ASHA American Speech-Language-Hearing Association

AJSLP

Review Article

Assessment and Diagnostic Standards of Apraxia of Speech in Chinese-Speaking Adults and Children: A Scoping Review

Eddy Chun Ho Wong,[a] Min Ney Wong,[a,b] and Shelley L. Velleman[c]

[a] Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University, Hung Hom [b] Research Centre for Language, Cognition, and Neurosciences, Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University, Hung Hom [c] Department of Communication Sciences and Disorders, The University of Vermont, Burlington

ARTICLE INFO

Article History:
Received November 8, 2021
Revision received February 24, 2022
Accepted August 22, 2022

Editor-in-Chief: Katherine C. Hustad
Editor: Julie Case

https://doi.org/10.1044/2022_AJSLP-21-00355

ABSTRACT

Purpose: Apraxia of speech (AOS) and childhood apraxia of speech (CAS) are motor-based speech disorders that have been well studied in Indo-European languages. There is limited understanding of these disorders in speakers of Sino-Tibetan languages, such as Chinese. The purpose of this study is to review methods used in research studies for the assessment and diagnosis of AOS and CAS in Chinese speakers.
Method: This scoping review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews guidelines. Articles with a focus on AOS or CAS in Chinese speakers were systematically searched in seven English and six Chinese databases. Three reviewers performed independent screening, data extraction, and quality assessment after obtaining 100% agreement on the prescreening exercise. A qualitative analysis was conducted to rate the quality of diagnoses, ranging from high (Level I) to low (Level III), with Level IV assigned to studies for which the appropriate rating was unclear due to insufficient evidence.
Results: Twenty-eight AOS articles and five CAS articles were identified. A variety of assessment and diagnostic methods were reported. No study of Chinese speakers with AOS or CAS received a rating of Level I. The highest level achieved was Level IIIa for both AOS and CAS studies.
Conclusions: There is no reliable and valid test or method for the diagnosis of AOS or CAS in Chinese speakers. The current gold standard of diagnosis is based upon expert perceptual judgment. Further single-language and cross-linguistic investigations of AOS and CAS and the future development of assessment and diagnostic methods are recommended.

Apraxia of speech (AOS) is a type of motor-based speech disorder that can be caused by stroke, brain injury, or degenerative diseases as an acquired form (Duffy, 2020). A similar disorder with an onset in childhood, with known or unknown origins, is currently called childhood apraxia of speech (CAS; American Speech-Language-Hearing Association [ASHA], 2007). Speakers with either CAS or AOS have a core impairment in motor planning and programming of speech movements and are characterized by disrupted articulation and prosody (ASHA, 2007; Duffy, 2020). Clinical features and diagnostic methods have been studied extensively in English speakers, while the understanding of both CAS and AOS in Chinese speakers is limited.

Clinical Features of CAS and AOS

There are numerous features reported in the literature as symptoms of CAS and/or AOS in English (ASHA,

Correspondence to Eddy Chun Ho Wong: eddychwong@gmail.com.
Disclosure: *The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.*

