages, processes and representations. *Cognition*, 42(1-3), 1-22.
- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22(1), 1-38.
- Lewis, B. A., Avrich, A. A., Freebairn, L. A., Taylor, H. G., Iyengar, S. K., & Stein, C. M. (2011). Subtyping children with speech sound disorders by endophenotypes. *Topics in Language Disorders*, 31(2), 112-127.
- Maas, E., Mailend, M.-L., & Guenther, F. H. (2015). Feedforward and feedback control in apraxia of speech: Effects of noise masking on vowel production. *Journal of Speech, Language, and Hearing Research*, 58(2), 185-200.
- Maassen, B. (2002). Issues contrasting adult acquired versus developmental apraxia of speech. *Seminars in Speech and Language*, 23(4), 257-266.
- 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.
- Mailend, M.-L., & Maas, E. (2013). Speech motor programming in apraxia of speech: Evidence from a delayed picture-word interference task. *American Journal of Speech-Language Pathology*, 22(2), S380-S396.
- Mandelli, M. L., Caverzasi, E., Binney, R. J., Henry, M. L., Lobach, I., Block, N., Amirbekian, B., Dronkers, N., Miller, B. L., Henry, R. G., & Gorno-Tempini, M. L. (2014). Frontal white matter tracts sustaining speech production in primary progressive aphasia. *Journal of Neuroscience*, 34(29), 9754-9767.
- McNeil, M. R., Pratt, S. R., & Fossett, T. R. D. (2004). The differential diagnosis of apraxia of speech. In B. R. Maassen, R. Kent, H. Peters, P. van Lieshout, & W. Hulstijn (Eds.), *Speech motor control in normal and disordered speech* (pp. 389-413). Oxford University Press.
- McNeil, M. R., Robin, D. A., & Schmidt, R. A. (2009). Apraxia of speech: Definition, differentiation, and treatment. In M. R. McNeil (Ed.), *Clinical management of sensorimotor speech disorders* (2nd ed., pp. 249-268). Thieme.
- Melle, N., & Gallego, C. (2012). Differential diagnosis between apraxia and dysarthria based on acoustic analysis. *The Spanish Journal of Psychology*, 15(2), 495-504.
- Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., Stewart, L. A., & PRISMA-P Group. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews*, 4(1), Article 1.
- Moss, A., & Grigos, M. I. (2012). Interarticulatory coordination of the lips and jaw in childhood apraxia of speech. *Journal of Medical Speech-Language Pathology*, 20(4), 127-132.
- Mumby, K., Bowen, A., & Hesketh, A. (2007). Apraxia of speech: How reliable are speech and language therapists' diagnoses. *Clinical Rehabilitation*, 21(8), 760-767.
- 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.

- Murray, E., McCabe, P., Heard, R., & Ballard, K. J. (2015). Differential diagnosis of children with suspected childhood apraxia of speech. *Journal of Speech, Language, and Hearing Research*, 58(1), 43-60.
- Murray, E., Thomas, D., & McKechnie, J. (2019). Comorbid morphological disorder apparent in some children aged 4-5 years with childhood apraxia of speech: Findings from standardised testing. *Clinical Linguistics & Phonetics*, 33(1-2), 42-59.
- New, A. B., Robin, D. A., Parkinson, A. L., Duffy, J. R., McNeil, M. R., Piguet, O., Hornberger, M., Price, C. J., Eickhoff, S. B., & Ballard, K. J. (2015). Altered resting-state network connectivity in stroke patients with and without apraxia of speech. *NeuroImage: Clinical*, 8, 429-439.
- Nijland, L., Maassen, B., & van der Meulen, S. (2003). Evidence of motor programming deficits in children diagnosed with DAS. *Journal of Speech, Language, and Hearing Research*, 46(2), 437-450.
- Nijland, L., Maassen, B., Van der Meulen, S., Gabreëls, F., Kraaimaat, F. W., & Schreuder, R. (2002). Coarticulation patterns in children with developmental apraxia of speech. *Clinical Linguistics & Phonetics*, 16(6), 461-483.
- Nijland, L., Maassen, B., Van der Meulen, S., Gabreëls, F., Kraaimaat, F. W., & Schreuder, R. (2003). Planning of syllables in children with developmental apraxia of speech. *Clinical Linguistics & Phonetics*, 17(1), 1-24.
- Overby, M., Belardi, K., & Schreiber, J. (2019). A retrospective video analysis of canonical babbling and volubility in infants later diagnosed with childhood apraxia of speech. *Clinical Linguistics & Phonetics*, 34(7), 634-651.
- Overby, M., & Caspari, S. S. (2015). Volubility, consonant, and syllable characteristics in infants and toddlers later diagnosed with childhood apraxia of speech: A pilot study. *Journal of Communication Disorders*, 55, 44-62.
- Overby, M., Caspari, S. S., & Schreiber, J. (2019). Volubility, consonant emergence, and syllabic structure in infants and toddlers later diagnosed with childhood apraxia of speech, speech sound disorder, and typical development: A retrospective video analysis. *Journal of Speech, Language, and Hearing Research*, 62(6), 1657-1675.
- Parrell, B., Agnew, Z., Nagarajan, S., Houde, J., & Ivry, R. B. (2017). Impaired feedforward control and enhanced feedback control of speech in patients with cerebellar degeneration. *Journal of Neuroscience*, 37(38), 9249-9258.
- Patel, R., Connaghan, K., Franco, D., Edsall, E., Forgit, D., Olsen, L., Ramage, L., Tyler, E., & Russell, S. (2013). "The caterpillar": A novel reading passage for assessment of motor speech disorders. *American Journal of Speech-Language Pathology*, 22(1), 1-9.
- Peter, B., & Stoel-Gammon, C. (2008). Central timing deficits in subtypes of primary speech disorders. *Clinical Linguistics & Phonetics*, 22(3), 171-198.
- Preston, J. L., Molfese, P. J., Gumkowski, N., Sorcinelli, A., Harwood, V., Irwin, J. R., & Landi, N. (2014). Neurophysiology of speech differences in childhood apraxia of speech. *Developmental Neuropsychology*, 39(5), 385-403.
- Robin, D. A., Jacks, A., Hageman, C., Clark, H. C., & Woodworth, G. (2008). Visuomotor tracking abilities of speakers with apraxia of speech or conduction aphasia. *Brain and Language*, 106(2), 98-106.
- Rudzicz, F., Namasivayam, A. K., & Wolff, T. (2012). The TORGO database of acoustic and articulatory speech from speakers with dysarthria. *Language Resources and Evaluation*, 46(4), 523-541.

- Rusz, J., Hlavnička, J., Tykalová, T., Novotný, M., Dušek, P., Šonka, K., & Růžička, E. (2018). Smartphone allows capture of speech abnormalities associated with high risk of developing Parkinson's disease. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 26(8), 1495-1507.
- Rutkove, S. B., Qi, K., Shelton, K., Liss, J., Berisha, V., & Shefner, J. M. (2019). ALS longitudinal studies with frequent data collection at home: Study design and baseline data. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, 20(1-2), 61-67.
- Schmid, G., & Ziegler, W. (2006). Audio-visual matching of speech and non-speech oral gestures in patients with aphasia and apraxia of speech. *Neuropsychologia*, 44(4), 546-555.
- Scholl, D. I., McCabe, P. J., Heard, R., & Ballard, K. J. (2018). Segmental and prosodic variability on repeated polysyllabic word production in acquired apraxia of speech plus aphasia. *Aphasiology*, 32(5), 578-597.
- Shriberg, L. D., Aram, D. M., & Kwiatkowski, J. (1997a). Developmental apraxia of speech: II. Toward a diagnostic marker. *Journal of Speech, Language, and Hearing Research*, 40(2), 313-337.
- Shriberg, L. D., Aram, D. M., & Kwiatkowski, J. (1997b). Developmental apraxia of speech: III. A subtype marked by inappropriate stress. *Journal of Speech, Language, and Hearing Research*, 40(2), 313-337.
- Shriberg, L. D., Green, J. R., Campbell, T. F., McSweeney, J. L., & Scheer, A. R. (2003). A diagnostic marker for childhood apraxia of speech: The coefficient of variation ratio. *Clinical Linguistics & Phonetics*, 17(7), 575-595.
- 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.
- 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(2), 487-519.
- Shriberg, L. D., Strand, E. A., Fourakis, M., Jakielski, K. J., Hall, S. D., Karlsson, H. B., Mabie, H. L., McSweeney, J. L., Tilkens, C. M., & Wilson, D. L. (2017a). A diagnostic marker to discriminate childhood apraxia of speech from speech delay: I. Development and description of the Pause Marker. *Journal of Speech, Language, and Hearing Research*, 60(4), S1096-S1117.
- Shriberg, L. D., Strand, E. A., Fourakis, M., Jakielski, K. J., Hall, S. D., Karlsson, H. B., Mabie, H. L., McSweeney, J. L., Tilkens, C. M., & Wilson, D. L. (2017b). A diagnostic marker to discriminate childhood apraxia of speech from speech delay: II. Validity studies of the Pause Marker. *Journal of Speech, Language, and Hearing Research*, 60(4), S1118-S1134.
- Shriberg, L. D., Strand, E. A., Fourakis, M., Jakielski, K. J., Hall, S. D., Karlsson, H. B., Mabie, H. L., McSweeney, J. L., Tilkens, C. M., & Wilson, D. L. (2017c). A diagnostic marker to discriminate childhood apraxia of speech from speech delay: III. Theoretical coherence of the Pause Marker with speech processing deficits in childhood apraxia of speech. *Journal of Speech, Language, and Hearing Research*, 60(4), S1135-S1152.
- Shriberg, L. D., Strand, E. A., Fourakis, M., Jakielski, K. J., Hall, S. D., Karlsson, H. B., Mabie, H. L., McSweeney, J. L., Tilkens, C. M., & Wilson, D. L. (2017d). A diagnostic marker to discriminate childhood apraxia of speech from speech delay: IV. The Pause Marker index. *Journal of Speech, Language, and Hearing Research*, 60(4), S1153-S1169.