2007; Duffy, 2020). For diagnosing CAS, ASHA (2007) identified a variety of clinical features from the early CAS research and proposed three consensus features, including inconsistent errors on consecutive repeated productions of the same stimulus, lengthened and disrupted coarticulatory transitions between sounds and syllables, and dysprosody. The validity of the three features has been demonstrated in speakers with CAS (Chenausky et al., 2020), and their diagnostic accuracy has been reported in different independent studies. Iuzzini-Seigel et al. (2017) reported the sensitivity (70%) and specificity (80%) of inconsistency in different productions of the same word at different times for differentiating children with CAS from those with other types of speech sound disorders (SSDs) and those with language impairment. Shriberg et al. (1997) demonstrated the sensitivity (58%) and specificity (42%) of using inappropriate stress to differentiate children with CAS from children with speech delay in a sample of 20 children. In addition, Shriberg et al. (2012) described 10 segmental and suprasegmental features of CAS based on Strand's Diagnostic Checklist (Shriberg et al., 2011), including (a) difficulty achieving initial articulatory configurations and transitions into vowels, (b) syllable segregation, (c) lexical stress errors or equal stress, (d) vowel or consonant distortions including distorted substitutions, (e) nonspeech groping, (f) intrusive schwa, (g) voicing errors, (h) slow rate, (i) slow diadochokinetic (DDK) rate, and (j) increased difficulty with longer or more phonetically complex words. Murray et al. (2015) demonstrated the utility of some of the same features, including syllable segregation, lexical stress matches, phoneme accuracy of polysyllabic words, and articulatory accuracy on repetition of [pətəkə] for identifying CAS. Other speech features such as voicing errors, reduced accuracy of syllable structures, lexical stress errors, and syllable deletion also have been suggested to be useful in differentiating school-age children with CAS from those with other types of SSDs (Benway & Preston, 2020). However, it is important to note that the validity and diagnostic accuracy of some CAS symptoms that have often been reported, such as nonspeech groping behavior (Allison et al., 2020), have not yet been examined.

For adults with AOS, similarly, there is a variety of segmental and suprasegmental clinical features identified from the literature, such as phoneme distortions and distorted substitutions or additions, reduced speech rate, syllable segregation, and equal stress across syllables (Duffy, 2020). However, limited support has been found for the diagnostic accuracy of these features in making a differential diagnosis of AOS. Croot et al. (2012) is the only study that has demonstrated that phonetic distortions, syllable segregation, and equal or excess stress have high sensitivity (89%) to differentiate speakers with progressive AOS from speakers with aphasia only. Allison et al. (2020) concluded that there is still a lack of clinical marker(s) that have empirical diagnostic accuracy for both CAS and AOS.

Possible Clinical Features of CAS and AOS in Chinese

Our current understanding of CAS and AOS is mostly based on English speakers or those who speak other Indo-European languages, such as Dutch, Swedish, or German. In contrast, Chinese is a group of Sino-Tibetan language varieties. It is spoken by 1.3 billion people around the world as their first language (The Ethnologue 200, 2021). Chinese includes Mandarin (the official language of mainland China and Taiwan), Cantonese (the main language spoken in Hong Kong), and other languages. The phonologies of Mandarin and Cantonese are different. Apart from the different numbers and types of consonants, vowels, and diphthongs, Mandarin has triphthongs, which are not present in Cantonese. They also differ in terms of the number and characteristics of lexical tones; Cantonese has six and Mandarin has four lexical tones, whereas English does not use lexical tone at all.

Some studies have claimed that motor speech disorders might manifest differently in different languages (Whitehill, 2010; E. C. H. Wong et al., 2020). Among the limited investigations of CAS and AOS in Chinese speakers, E. C. H. Wong et al. (2020) compared the clinical features of English and Cantonese speakers with CAS. They suggested that some clinical features, such as poor speech intelligibility, poor DDK performance, inconsistent errors, and groping behaviors, would be found in both English and Cantonese speakers with CAS. Other features, such as lexical stress errors, intrusive schwa, and voicing errors, that have been found in English speakers with CAS may not be found in Cantonese speakers with CAS. The authors explained that, in the absence of consonant clusters in Cantonese, it is unlikely to observe intrusive schwa in clusters in these speakers. However, it is possible to find schwa insertions between syllable and word boundries (Chenausky et al., 2020). In addition, all Cantonese consonants are voiceless, although some of these voiceless consonants, such as [p] and [p^h], contrast with respect to aspiration. Thus, it is also unlikely to observe voicing errors in Cantonese. Instead, difficulty in producing contrastive aspiration has been commonly reported in Cantonese phonological development (To et al., 2013). Voicing and aspiration have been considered to be similar from a motor perspective (Whitehill, 2010; E. C. H. Wong et al., 2020). Regarding suprasegmentals, Cantonese is a syllable-timed language, in which syllable duration varies very little, so it is unlikely to observe lexical stress errors in Cantonese. Instead, Cantonese words are contrasted by lexical tones. Therefore, features like deaspiration, lexical tone errors, and weak tone perception and production may be found in Cantonese speakers with CAS but may not be found in English speakers

with CAS (E. C. H. Wong et al., 2020). Furthermore, preliminary evidence of lexical tone production errors, tone perception difficulty, and tone sequencing difficulty in Cantonese-speaking children with CAS has also been reported (E. C. H. Wong et al., 2021, 2022). These tone-related clinical features are new to the literature as they have not been reported in speakers of Indo-European languages with CAS. These newly identified features also suggest that CAS may manifest differently in different languages.

The possible different clinical manifestations of AOS in English versus Chinese can also be inferred from some currently available assessment tools. The assessment standard suggested by the China Rehabilitation Research Centre (CRRC) was recently reported in the English literature (e.g., You et al., 2019). The initial version of the CRRC assessment standard was proposed by S. Li and Shirasaka (1994), and a few other versions have been developed since then (S. Li, 2008; Z. Wang & Li, 2013; Wei & Li, 2000). Z. Wang and Li (2013) reported the use of a set of Chinese materials "referred from" (Z. Wang & Li, 2013, English abstract, p. 70) the Motor Speech Evaluation (MSE; Wertz et al., 1984) in their study. The approach consists of using perceptual judgment of participants' performance on several tasks, including triphthong imitation ([aui] and [iua]), single-syllable imitation ([pa], [ta], and [ka]), trisyllable imitation ([pataka]), and phrase imitation. Although the validity and diagnostic accuracy of these assessment standards in making an AOS diagnosis have not been reported, the inclusion of a triphthong imitation task highlights that there may be a need to examine triphthong production in people with AOS. Examination of triphthong production (i.e., English triphthongs [aɪə˞], [ɔɪə˞], [aʊə˞]) has never been reported for English speakers with AOS.

Diagnostic Methods in English Speakers

While awaiting the establishment of a definitive set of criteria for CAS and AOS, the most common approach for diagnosis is using a checklist that comprises a set of clinical features. For English speakers with CAS, ASHA's three consensus features have been commonly used in research as the only one or one of the requirements to confirm CAS diagnoses in participants, with a set of predefined working definitions and observation criteria. For example, Grigos and Case (2018) required participants to show all three consensus features at least 3 times in more than one speaking context, including single words, connected speech, and syllable sequencing tasks. Thomas et al. (2016) provided a set of relatively objective criteria to confirm CAS diagnoses in their participants. They included participants who had more than 40% inconsistency on the Inconsistency assessment of the Diagnostic Evaluation of Articulation and Phonology (DEAP; Dodd et al., 2006; for children under the age of 11 years), exhibited syllable segregation within at least 10 words, and had a minimum of 15% stress pattern mismatches on the Test of Polysyllables (Gozzard et al., 2006). In addition, Strand's 10-point checklist is a popular diagnostic tool for research (e.g., Shriberg et al., 2017a; Zuk et al., 2018). For example, Shriberg et al. (2011) proposed that the diagnosis of CAS can be made based upon at least four of these 10 features across three or more of the Madison Speech Assessment Protocol tasks. Murray et al. (2015) used the same checklist and demonstrated that syllable segregation, lexical stress mismatches, reduced phoneme accuracy of polysyllabic words, and reduced articulatory accuracy on repetition of [pətəkə] have 91% diagnostic accuracy against expert diagnosis in their sample of children with CAS and without CAS (i.e., dysarthria, phonological disorders, and submucosal cleft).