- Shriberg, L. D., Tomblin, J. B., & McSweeney, J. L. (1999). Prevalence of speech delay in 6-year-old children and comorbidity with language impairment. *Journal of Speech, Language, and Hearing Research*, 42(6), 1461-1481.
- Šimundić, A.-M. (2009). Measures of diagnostic accuracy: Basic definitions. *eJIFCC*, 19(4), 203-211.
- Stackhouse, J., & Wells, B. (1997). *Children's speech and literacy difficulties: A psycholinguistic framework* (Vol. 1). Wiley.
- Strand, E. A., Duffy, J. R., Clark, H. M., & Josephs, K. (2014). The Apraxia of Speech Rating Scale: A tool for diagnosis and description of apraxia of speech. *Journal of Communication Disorders*, 51, 43-50.
- Strand, E. A., Mccauley, R. J., Weigand, S. D., Stoeckel, R. E., & Baas, B. S. (2013). A motor speech assessment for children with severe speech disorders: Reliability and validity evidence. *Journal of Speech, Language, and Hearing Research*, 56(2), 505-520.
- 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(6), 1595-1609.
- Terband, H., Maassen, B., Guenther, F. H., & Brumberg, J. (2014). Auditory-motor interactions in pediatric motor speech disorders: Neurocomputational modeling of disordered development. *Journal of Communication Disorders*, 47, 17-33.
- Teverovsky, E. G., Bickel, J. O., & Feldman, H. M. (2009). Functional characteristics of children diagnosed with childhood apraxia of speech. *Disability and Rehabilitation*, 31(2), 94-102.
- Thoonen, G., Maassen, B., Gabreëls, F., & Schreuder, R. (1999). Validity of maximum performance tasks to diagnose motor speech disorders in children. *Clinical Linguistics & Phonetics*, 13(1), 1-23.
- Thoonen, G., Maassen, B., Gabreëls, F., Schreuder, R., & de Swart, B. (1997). Towards a standardised assessment procedure for developmental apraxia of speech. *International Journal of Language & Communication Disorders*, 32(1), 37-60.
- Tourville, J. A., & Guenther, F. H. (2011). The DIVA model: A neural theory of speech acquisition and production. *Language and Cognitive Processes*, 26(7), 952-981.
- Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., ... Straus, S. E. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine*, 169(7), 467-473.
- Trupe, L. A., Varma, D. D., Gomez, Y., Race, D., Leigh, R., Hillis, A. E., & Gottesman, R. F. (2013). Chronic apraxia of speech and Broca's area. *Stroke*, 44(3), 740-744.
- Tükel, Ş., Björelus, H., Henningson, G., McAllister, A., & Eliasson, A. C. (2015). Motor functions and adaptive behaviour in children with childhood apraxia of speech. *International Journal of Speech-Language Pathology*, 17(5), 470-480.
- Utianski, R. L., Caviness, J. N., Worrell, G. A., Duffy, J. R., Clark, H. M., Machulda, M. M., Whitwell, J. L., & Josephs, K. A. (2019). Electroencephalography in primary progressive aphasia and apraxia of speech. *Aphasiology*, 33(11), 1410-1417.
- Utianski, R. L., Whitwell, J. L., Schwarz, C. G., Duffy, J. R., Botha, H., Clark, H. M., Machulda, M. M., Senjem, M. L., Knopman, D. S., Petersen, R. C., Jack, C. R., Jr., Lowe, V. J., & Josephs, K. A. (2018). Tau uptake in agrammatic primary progressive aphasia with and without apraxia of speech. *European Journal of Neurology*, 25(11), 1352-1357.

- Utianski, R. L., Whitwell, J. L., Schwarz, C. G., Senjem, M. L., Tosakulwong, N., Duffy, J. R., Clark, H. M., Machulda, M. M., Petersen, R. C., Jack, C. R., Jr., Lowe, V. J., & Josephs, K. A. (2018). Tau-PET imaging with [18F]AV-1451 in primary progressive apraxia of speech. *Cortex*, 99, 358-374.
- Velleman, S. L., & Shriberg, L. D. (1999). Metrical analysis of the speech of children with suspected developmental apraxia of speech. *Journal of Speech, Language, and Hearing Research*, 42(6), 1444-1460.
- Vergis, M. K., Ballard, K. J., Duffy, J. R., McNeil, M. R., Scholl, D., & Layfield, C. (2014). An acoustic measure of lexical stress differentiates aphasia and aphasia plus apraxia of speech after stroke. *Aphasiology*, 28(5), 554-575.
- Wambaugh, J. L., Duffy, J. R., McNeil, M. R., Robin, D. A., & Rogers, M. A. (2006). Treatment guidelines for acquired apraxia of speech: A synthesis and evaluation of the evidence. *Journal of Medical Speech-Language Pathology*, 14(2), 15-34.
- Wertz, R. T., LaPointe, L. L., & Rosenbek, J. C. (1984). Apraxia of speech in adults: The disorder and its management. Grune & Stratton.
- Whiting, P., Rutjes, A. W., Reitsma, J. B., Bossuyt, P. M., & Kleijnen, J. (2003). The development of QUADAS: A tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Medical Research Methodology*, 3(1), Article 25.
- Whiting, P. F., Rutjes, A. W., Westwood, M. E., Mallett, S., Deeks, J. J., Reitsma, J. B., Leeflang, M. M., Sterne, J. A. C., & Bossuyt, P. M. (2011). QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine*, 155(8), 529-536.
- Whitwell, J. L., Duffy, J. R., Strand, E. A., Machulda, M. M., Senjem, M. L., Gunter, J. L., Kantarci, K., Eggers, S. D., Jack, C. R., Jr., & Josephs, K. A. (2013). Neuroimaging comparison of primary progressive apraxia of speech and progressive supranuclear palsy. *European Journal of Neurology*, 20(4), 629-637.
- Whitwell, J. L., Duffy, J. R., Strand, E. A., Xia, R., Mandrekar, J., Machulda, M. M., Senjem, M. L., Lowe, V. J., Jack, C. R., Jr., & Josephs, K. A. (2013). Distinct regional anatomic and functional correlates of neurodegenerative apraxia of speech and aphasia: An MRI and FDG-PET study. *Brain and Language*, 125(3), 245-252.
- Wilson, S. M., Henry, M. L., Besbris, M., Ogar, J. M., Dronkers, N. F., Jarrold, W., Miller, B. L., & Gorno-Tempini, M. L. (2010). Connected speech production in three variants of primary progressive aphasia. *Brain*, 133(7), 2069-2088.
- Worthey, E. A., Raca, G., Laffin, J. J., Wilk, B. M., Harris, J. M., Jakielski, K. J., Dimmock, D. P., Strand, E. A., & Shriberg, L.D. (2013). Whole-exome sequencing supports genetic heterogeneity in childhood apraxia of speech. *Journal of Neurodevelopmental Disorders*, 5(1), Article 29.
- Ziegler, W. (2002). Task-related factors in oral motor control: Speech and oral diadochokinesis in dysarthria and apraxia of speech. *Brain and Language*, 80(3), 556-575.
- Ziegler, W., Aichert, I., & Staiger, A. (2012). Apraxia of speech: Concepts and controversies. *Journal of Speech, Language, and Hearing Research*, 55(5), S1485-S1501.
- Zuk, J., Iuzzini-Seigel, J., Cabbage, K., Green, J. R., & Hogan, T.P. (2018). Poor speech perception is not a core deficit of childhood apraxia of speech: Preliminary findings. *Journal of Speech, Language, and Hearing Research*, 61(3), 583-592.