For English-speaking adults with AOS, the use of the Apraxia of Speech Rating Scale (ASRS; Strand et al., 2014) has been reported. This is a diagnostic tool with 16 items that aims to quantify the presence or absence, relative frequency, and severity of speech characteristics frequently associated with AOS. Examiners are required to score the items on a 5-point scale (where 0 indicates *not present*, 1 indicates *detectable but infrequent*, 2 indicates *frequent but not pervasive*, 3 indicates *nearly always evident but not marked in severity*, and 4 indicates *nearly always evident and marked in severity*). It has been recommended for clinical use to identify AOS because of its validity and high sensitivity (96%) and specificity (100%), as well as high intrarater and interrater reliability (ranging from .87 to .98). The scale was later refined to a 13-item version by another research team (Utianski et al., 2018). The items were organized into (a) articulatory features (e.g., distorted sound substitutions and additions), (b) prosodic features (e.g., syllable segmentation within and across words), and (c) other features (e.g., articulatory groping, off-target speech alternating motion rates). High interrater reliability was reported (i.e., intraclass correlation coefficient = .954; Utianski et al., 2018; Wambaugh et al., 2019). However, it remains to be seen whether this instrument differentiates AOS from other speech diagnoses, such as dysarthria.

As mentioned, CAS or AOS diagnoses are mainly based on expert perceptual judgment of the presence of clinical features from checklists. Nevertheless, one standardized criterion-referenced test is available. The Dynamic Evaluation of Motor Speech Skills (DEMSS), which aims to understand difficulties in speech motor planning and programming and aids clinicians in differential CAS diagnosis (Strand & McCauley, 2019), is the only assessment tool for CAS with validity and reliability. High percentages of agreement for perceptual judgments for test–retest, inter- and intrajudge reliability (88%–90%), high sensitivity (0.70) and specificity (0.97), and both construct and content validity were reported (Strand & McCauley, 2019). Although the reliability and validity reported support its clinical use

(Strand et al., 2013), the utility of this relatively new test for both clinical and research practice is still being explored. Adaptations of the DEMSS to other stress-timed languages, such as Swedish and Brazilian Portuguese, have been reported (Gubiani et al., 2021; Rex et al., 2021). Nevertheless, it is unclear whether this tool could be applied to syllable-timed languages (e.g., Chinese), given that the content and the scoring system are strongly focused on features suggested for the diagnosis of CAS in English. For children with CAS, the DEAP Inconsistency Assessment (Dodd et al., 2006) is also commonly used in the United Kingdom and Australia to provide valuable information on inconsistency of productions of the same word at different points in time (i.e., not consecutively). The result is helpful to determine if the child has the inconsistent errors that are often considered to be a hallmark of CAS.

In addition to the available tests, the syllable repetition task (SRT; Shriberg et al., 2012) is a process-oriented task that aims to identify specific types of deficits in the process of speech production in children. Four scores, including the competence score, encoding score, memory score, and transcoding score, can be calculated after converting perceptual judgments to a quantitative scale. The use of the SRT for revealing underlying speech motor planning and programming deficits and assisting the clinician to make a CAS diagnosis has been supported in the literature (Rvachew & Matthews, 2017).

Quantitative measurements have also been proposed for identifying CAS in children. For example, two acoustic measures that report diagnostic accuracy have been suggested. They are the pause marker (PM; Shriberg et al., 2017a) and the maximum performance tasks (MPT; Thoonen et al., 1999). Shriberg et al. (2017b) reported that the PM has high sensitivity (86.8%) and specificity (100%) for differentiating children with CAS from those with other SSDs. However, it has not yet been applied as a clinical standard. Thoonen et al. (1999) reported that the MPT has high sensitivity (89%) and specificity (100%) for differentiating children with CAS from those with dysarthria, but no other SSDs were reported. In addition, kinematic measures have also been proposed to quantify the performance of children with CAS compared with those without CAS. Case and Grigos (2016) and Grigos et al. (2015) demonstrated that children with CAS can be differentiated from those with speech delay and typical development by looking at the movement variability of the jaw and lips.

For adults with AOS, the Apraxia Battery for Adults–Second Edition (ABA-2; Dabul, 2000) is a standardized test. However, due to the advancement of research evidence, some characteristics included in this test are no longer considered relevant for AOS diagnosis. For example, longer latency times for polysyllabic words can be attributed to deficits in the word retrieval process in patients with aphasia (Galletta & Goral, 2018). Moreover, phonemic perseverative errors can also be found in patients with aphasia (Pilkington et al., 2017), and inconsistent articulatory errors have been removed as a feature of AOS (Duffy, 2013; Haley et al., 2021; McNeil et al., 2017; Staiger et al., 2012; Strand et al., 2014; Utianski et al., 2018). As of now, there are no other standardized tests for adults with AOS. Instead, quantitative measures have been proposed for assisting in making AOS diagnoses, in addition to the ASRS. The Pairwise Variability Index for vowel duration (PVIdur), an acoustic measure of relative stress in multisyllabic words with a weak–strong stress pattern, was suggested to be useful in differentiating aphasia with AOS from aphasia without AOS (Ballard et al., 2016), but not from dysarthria (Melle & Gallego, 2012). Further empirical investigation is therefore needed to confirm whether lexical stress errors are actually a problem for adults with AOS, given that neither version of the ASRS includes lexical stress errors as a feature of adult AOS (Strand et al., 2014; Wambaugh et al., 2019). Haley and Jacks (2019) identified limitations of the PVIdur and suggested the word syllable duration (WSD) measure (Haley et al., 2012) as a more stable and informative alternative for differential diagnosis of AOS. Kinematic measures have also been employed to investigate the articulatory control of speakers with AOS (Bartle-Meyer, Goozée, & Murdoch, 2009; Bartle-Meyer, Goozée, Murdoch, & Green, 2009). However, further investigation with larger samples is still needed to confirm the utility of kinematic measures to inform AOS diagnoses.

Diagnostic Methods in Chinese Speakers

In contrast to the variety of diagnostic methods reported for English speakers with CAS and AOS, there have been limited reports for Chinese speakers. Recently, E. C. H. Wong et al. (2021, 2022) used a relatively objective diagnostic approach modified from Murray et al. (2015) for CAS diagnosis. In particular, the CAS diagnosis was given when the children showed five clinical features across four tasks. The features were (a) inconsistent errors on consonants and vowels in repeated productions of syllables or words; (b) lengthened and disrupted coarticulatory transitions between sounds and syllables; (c) lexical tone errors; (d) reduced accuracy in multisyllabic words, that is, the percentage of phonemes correct (PPC) in polysyllabic words or sentences is less than 60%; and (e) reduced phonetic accuracy on the DDK task, that is, the PPC in stimuli with three different syllables on DDK tasks is less than 70%. The working definition of each feature was presented in the original article. The four tasks used were (a) a speech sample; (b) the Hong Kong Cantonese Articulation Test (Cheung et al., 2006), a standardized articulation test for Cantonese speakers; (c) a DDK task; and (d) imitation of polysyllabic words. Because

clinical features were used for diagnosis, the authors also provided diagnostic criteria: The first three features were each observed at least once on all of the tasks except inconsistent errors in the speech sample. The fourth clinical feature was observed at least once on all of the tasks except the DDK task, whereas the fifth clinical feature was observed on the DDK task only.

For Chinese-speaking adults with AOS, the assessment standard suggested by CRRC has been commonly reported in research across the previous decades (e.g., C. Chen et al., 2020; H. Li et al., 2002). These assessment tasks include imitation of triphthongs ([aui] and [iua]), consonant–vowel syllables ([pa], [ta], and [ka]), a trisyllable ([pataka]), and phrases. In addition, the use of the ABA-2 has been reported by You et al. (2019). However, these authors failed to report the details of how this English standardized test was used on Chinese speakers with AOS.