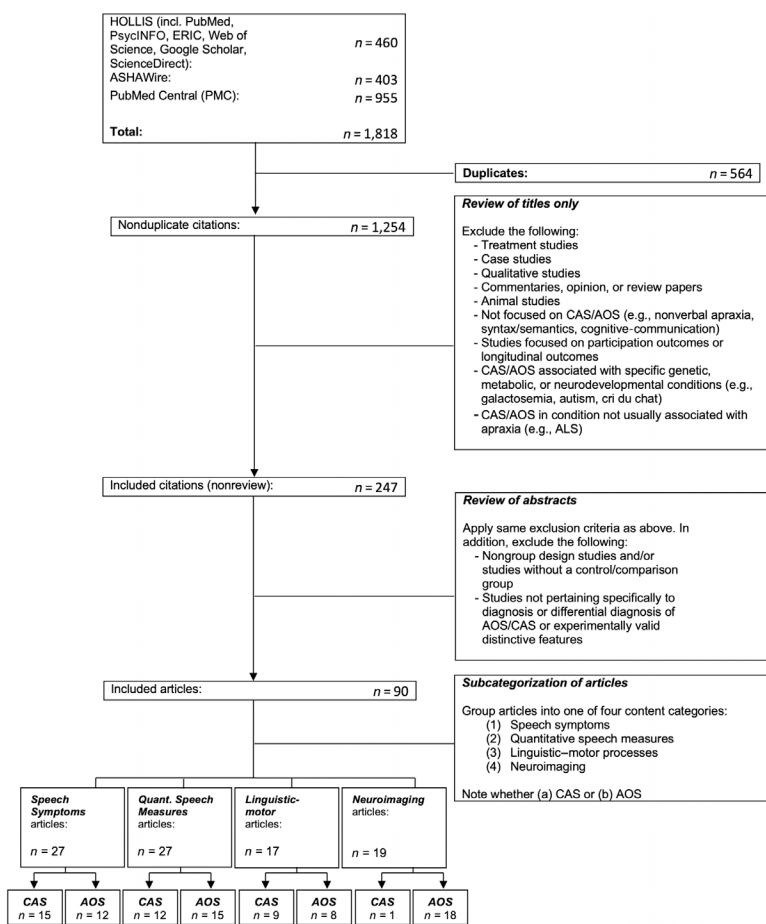


Figure 1. Summary of article search procedures. ALS = amyotrophic lateral sclerosis; AOS = acquired apraxia of speech; CAS = childhood apraxia of speech.

Table 1. Four main content categories for articles based on primary methodological approach.

Category	Methodological approach
Speech symptoms	Identification of CAS/AOS by describing surface speech characteristics using perceptual or clinician judgment, phonetic transcription, or analysis of error patterns (e.g., error counts, categorization of error types).
Quantitative speech measures	Identification of CAS/AOS by quantifying surface features using objective acoustic and/or kinematic measurements (e.g., formant measures, acoustic measures of lexical stress, speech rate, pause durations).
Impaired linguistic-motor processes	Identification of CAS/AOS using experimental paradigms to isolate planning/programming deficits from higher level linguistic or lower level motor execution deficits. Studies using this approach will typically introduce interference at planning/programming stages of speech production (e.g., masking noise/bite-block interfering with normal speech feedback).
Neuroimaging	Identification of unique patterns of atrophy/hypometabolism that may be characteristic of CAS/AOS using imaging modalities such as magnetic resonance imaging, functional magnetic resonance imaging, diffusion tensor imaging, and positron emission tomography imaging.

Note. CAS = childhood apraxia of speech; AOS = acquired apraxia of speech.

Table 2. Charting of CAS studies in speech symptoms category.

Study	Participants		Age range (years;months)	Method/task	Dependent measures	Main results	Sensitivity/ Group n specificity?
	Group	n					
Thoonen et al. (1997)	CAS	11	6;2–7;9	Phonetic transcription; real-word and nonword repetition	- Consonant accuracy and error type	- Higher rate of errors in the CAS group - Rate of substitution errors correlated with severity - The TD group showed larger benefit of real words vs. nonwords compared to CAS group	N
	TD	11	6;0–7;11				
Shriberg et al. (1997a, 1997b)	sCAS	19	4;7–14;11	Phonetic transcription, prosodic coding; conversational speech samples	- Segmental accuracy, intelligibility index, prosodyvoice profile	- Inappropriate stress may be a diagnostic marker for CAS	Y
	SD	73	Age-matched				
Velleman & Shriberg (1999)	sCAsa	15	4;9–14;11	Phonetic transcription, lexical stress coding; conversational speech	- Lexical metrical patterns, syllable omissions, vowel augmentation	- Lexical stress errors were similar between groups	N
	SD	15	3;3–12;10			- Syllable omissions persisted to later ages in the sCAS group	
Bahr (2005)	CAS	5	4;0–7;0	Clinical rating, acoustic analysis; CVC sequences from the Gesture Articulation Test	- Accuracy of gesture use	- The CAS and SD groups had similar number and type of speech gesture errors	N
	TD	5	4;0–7;0		- F2 slope, word duration	- The CAS group had longer word durations than the SD and TD groups	
	SD	5	4;0–7;0				

Highman et al. (2008)	sCAS	20	M = 4;0	Parent report (retrospective)	- Parent report on early vocalizations, babbling, and feeding behavior	- The sCAS and SLI groups had fewer infant vocalizations than the TD group	N
	LI	20	M = 5;0				
	TD	20	M = 5;1			- The sCAS group had less babbling than the LI/TD groups	
Aziz et al. (2010)	sCAS	10	4;0–6;0	Parent report, clinical rating, phonetic transcription; standardized testing, oral motor exam, spontaneous speech, nursery rhyme	- Segmental accuracy	- The sCAS group had lower segmental accuracy, increased difficulty with polysyllabic words and consonant clusters, and deficits in prosody compared to SD and TD groups	N
	SD	10	4;0–6;0		- Syllable shape accuracy		
	TD	10	4;0–6;0		- Maximum repetition rate - Prosodic accuracy		
Lewis et al. (2011)	SD	74	4;0–7;0	Clinical rating; standardized testing (phonological awareness, vocabulary, speeded naming), oral motor assessment	- DDK rate	- All 3 groups had deficits in phonological memory	N
	SD + LI	94	4;0–7;0	- Standardized test scores		- DDK rate did not differentiate groups	
	CAS	41	4;0–7;0			- The SSD + LI and CAS groups had lower vocabulary and phonological awareness scores than the SSD-only group	

Strand et al. (2013)	CAS	20	3;0–6;7	Clinical rating; standardized testing (DEMSS)	- Clusters based on DEMSS subscores	- DEMSS largely differentiated children with CAS, mild CAS, and other speech disorders (compared to expert diagnosis)	Y
	SD	61	3;0–6;3				
Murray et al. (2015)	CAS	28	4;0–12;0	Phonetic transcription, lexical stress judgment; standardized testing, spontaneous speech sample, oral motor assessment	- 24 quantitative measures of segmental accuracy, rate, and presence of clinical features	- Model containing syllable segregation, lexical stress matches, PPC of polysyllables, and DDK accuracy had 91% diagnostic accuracy against expert diagnosis	Y
	CAS+	4					
	Non-CAS ^b	15					
Overby & Caspari (2015)	TD	2	4;5–6;4	Phonetic transcription; home videos from birth to age of 2 years: retrospective analysis	- Number of vocalizations	- The CAS group had fewer resonant and nonresonant productions, reduced phonetic inventories and limited syllable shapes at young ages compared to the TD group	N
	CAS	4	3;0–4;5		- Syllable shapes - Consonant inventories - Volubility		
Iuzzini-Seigel et al. (2017)	CAS	10	4;7–17;8	Phonetic transcription; word and sentence repetition	- Token-to-token inconsistency, phonemic inconsistency	- Token-to-token inconsistency was sensitive and specific in differentiating between the CAS group from the SD and LI groups, especially in simpler stimuli	
	CAS + LI	10	4;7–17;8				
	SD	10	4;7–17;8				
	LI	9	4;7–17;8				
	TD	9	4;7–17;8				

Keske-Soares et al. (2018)	CAS	6	4;6–5;8	Standardized testing (DEMSS–Brazilian Portuguese version)	- DEMSS–Brazilian Portuguese subscores	- The CAS group had lower scores in accuracy and consistency than the SD and TD groups	N
	SD	6	4;6–5;8				
	TD	6	4;6–5;8				
Overby, Caspari, & Schreiber (2019)	CAS	7	3;5–8;8	Phonetic transcription; home videos from birth to age of 2 years: retrospective analysis	- Volubility; age of resonant consonant emergence, consonant diversity and frequency, syllable structure diversity and frequency	- Children later diagnosed with CAS were less voluble, used fewer resonant consonants, and had less diverse phonetic repertoires at young ages, and acquired resonant consonants later than children with SD and TD	N
	SD	5	3;5–8;8				
	TD	5	3;5–8;8				
Overby, Belardi, & Schreiber (2019)	CAS	10	3;0–8;11	Coding of home videos in three age brackets (7–12, 13–18, and 19–24 months): retrospective analysis	- Number of canonical babbles, number of noncanonical babbles, volubility, canonical babbling ratio	- Children later diagnosed with CAS used fewer canonical babbles, had lower volubility, and had later onset of canonical babbling compared to the SD and TD groups	N
	SD	4	3;0–8;11				
	TD	6	3;0–8;11				

Note. CAS = childhood apraxia of speech, developmental apraxia of speech, speech disorder–developmental apraxia of speech; Y/N = yes/no; TD = typically developing; sCAS = suspected childhood apraxia of speech; SD = speech sound disorder, phonological disorder, articulation disorder, multiple phonological disorder; CVC = consonant–vowel–consonant; F2 = second formant; SLI = specific language impairment; LI = language impairment; DDK = diadochokinetic; DEMSS = Dynamic Evaluation of Motor Speech Skill; PPC = percentage phonemes correct.

^asCAS for this study was called SD-DAS and split into two groups: SD-DASi (with inappropriate prosody) and SD-DASa (with appropriate prosody). ^bNon-CAS included dysarthria, phonological disorder, and submucosal cleft.

Table 3. Charting of acquired apraxia of speech (AOS) studies in speech symptoms category.

	Participants						
--	--------------	--	--	--	--	--	--

Study	Group	n	Method/task	Dependent measures	Main results	Sensitivity/ Group n specificity?
Ziegler (2002)	strAOS	15	Clinician rating, acoustic measures; real- word/nonword repetition (sentence)	Speech rate measures for sentence production and DDK, perceptual severity, and rate measures	<ul style="list-style-type: none"> - Rate was slowed in both AOS and dysarthria (except PD) groups for sentence repetition task - AOS groups showed more syllable isochrony and disfluency compared to the dysarthria group - DDK was slowed for dysarthria group (except PD), but not AOS group 	
	strDYS.	125				
	HC	32				
Mumby et al. (2007)	strAOS + APH.	23	Clinician rating; standardized testing, oral motor exam	Presence and severity of AOS	<ul style="list-style-type: none"> - Inter- and intrarater reliability was high for diagnosing both presence and severity of AOS 	N
	strAPH	19				
Ash et al. (2010)	prAOS + APH.	16	Phonetic transcription; spontaneous speech; MRI	Error count + type, cortical atrophy	<ul style="list-style-type: none"> - PNFA had significantly greater number of total errors compared to HC - 82% of errors produced by PNFA were phonemic (cf. phonetic) - Cortical atrophy in prefrontal regions bilaterally and LH perisylvian regions 	N
Croot et al. (2012)	prAOS + APH	9	Phonetic transcription; spontaneous speech + real-	Error type (apraxic vs. phonological), PiB-PET status	<ul style="list-style-type: none"> - Apraxic errors had high sensitivity for nfvPPA while phonological errors had high specificity for lvPPA 	Y

			word/nonword repetition; PiB-PET		- PiB negativity was associated with nfVPPA	
	prAPH	14				
Haley et al. (2012)	strAOS + APH**includes probable AOS	31.	Phonetic transcription, clinician rating, acoustic measures; real-word/nonword repetition	Error counts (segment substitution/error/distortion, revision, prolongation), word/segment duration, scanning index, DDK rate	- strAOS + APH group was differentiable from HC and strAPH group on most operationalized and acoustic measures evaluated, with the exception of the sentence scanning index - Operationalized metrics showed good interrater reliability	N
	strAPH.	8				
	aHC	20				
Haley et al. (2013)	strAOS + APH	15	Phonetic transcription; real-word/nonword repetition	Error consistency (consistency of error location, variability of error type, error token variability, total token variability)	- No between-groups differences in error consistency metrics for strAOS + APH compared to strAPH	N
	strAPH	11				
Strand et al. (2014)	prAOS	23	Clinician rating; standardized testing; real-word/nonword repetition	Inter- and intrajudge ICC for ASRS	- Inter- and intrajudge ICC measures were high (> .9) for AOS characteristics identified as present	N
	prAOS + APH	33				
	prAPH	78				
Cunningham et al. (2016)	strAOS + APH	7	Phonetic transcription; real-word/ nonword repetition	Error count (distortion errors)	- strAOS + APH group made a greater number of distortion errors compared to strAPH group	N
	strAPH	7				

Bislick et al. (2017)	strAOS + APH	10	Phonetic transcription; real-word/nonword repetition	Error consistency (location + type)	<ul style="list-style-type: none"> - No between-groups differences in consistency of error location - strAOS + APH group showed greater variability of error type, but only in blocked condition - strAOS + APH group produced more phonetic errors than strAPH group 	N
	strAPH	10				
Haley et al. (2017)	strAOS + APH	33	Phonetic transcription; real-word/nonword repetition	Error count (distortion and distorted substitution errors)	<ul style="list-style-type: none"> - strAOS + APH group produced significantly more distortion and distorted-substitution errors compared to strAPH group 	N
	strAPH	33				
Jonkers et al. (2017)	strAOS + APH	30	Clinician rating; standardized testing; real-word/nonword repetition	Inter- and intrarater reliability for eight speech features; feature count	<ul style="list-style-type: none"> - Presence of at least 3/8 candidate diagnostic speech features was predictive of AOS (cf. aphasia only, dysarthria) in 88% of cases - Within AOS group, marked variability in which signs were present/diagnostic of AOS 	N
	strAPH.	10				
	strDYS	10				
	HC	35				
Duncan et al. (2019)	prAOS + APH	18	Clinician rating; standardized testing; oral motor exam	Presence and severity of AOS; interrater reliability for 14 ASRS features	<ul style="list-style-type: none"> - Interrater agreement was high for diagnosing presence and severity of AOS, but lower for specific speech features 	N

					- Articulatory groping and increased errors with increased length/complexity were the speech features most predictive of AOS severity	
	prAPH	33				

Note. Most progressive aphasia studies reported results using consensus criteria groupings (nonfluent variant primary progressive aphasia [nfvPPA], logopenic variant primary progressive aphasia [lvPPA], semantic variant primary progressive aphasia [svPPA]) or Mayo criteria (primary progressive apraxia of speech [PPAOS]). We have relabeled those as follows: PPAOS is considered an AOS group; lvPPA and svPPA are considered APH groups. nfvPPA is considered an AOS + APH group, unless authors specified which of two consensus criteria were met; in these cases, nfvPPA with agrammatism only was considered an APH group, whereas nfvPPA with motor speech impairment only was considered an AOS-only group. str = poststroke or other acute acquired etiology; AOS = AOS without comorbid language deficits; DDK = diadochokinetic/diadochokinetic rate; PD = Parkinson's disease; Y/N = yes/no; DYS = dysarthria-only group (no AOS, no aphasia); HC = healthy control; AOS + APH = AOS with comorbid language impairment; APH = aphasia-only deficits (no AOS); pr = progressive etiology; MRI = magnetic resonance imaging; PNFA = progressive nonfluent aphasia; LH = left hemisphere; PiB = Pittsburgh compound B; PET = positron emission tomography; ICC = intraclass correlation coefficient; ASRS = Apraxia of Speech Rating Scale.

Table 4. Charting of CAS studies in quantitative speech measures category.