Purpose of This Study

Given that research on CAS and AOS in Chinese speakers is just emerging, there is a need to review the diagnostic methods reported in studies focusing on Chinese speakers with CAS and AOS to inform future investigations. The purpose of this study was to report the diagnostic methods that have been used on Chinese speakers with CAS or AOS and to examine the quality of these diagnostic methods in terms of assessment tasks, assessment methods (perceptually judged or based on quantitative data and criteria), diagnostic criteria and accuracy, and information about examiner qualifications. In addition to reviewing articles written in English, we aimed to extend our review to articles written in Chinese. To the best of our knowledge, this is the first ever review of CAS and AOS in Chinese speakers including articles written in Chinese.

Method

This scoping review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews guidelines (Tricco et al., 2018) to report the findings, which resulted in assessing the scope of the literature and synthesizing evidence on the topic of assessment and diagnostic methods for Chinese speakers with AOS.

Eligibility Criteria

The articles included in this review focused on speakers with CAS and/or AOS. The inclusion criteria were as follows: (a) published in peer-reviewed or non–peer-reviewed journals; (b) published between January 1980 and October 2020; (c) written in English, or traditional or simplified Chinese; and (d) included at least one Chinese speaker who had been diagnosed with AOS or CAS. Unpublished master's or doctoral theses with a focus on CAS and/or AOS were also included. We decided to include articles from 1980 because the development of speech-language pathology in Chinese-speaking countries (e.g., China) started in the 1980s. Articles were excluded if they were (a) animal studies; (b) commentaries, opinion, or review articles; or (c) not focused on CAS or AOS (e.g., focused on nonverbal apraxia). The articles in this review were maximized with this set of criteria.

Information Sources and Search

A total of 13 electronic databases were searched, including seven databases for English and six for Chinese articles. The seven English databases were (a) AMED Allied and Complementary Medicine, (b) Cumulative Index to Nursing and Allied Health Literature, (c) Embase, (d) Medline, (e) PubMed, (f) Scopus, and (g) Web of Sciences. The six Chinese databases were (a) Chinese Electronic Periodical Services via Airiti Library, (b) Chinese Social Sciences Citation Index, (c) China Academic Journal Full-text Database via CNKI, (d) National Digital Library of Theses and Dissertations in Taiwan, (e) Index to the Taiwan Periodical Literature System, and (f) Wan Fang Data.

Different search terms were used for English and Chinese electronic databases. The search terms used in English databases were as follows: ("apraxia" OR "dyspraxia" OR "apraxia of speech" OR "childhood apraxia of speech" OR "apraxic speech" OR "dyspraxic speech" OR "verbal dyspraxia") AND ("Chinese" OR "Mandarin" OR "Cantonese"). The search terms used in Chinese databases were as follows: "言語失用" (traditional Chinese script, meaning "speech apraxia") OR "言语失用" (simplified Chinese script, meaning "speech apraxia"). These Chinese terms were used to refer to disorders in both children and adults with no other terms noted in Chinese research or clinical practice. In addition to systematic searches of databases, additional sources of evidence included the reference lists of included articles. The search results were imported to EndNote (a software citation manager) and Rayyan (an online web manager for systematic reviews) for screening of eligibility.

Selection of Sources of Evidence

Screening was performed to determine the eligibility of searched articles. All three authors participated in the screening, data collection, and quality assessment of diagnoses.