Study	Participants Group	n	Age range years/years;months	Method/task	Dependent measures	Main results	Sensitivity/Group n specificity?
Thoonen et al. (1999)	TD	11	6–8	Acoustic measures; maximum performance tasks	- Maximum phonation duration, maximum fricative duration, rate and accuracy of DDK	- Assessment protocol of maximum performance tasks had 89%–100% sensitivity and 97% specificity for differential diagnosis of CAS and dysarthria	Y
	Dys.	9	6–10				
	CAS.	11	6–8				

	SD.	11	4–10				
Maasen et al. (2001)	CAS	6	5;0–5;11	Acoustic measures; structured phrases	- F2 trajectories	- CAS group had greater anticipatory coarticulation and more variable formant trajectories than TD group	N
	TD	6	5;0–5;11				
Nijland et al. (2002)	CAS	9	5;0–6;10	Acoustic measures; nonword repetition	- F2 trajectories	- CAS group had more variable coarticulation, less distinction between vowels than control groups	N
	TD	6	4;9–5;11				
	HC	6	20–30				
Munson et al. (2003)	sCAS	5	3;09–8;10	Acoustic measures, perceptual rating; troachic and iambic nonword repetition	- Vowel duration, F0, timing of F0 peak, intensity	- sCAS and SD groups both produced acoustic differences between stressed and unstressed syllables, but listeners judged the sCAS group to have fewer correct lexical stress productions	N

	SD	5			- Perceptual judgments of lexical stress		
Nijland, Maassen, Van der Meulen, Gabreëls, et al., (2003)	CAS	6	5;0–5;11	Acoustic measures; phrase repetition	- F2 trajectories; segmental durations	- Children with CAS had stronger anticipatory coarticulation and reduced prosodic contrasts compared to the TD group	N
	TD	6	4;9–5;11				
Shriberg et al. (2003)	sCAS	11	3;3–10;10	Acoustic measures; real-word repetition	- Lexical stress ratio (LSR)	- Children with sCAS had more extreme LSR values than children with SD	N
	SD	24	3;4–12;0				
Moss & Grigos (2012)	CAS	6	3;0–7;0	Kinematic measures; real-word repetition (1–3 syllables)	- Lip and jaw spatial coupling, temporal coupling, and spatiotemporal index (STI)	- No group differences in spatiotemporal coupling, but CAS group had more variable movements	N
	TD	6					
	SD	6					
Grigos et al. (2015)	CAS	11	3;1–7;2	Kinematic measures; real-word repetition (increasing word length)	- Jaw and lip movement duration, velocity, displacement, and STI	- The CAS group had significantly higher variability in movement; movement	N

						duration and variability differences between the CAS group and the SD group increased as word length increased	
	SD	11	3;2-7;8				
	TD	11	. 3;1-7;0				
Case & Grigos (2016)	CAS	8	5;4-5;7	Kinematic measures, phonetic transcription; novel-word learning	- Segmental accuracy, token-to-token consistency	- CAS group improved consonant accuracy and consistency with practice	N
	TD	8	5;0-5;7		- Lip and jaw movement duration and STI	- Increased variability in lip and jaw movements in CAS group that did not change with practice	
Shriberg et al. (2017b)	CAS	60	4;0-23;0	Acoustic-aided scoring of pauses; 17 speech tasks from Madison Speech Assessment Protocol	- Pause Marker (PM) scores from continuous speech sample	- PM scores had high sensitivity and specificity for identifying speakers with CAS vs. other SDs	Y
	AOS	31	50;0-78;0				
	SD	205	3;0-9;0				

Shriberg et al. (2017d)	CAS	37	4;0–23;0	Acoustic-aided scoring of pauses; acoustic and perceptual measures of speech, prosody, and voice precision stability	- Pause Marker index (i.e., severity metric based on PM scores)	- The Pause Marker index ratings significantly correlated with other measures of CAS precision and stability, suggesting this measure can be used to index severity of CAS	N
	CND	46	3;0–10;0				
	AOS	22	53;0–84;0				
	SD	202	3;0–9;0				
Kopera & Grigos (2019)	CAS	7	3;9–7;2	Acoustic measures, kinematic measures; production of multisyllable word in connected speech	For stressed and unstressed syllables: - Vowel duration, F0 - Jaw movement duration, displacement - Pairwise variability index (PVI): kinematic and acoustic	- CAS group showed reduced jaw movement duration contrast between stressed and unstressed syllables compared to TD group; no other acoustic or kinematic PVI measurements differed between groups	N
	SD	8	4;1–6;7				

	TD	9	4;1-7;0			
--	----	---	---------	--	--	--

Note. TD = typically developing; Y/N = yes/no; Dys = dysarthria; CAS = childhood apraxia of speech, developmental apraxia of speech, speech disorder–developmental apraxia of speech (sCAS = suspected childhood apraxia of speech); DDK = diadochokinetic rate; SD = speech sound disorder, phonological disorder, articulation disorder, multiple phonological disorder; F2 = second formant; HC = healthy adult control; F0 = fundamental frequency; AOS = adult apraxia of speech; CND = complex neurodevelopmental disorder.

Table 5. Charting of acquired apraxia of speech (AOS) studies in quantitative speech measures category.

Study	Participants Group	n	Method/task	Dependent measures	Main results	Sensitivity/Group n specificity?
Haley & Overton (2001)	strAOS + APH	10	Acoustic measures; realword/nonword repetition (multisyllabic words)	Vowel duration of mono- vs. polysyllabic words	- Vowel duration is longer in di- and trisyllabic words (cf. monosyllabic words)	N
	strAPH	10				
	HC	10				
Haley (2002)	strAOS + APH.	10	Acoustic measures; realword/nonword repetition	Fricative (/s/, /j/) segment duration, first spectral moment	- Fricative segment duration was longer for strAOS + APH group compared to healthy controls only - Aberrant phonetic productions of fricatives were observed in both the strAOS + APH and strAPH groups, indicating that	N

					this type of phonetic error was not unique to individuals with a diagnosis of AOS	
	strAPH	10				
	HC	10				
Bartle-Meyer et al. (2009)	strAOS + APH	5	Kinematic measures; DDK	Covariance values between articulators of interest (tongue x jaw, tongue tip x tongue back)	- Articulatory coupling was greater for the majority (4/5) of strAOS + APH patients as compared healthy controls	N
	HC	12				
Jacks et al. (2010)	strAOS + APH	7	Acoustic measures; real-word/nonword repetition	Vowel acoustic measures (absolute Bark formant values, vowel space area, intervowel distance, individual trial-to-trial formant variability)	- No significant between-groups differences (strAOS + APH vs. HC) on any acoustic vowel measures	N
	HC (database)	–				
Wilson et al. (2010)	prAOS + APH	14	Acoustic measures, phonetic transcription; MRI-structural; spontaneous speech	(maximum) speech rate, # distortions, # phonological paraphasias	- Speech rate, particularly maximum speech rate, was reduced for the nfvPPA group compared to other subtypes and HCs - nfvPPA patients had a greater	N

					number of sound distortions compared to other subtypes and HCs	
	prAPH	36				
	NOS	10				
	HC	10				
Courson et al. (2012)	strAOS (French)	4	Acoustic measures; realword/nonword repetition (multisyllabic words)	PVI for vowel duration	- Both strAOS groups (English and French) had lower PVI for vowel duration values compared to HC	N
	strAOS (English).	9				
	HC (French).	4				
	HC (English).	9				
Melle & Gallego (2012)	strAOS + APH	4	Acoustic measures; DDK (+ vowel alteration)	Magnitude/rate/regularity F2 variation, average AMR duration/rate, average SMR duration/rate	- AMR-based measures distinguished between strAOS + APH and HC groups - SMR-based measures distinguished between strAOS + APH and dysarthria groups	N
	strDYS	4				
	HC	15				
Patel et al. (2013)	strAOS + APH.	4	Acoustic measures; passage reading	Passage reading rate, pause frequency,	- Both AOS and dysarthria groups	N

				variation in F0 and intensity, error counts	produced a greater number of errors on complex words - Errors of inconsistency were more common among AOS compared to dysarthria participants	
	strDYS	10				
	HC	7				
Ballard, Savage, et al. (2014) . .	prAOS + APH	20	Acoustic measures, MRI structural, PET (PiB); spontaneous speech, realword/nonword repetition (multisyllabic words)	PVI for vowel duration, peak intensity, syllable segregation (proportion silence time, duration of silences), VBM	- PVI for vowel duration differentiated the nvPPA group from lvPPA and HC groups and was also highly consistent with expert judgment of AOS presence - VBM analysis showed the PVI for vowel duration was related to gray matter intensity in the precentral gyrus, SMA, and IFG regions bilaterally (for nvPPA only)	Y

	prAPH	21				
	HC	17				
Vergis et al. (2014)	strAOS + APH	9	Acoustic measures; realword/nonword repetition (multisyllabic words)	Pairwise variability index (PVI) for vowel duration and peak intensity	- strAOS + APH group demonstrated significantly lower PVI for vowel duration for words with weak–strong stress compared to strAPHand HC groups - No group differences in PVI for intensity	N
	strAPH	8				
	HC	8				
Ballard et al. (2016)	strAOS + APH	35	Acoustic measures, clinician rating; spontaneous speech, real-word/nonword repetition (multisyllabic words), words of inc. length	15 model predictor variables including acoustics and clinician-rated measures	- 2 measures distinguished between strAOS + APH and strAPH groups: (1) speech errors with words of increasing length and (2) relative vowel duration in 3-syllable words with weak–strong stress pattern	Y
	strAPH	37				