To achieve agreement among the three authors on procedures and standards for screening, data extraction, and quality assessment of diagnoses, a prescreening exercise was conducted. In the exercise, five English articles with titles, abstracts, and keywords randomly selected from the English databases were distributed to the authors. The authors independently decided whether the articles met the inclusion criteria for full-text review based on the titles, abstracts, and keywords. All three authors agreed to include three out of five articles with 100% agreement. One English article that had been included in this manner was randomly selected for trial data extraction and quality assessment. The authors were instructed to independently extract data by using an Excel spreadsheet and to rate the quality of the diagnoses based on a predeveloped rating scale (see Table 1). Disagreement on the quality assessment was resolved in the follow-up online meeting, resulting in a consensus on screening, data extraction, and quality assessment procedures and standards. The consensus rating was archived as the final decision on the quality assessment of that article. A second trial of the exercise was subsequently carried out to confirm agreement on procedures and standards. Another five English articles with titles, abstracts, and keywords randomly selected from the databases were distributed to the authors for screening, data extraction, and quality assessment. The authors agreed to include four out of five articles with 100% agreement. One of the four articles was randomly selected and given to the authors for data extraction and quality assessment. Only two out of three authors (66.7%) agreed on the rating. Therefore, a final round of data extraction and quality assessment was performed. In this round, another English article that was agreed to be included was randomly selected and given to the authors. All three authors agreed on the rating at the 100% level on screening and quality assessment in this round. Another prescreening exercise was conducted between the first and second authors for data extraction and quality assessment of a randomly selected Chinese article because they were the only Chinese speakers who screened and reviewed articles written in Chinese. Agreement at the 100% level was achieved. Thenceforth, the authors screened articles, extracted data, and rated the quality of diagnoses independently.

Data Charting Process

A Microsoft Excel spreadsheet was used for extracting data from the articles that met the eligibility criteria. The data were extracted mainly from the "participants" or "confirmation of diagnosis" sections of the articles. Other parts of the articles, such as "limitations," were also reviewed to extract possible additional information related to diagnostic methods. The data were entered into the Excel spreadsheet for further qualitative and quantitative analyses. The data items included (a) participants' ages and genders, (b) initial diagnoses, (c) comorbidities or coexisting condition(s), (d) assessment tasks and methods, (e) diagnostic criteria for AOS or CAS, (f) diagnostic accuracy of the tasks, and (g) information about the examiners' qualifications.