Basilakos et al. (2017)	strAOS + APH	20	Acoustic measures; spontaneous speech	PVI for vowel duration, proportion of distortion errors, VOT variability, amplitude envelope modulation spectrum	<ul style="list-style-type: none"> - Classification accuracy for AOS was over 90% for all variables together - Envelope modulation spectrum variables had the greatest effect on classification 	Y
	strAPH	24				
	DC	13				
Duffy et al. (2017)	prAOS	21	Acoustic measures; realword/nonword repetition (multisyllabic words), sentence repetition	Repetition rate for 1- to 4-syllable words + sentences, duration of word, sentence production, PVI for vowel duration	<ul style="list-style-type: none"> - PPAOS group had longer durations and a reduced rate for both single words and sentences compared to all other groups - PPAOS group had a reduced PVI compared to all other groups - Diagnostic accuracy was highest for identifying PPAOS based on acoustic metrics for longer multisyllabic 	Y

					words and sentences	
	prAPH	26				
	HC	11				
Scholl et al. (2018).	strAOS + APH	20	Acoustic measures, phonetic transcription; real-word/ nonword repetition (multisyllabic words)	PVI for vowel duration, error variability, no. of errors, no. of errors over consecutive repetitions	- strAOS + APH group had a greater number of errors overall, greater error variability, reduced improvement across consecutive repetitions, and reduced PVI compared to strAPH group - PVI measure was a stronger predictor of AOS presence than error variability measures	Y
	strAPH	21				
Haley & Jacks (2019)	strAOS + APH	7	Acoustic measures; realword/nonword repetition (multisyllabic words)	PVI for vowel duration, F0, and intensity, lexical stress ratio, word syllable duration	- 3 duration-based acoustic measures differentiated strAOS + APH from both strAPH and HC groups: PVI for vowel duration, lexical stress ratio and word	N

					syllable duration - Diagnostic overlap was smallest for word syllable duration measure, which also had the highest interrater reliability	
	strAPH	9				
	HC	19				

Note. str = poststroke or other acute acquired etiology; AOS + APH = AOS with comorbid language impairment; Y/N = yes/no; APH = aphasia-only deficits (no AOS); HC = healthy control; AOS = AOS without comorbid language deficits; DDK = diadochokinetic rate; pr = progressive etiology; MRI-structural = structural magnetic resonance imaging; NOS = diagnosis not otherwise specified, e.g., semantic dementia, behavioral variant frontotemporal dementia; nfvPPA = nonfluent variant primary progressive aphasia; F2 = second formant; AMR = alternating motion rate; DYS = dysarthria-only group (no AOS, no aphasia); SMR = sequential motion rate; F0 = fundamental frequency; PET = positron emission tomography; PiB = Pittsburgh compound B; lvPPA = logopenic variant primary progressive aphasia; VBM = voxel-based morphometry; SMA = supplementary motor area; IFG = inferior frontal gyrus; VOT = voice onset time; DC = other disease control; PPAOS = primary progressive apraxia of speech.

Table 6. Charting of CAS studies in impaired linguistic–motor processes category.

Study	Participants		Age range (years/years;months)	Method/task	Dependent measures	Main results	Sensitivity/Group n specificity?
	Group	n					
Nijland, Maassen, & van der Meulen, (2003)	CAS	5	5;0–6;10	Acoustic measures; phrase repetition under normal speaking and bite block conditions	- F2 trajectory	- Bite block did not affect anticipatory coarticulation for TD and healthy adult speakers, but had large effect on	N

						coarticulation for children with CAS, suggesting motor planning difficulty	
	TD	5	5;0–6;10				
	HC	6	20–30				
Peter & Stoel-Gammon (2008)	sCAsa	11	4;7–6;6	Acoustic measures, behavioral rating; nonword imitation, rhythm imitation	- % accuracy in imitation, vowel duration, rate	- Low timing accuracy was associated with a high number of CAS characteristics	N
	TD	11	4;10–6;9				
Froud & Khamis-Dakwar (2012)	CAS	5	5;1–8;3	EEG; oddball paradigm with syllables	- Mismatch negativity (MMN) responses	- CAS group showed different MMN responses to allophonic and phonemic contrasts than the TD group, suggesting phonological involvement in CAS	N
	TD	5	5;3–8;9				
Shriberg et al. (2012)	CAS	40	5;0–50;0	Nonword repetition (syllable repetition task [SRT])	- SRT scores: encoding (% substitution errors within manner class), transcoding (additions),	- CAS group had lower SRT scores in multiple domains (encoding, transcoding,	N

					memory (greater difficulty with increasing length)	and memory) compared to controls	
	TD	119	3;0–7;0+				
	SD	140	3;0–7;0+				
	SD + LI	70	3;0–7;0+				
Preston et al. (2014)	CAS	8	9;0–15;0	EEG; monosyllabic and multisyllabic word production	- Event-related potentials (ERPs)	- CAS group had reduced ERP amplitude of signal reflecting phonological encoding while saying multisyllabic words relative to monosyllabic words	N
	TD	13	9;0–15;0				
Iuzzini-Seigel et al. (2015)	CAS	9	6;1–17;6	Acoustic measures; nonword repetition with and without auditory masking	- Voice onset time and vowel space area	- Auditory masking only affected speech of children with CAS, suggesting overreliance on auditory feedback in CAS	N
	SD	10					
	TD	11					
Shriberg et al. (2017c)	CAS	37	4;0–23;0	Acoustic measures, phonetic transcription, prosodyvoice coding;	- PM scores, SRT scores, and percentage consonants correct	- Findings support the presence of deficits in both encoding and transcoding of phonemic	N

				syllable repetition, conversational speech		representations in CAS	
	AOS	22	45;0–84;0				
	SD	205	3;0–9;0				
Zuk et al. (2018)	CAS	7	4;7–17;3	Behavioral response: same–different judgments of /da–/ga/ stimuli	- Discrimination threshold, /da–/ga/ F3 onset frequency	- CAS-only group showed no speech perception differences from TD group; all LI groups showed poorer syllable discrimination than non-LI groups	N
	CAS + LI	6	5;4–12;4				
	LI	7	7;8–12;0				
	SD	12	6;4–9;11				
	TD	15	7;10–16;9				
Ingram et al. (2019)	CAS	9	5;0–6;11	Behavioral response: detection of vowel duration differences	- % accuracy in making same–different judgments regarding vowel length	- Children with CAS exhibited deficits in detecting vowel duration differences compared to TD group, suggesting possible perceptual component	N
	TD	14	5;0–6;11				

Note. CAS = childhood apraxia of speech, developmental apraxia of speech, speech disorder–developmental apraxia of speech (sCAS = suspected childhood apraxia of speech); F2 = second formant; N = no; TD = typically developing; HC = healthy control; EEG = electroencephalography; SD = speech sound disorder, phonological disorder, articulation disorder, multiple phonological disorder; LI = language impairment; PM = Pause Marker; AOS = acquired apraxia of speech; F3 = third formant.

^asCAS 4 and 9 speech characteristics of CAS.

Table 7. Charting of acquired apraxia of speech (AOS) studies in impaired linguistic–motor processes category.

Study	Participants Group	n	Method/task	Dependent measures	Main results	Sensitivity/Group n specificity?
Schmid & Ziegler (2006)	strAOS + APH	7	Error rates (correct/incorrect); discrimination task	Error rate across 4 presentation modes: auditory, visual, bimodal, cross-modal	- Error rates were greater for the AOS = APH and AOS groups across all presentation modes, compared to HCs	N
	strAPH	7				
	HC	14				
Ballard & Robin (2007)	strAOS + APH	8	Kinematic measures; visuomotor tracking	Jaw-target accuracy and variability measures	- HCs' jaw-target responses were more accurate and less variable compared to the strAOS + APH group	N
	HC	15				
Jacks (2008)	strAOS + APH	5	Acoustic measures; bite block	Vowel formant frequencies (F1, F2), Euclidean distance, acoustic	- At baseline (no bite block), production of vowels was less accurate for	N

				distance ratio, perceptual vowel quality rating	strAOS + APH group compared to HCs; however, after introduction of the bite block, accuracy decreased similarly across both groups	
	HC	5				
Robin et al. (2008)	strAOS	5	Kinematic measures; visuomotor tracking	Jaw-target accuracy measures in response to predictable vs. unpredictable feedback	- Accuracy was poorest for strAOS participants in response to predictable signal feedback, but similar to other groups in response to unpredictable signal feedback	N
	strAPH	4				
	HC	8				
Mailend & Maas (2013)	strAOS + APH	5	Reaction time; altered auditory feedback (interference paradigm)	Reaction times (RTs) across 2 conditions (no interference vs. interference) and between interference conditions (shared sounds vs. no shared sounds)	- Patients in strAOS + APH group had longer RTs in distractor vs. no distractor condition; no effect of condition was observed for HCs.	N

	strAPH	2				
	HC	9				
Jacks & Haley (2015)	strAOS + APH	10	Acoustic measures; masked + altered auditory feedback	Syllable rate, disfluency duration, vocal intensity	<ul style="list-style-type: none"> - Introduction of masked auditory feedback improved fluency (increased rate, decreased fluency duration, or both) for strAOS + APH group only - There was no positive effect on fluency in either group in the altered auditory feedback condition 	N
	HC	10				
Maas et al. (2015).	strAOS + APH	6	Acoustic measures; masked auditory feedback (noise masking)	Vowel contrast, variability, duration	<ul style="list-style-type: none"> - Vowel duration was longer and contrast was reduced under masking conditions for the strAOS + APH group compared to HCs; the strAPH group was not significantly 	N

					different compared to HCs - There were no significant differences in vowel variability across groups	
	strAPH	4				
	HC (younger).	11				
	HC (older)	12				
Ballard et al. (2018)a . .	strAOS + APH	8	Acoustic measures; masked + altered auditory feedback (F1 perturbation)	Vowel formant frequencies (% F1 change relative to baseline)	- strAOS + APH group showed adaptation to sustained F1 perturbation (sig. change in F1 to subsequent masked/unperturbed trials), whereas strAPH and HC groups showed no adaptation pattern	N
	strAPH	8				
	HC	10				

Note. str = poststroke or other acute acquired etiology; AOS + APH = AOS with comorbid language impairment; N = no; APH = aphasia-only deficits (no AOS); AOS = AOS without comorbid language deficits; HC = healthy control; F1 = first formant; F2 = second formant.

^aOnly Experiment 2 of Ballard et al. (2018) is summarized, as Experiment 1 does not include a control group.

Table 8. Charting of CAS studies in neuroimaging category.

	Participants						
Study	Group	n	Age range (years)	Method/task	Dependent measures	Main results	Sensitivity/specificity?

Fiori et al. (2016)	CAS	17	5–17	MRI, standardized speech and language testing, oral motor and motor speech testing	Fractional anisotropy (FA) used to generate connectome	Reduced structural connectivity and FA of speech-language networks in children with CAS compared to TD children	N
	TD	10	4–16				

Note. CAS = childhood apraxia of speech; MRI = magnetic resonance imaging; TD = typically developing; N = no.

Table 9. Charting of acquired apraxia of speech (AOS) studies in neuroimaging category.

Study	Participants		Method/task	Dependent measures	Main results	Sensitivity/specificity?
	Group	n				
Josephs et al. (2006).	prAOS	7	MRI-structural, SPECT, pathology, behavioral assessment	GM, WM atrophy	- AOS was primarily associated with atrophy in the premotor and supplementary motor cortices - All AOS cases had a pathological diagnosis characterized by underlying tau biochemistry.	N
	prAPH + AOS	3.				
	prAPH	7				

Josephs et al. (2012)	prAOS	12	MRI-structural, DTI, PiB-PET, FDG-PET, behavioral assessment	GM, WM atrophy; fractional anisotropy; mean diffusivity; voxel-wise metabolism	<ul style="list-style-type: none"> - For prAOS group compared to HC, GM atrophy was focused in superior lateral premotor cortex and supplementary motor area; WM loss was also focused in these regions + inferior premotor cortex and body of corpus callosum - prAOS group showed reduced fractional anisotropy and increased mean diffusivity of the superior longitudinal fasciculus - prAOS groups showed hypometabolism of superior lateral premotor cortex and supplementary motor area 	N
	HC	24				
Josephs et al. (2013)	prAOS	18	MRI-structural, DTI, PiB-PET, FDG-PET,	GM, WM atrophy; fractional anisotropy; mean diffusivity;	- Both proAOS and prAOS + APH groups showed atrophy and	N

			behavioral assessment	voxel-wise metabolism	hypometabolism in premotor cortex and midbrain, whereas prAPH groups showed imaging abnormalities in premotor, prefrontal, temporal, parietal lobes + caudate, insula	
	prAOS + APH	10				
	prAPH	9				
	HC	30				
Trupe et al. (2013)	strAOS + APH	17	MRI-structural, behavioral assessment	Voxel intensity vs. ABA-2 scores (voxel-based lesion–symptom mapping)	- AOS was associated with infarct in Broca’s area, anterior temporal cortex, and posterior insula; AOS severity was positively correlated with lesion volume	N
	strAPH	17				
Whitwell, Duffy, Strand, Machulda, et al. (2013) . .	prAOS	16	MRI-structural, DTI, behavioral assessment	GM, WM atrophy; fractional anisotropy; mean diffusivity	- Both PPAOS and NOS (dx = PSP-S) groups showed GM atrophy in supplementary motor area and WM atrophy in posterior frontal lobes	N

					- PPAOS group showed more focal GM atrophy in superior premotor cortex compared to more widespread (extending into prefrontal cortex) atrophy in PSP-S group	
	NOS	16				
	HC	20				
Whitwell, Duffy, Strand, Xia, et al. (2013) . . .	prAOS	17	MRI-structural, FDG-PET, behavioral assessment	GM atrophy, hypometabolism	- The left superior premotor volume was the only region that correlated with AOS severity (measured using the ASRS) - Neither inferior posterior frontal cortex (i.e., Broca's area) nor insula correlated with AOS severity; Broca's area instead correlated with severity of agrammatism	N
	prAOS + APH	18				
	prAPH	1				

Caso et al. (2014).	nvPPA (FTLD-tau)	9	MRI-structural, pathology, behavioral assessment	GM, WM atrophy; AOS severity ratings	<ul style="list-style-type: none"> - AOS was the most common feature at presentation regardless of FTLD subtype - prAOS (FTLD-tau) characterized by atrophy in GM of left posterior frontal regions and left frontal WM - prAOS (FTLD-TDP) characterized by atrophy in left posterior frontal GM only 	N
	nvPPA (FTLD-TDP).	2				
Josephs et al. (2014)	prAOS	13	MRI-structural, DTI, FDG-PET, behavioral assessment	Rates of whole-brain, ventricle, and midbrain volume atrophy; rates of regions GM atrophy, WM tract degeneration	<ul style="list-style-type: none"> - prAOS group had elevated rates of whole-brain atrophy, ventricular expansion, and midbrain atrophy - Increased rates of atrophy for prAOS group in prefrontal cortex, motor cortex, basal ganglia, and midbrain 	N
	HC	20				

Mandelli et al. (2014)	prAOS + APH	9	MRI-structural, DTI, behavioral assessment	Tract-specific DTI metrics	<ul style="list-style-type: none"> - Significant WM changes in the left intrafrontal and frontostriatal pathways were found in nvPPA, but not in lvPPA or svPPA - Correlations between tract-specific DTI metrics suggested a preferential role of a posterior premotor–SMA pathway in motor speech/AOS 	N
	prAPH	16				
	HC	21				
Basilakos et al. (2015).	strAOS + APH	18	MRI-structural, behavioral assessment	Voxel intensity vs. ASRS scores (voxel-based lesion–symptom mapping)	<ul style="list-style-type: none"> - Patterns of brain damage were at least partially dissociable for strAOS + APH vs. strAPH groups; AOS was most strongly associated with damage to cortical motor regions and somatosensory areas 	N

	strAPH	16				
Botha et al. (2015).	prAOS	40	MRI-structural, DTI	GM atrophy, fractional anisotropy, mean diffusivity	<ul style="list-style-type: none"> - Compared to controls, PPAOS group shows GM atrophy in bilateral premotor and SMA regions, middle cingulate gyri, Broca's area, insular gray matter. DTI abnormalities were observed in same regions and also implicated left uncinata fasciculus and bilateral superior longitudinal fasciculi - Direct comparison of PPAOS and nfvPPA groups revealed greater GM atrophy for nfvPPA group in left temporal, hippocampus and fusiform gyrus 	N
	prAOS + APH	12				
	prAPH	52				
	NOS	26				

New et al. (2015)	strAOS + APH.	15	fMRI (resting state), behavioral assessment	Mean gray, white matter signal intensity	- strAOS + APH group showed reduced connectivity between bilateral premotor regions; reduction of connectivity correlated with AOS severity	N
	strAPH	17				
	HC	18				
Itabashi et al. (2016)	strAOS	7	MRI-structural, behavioral assessment	Voxel intensity vs. diagnosis (voxel-based lesion-symptom mapping)	- Brain regions associated with AOS were centered on the left precentral gyrus	N
	strAOS + APH	15				
	DC	114				
Cerami et al. (2017)	prAOS + APH	19	FDG-PET, behavioral assessment	Voxel-wise metabolism	- Hypometabolism patterns differed across subtypes; among nfvPPA patients, parietal, subcortical and brainstem hypometabolism predict progression to corticobasal syndrome or progressive	N

					supranuclear palsy	
	prAPH	28				
	prDYS	3				
	NOS	5				
Botha et al. (2018).	prAOS	22	fMRI (resting state), MRI-structural, behavioral assessment	Gray, white matter signal intensity in intrinsic connectivity networks (ICNs); connectivity vs. apraxia severity (ASRS scores)	<ul style="list-style-type: none"> - prAOS group showed reduced connectivity in speech and language, face, salience, and left working memory ICNs - Reduced connectivity for prAOS group between right SMA and rest of speech and language ICN, which correlated with AOS severity 	N
	HC	44				
Utianski, Whitwell, Schwarz, Duffy, et al. (2018)	prAOS + APH	5	MRI-structural, tau-PET	Tau uptake, measured using ratio of cortical to cerebellar signal (SUVR) in ROIs	<ul style="list-style-type: none"> - Compared to HC group, prAPH groups showed uptake of tau in left frontal and parietal regions of interest, whereas prAOS + APH group showed uptake in bilateral SMA, frontal lobes, 	N

					precuneus, and precentral gyrus - prAOS + APH showed greater tau uptake in left precentral gyrus compared to prAPH group	
	prAPH	4				
	HC	27				
Utianski, Whitwell, Schwarz, Senjem, et al. (2018)	prAOS	7	MRI-structural, tau-PET, PiB-PET	Tau uptake (SUVR), ROI level and voxel level	- Compared to HC group, both prAOS + APH groups showed increased tau uptake in SMA, precentral gyrus, and Broca's area - prAOS group showed pattern of increased tau uptake only in superior (incl. SMA) and premotor cortices, and not in Broca's area	N
	prAOS + APH	7				
	HC	42				
Utianski et al. (2019)	prAOS	3	EEG, MRI-structural	Posterior dominant rhythm; clinical EEG read	- Patients with aphasia (prAPH and prAOS + APH groups) demonstrated theta slowing whereas the AOS-only group	N

					(prAOS) did not, and instead showed normal EEG patterns	
	prAOS + APH	2				
	prAPH	3				

Note. pr = progressive etiology; AOS = a group with apraxia and no comorbid language deficits (dysarthria status not accounted for); MRI-structural = structural magnetic resonance imaging; SPECT = single-photon emission computed tomography; GM = gray matter; WM = white matter; N = no; APH = a group with aphasia-only deficits (no AOS); DTI = diffusion tensor imaging; PiB = Pittsburgh compound B; PET = positron emission tomography; HC = healthy control; FDG = fluorodeoxyglucose; AOS + APH = AOS group with comorbid language impairment; str = poststroke or other acute acquired etiology; ABA-2 = Apraxia Battery for Adults–Second Edition; PPAOS = primary progressive apraxia of speech; NOS = diagnosis not otherwise specified, e.g., semantic dementia, unclassified primary progressive aphasia cases, behavioral variant frontotemporal dementia, progressive supranuclear palsy; dx = diagnosis; PSP-S = progressive supranuclear palsy syndrome; ASRS = Apraxia of Speech Rating Scale; nfvPPA = nonfluent variant primary progressive aphasia; FTLD-tau = frontotemporal lobar degeneration with tau pathology; FTLD-TDP = frontotemporal lobar degeneration with TDP-43 inclusions; lvPPA = logopenic variant primary progressive aphasia; svPPA = semantic variant primary progressive aphasia; SMA = supplementary motor area; fMRI = functional magnetic resonance imaging; DC = other disease control (e.g., individuals who have had a stroke but with no AOS or aphasia); DYS = dysarthria-only group (no AOS, no aphasia); SUVr = standardized uptake value ratio; ROIs = regions of interest; EEG = electroencephalography.

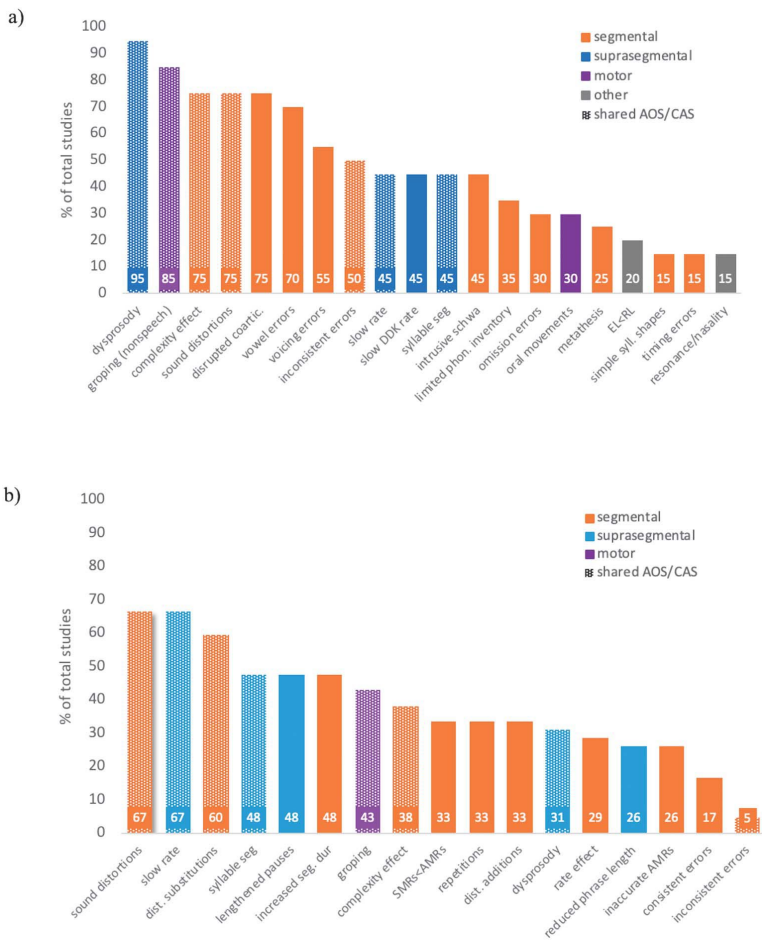


Figure 2. Distribution of inclusion criteria reported in studies since 2007 for (a) determining childhood apraxia of speech (CAS) diagnosis and (b) determining acquired apraxia of speech (AOS) diagnosis. AMRs = alternating motion rates; DDK = diadochokinetic; EL = expressive language; RL = receptive language; SMRs = sequential motion rates.

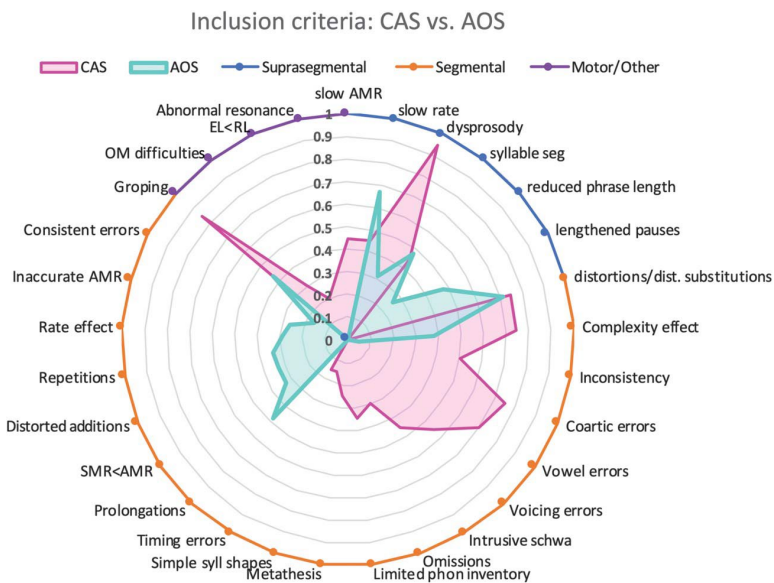


Figure 3. Comparison of inclusion criteria in childhood apraxia of speech (CAS) studies and apraxia of speech (AOS) studies published since 2007. Frequency of occurrence of each individual inclusion criterion is represented on the radial axis as a proportion of studies using the given criterion relative to the total number of CAS

(magenta) or AOS (teal) studies. Shaded areas indicate the degree of (non)overlap between features commonly used in CAS versus AOS studies. AMR = alternating motion rate; EL = expressive language; OM = oral motor; RL = receptive language; SMR = sequential motion rate.