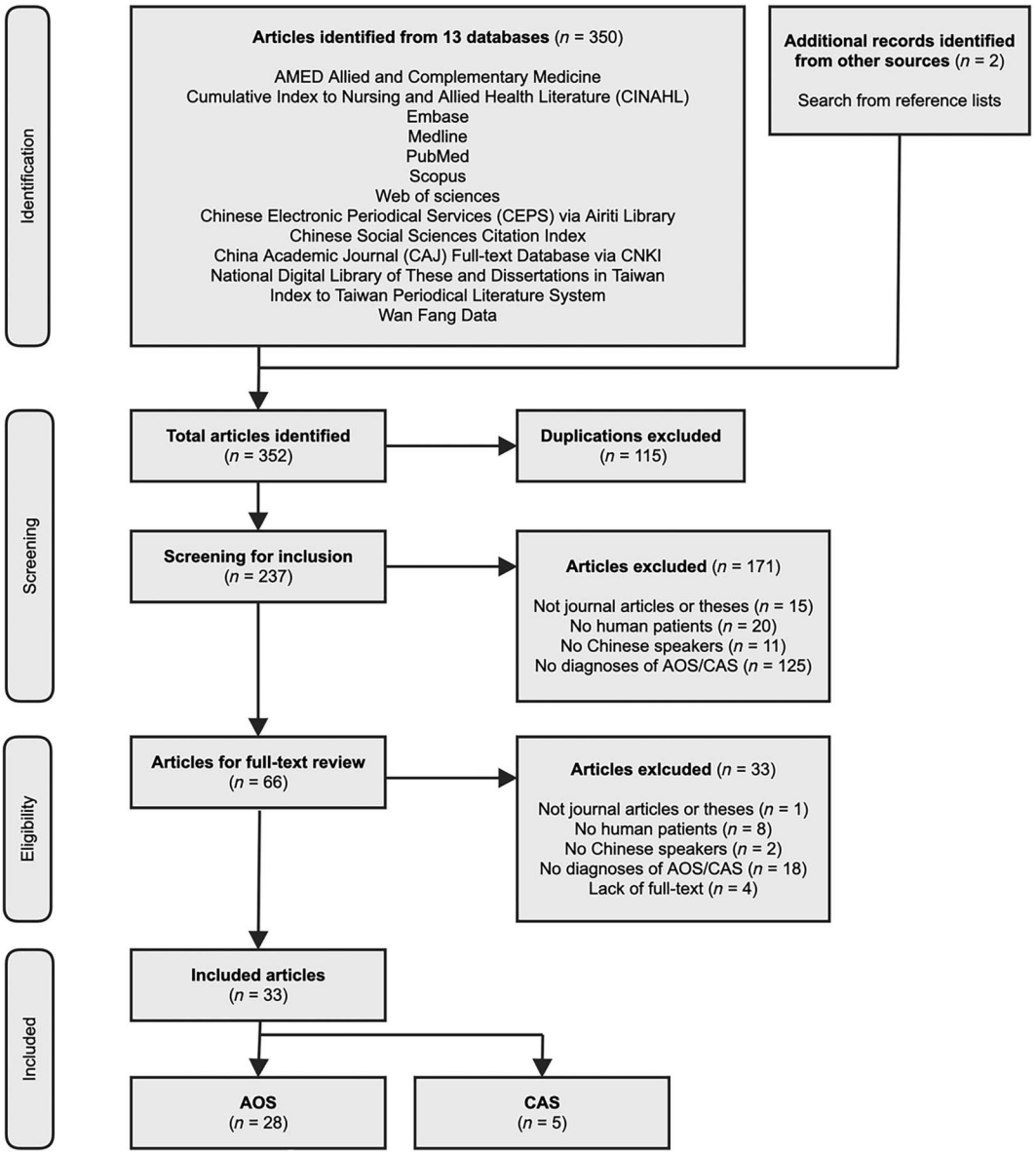
Figure 1 summarizes the article search procedures. There were 350 articles identified from 13 databases and two articles from the search of the reference lists of the included articles (n = 352). There were 217 English and 135 Chinese articles. After removing 115 duplicate articles, 125 English and 112 Chinese articles (n = 237) were screened based on titles, abstracts, and keywords. Among the 237 articles, 171 articles were excluded, including 15 articles that were not journal articles or theses (e.g., conference presentations), 20 articles that involved no human patients (e.g., animal studies, survey studies, review studies, or healthy speakers only), 11

Table 1. Levels of quality of the apraxia of speech or childhood apraxia of speech diagnoses in the articles included.

Method	Level of quality	Example
Reliable and valid tests or methods	Ia	Reliable and valid tests or methods + objective data + expert perceptual judgment
	Ib	Reliable and valid tests or methods + objective data
	Ic	Reliable and valid tests or methods + expert perceptual judgment
	Id	Reliable and valid tests or methods + nonexpert perceptual judgment
	Ie	Reliable and valid tests or methods only
Objective data/criteria	IIa	Objective data + expert perceptual judgment
	IIb	Objective data + nonexpert perceptual judgment
	IIc	Objective data only
Perceptual judgment	IIIa	Expert perceptual judgment only
	IIIb	Nonexpert perceptual judgment only
Unclear	IV	None

Note. Experts are defined as having at least 10 years of clinical experience in assessing or treating patients with apraxia of speech or childhood apraxia of speech.

Figure 1. Summary of article search procedures. AOS = apraxia of speech; CAS = childhood apraxia of speech.



articles that involved no Chinese patients, and 125 articles that involved no diagnoses of CAS or AOS. There were 66 articles left for full-text review. Thirty-three articles were excluded after full-text review for the following reasons: (a) not a journal article or thesis (n = 1), (b) no patients involved (e.g., review studies, commentaries, involved healthy speakers or typically developing children; n = 8), (c) no Chinese speakers involved (n = 2), (d) no diagnosis of CAS or AOS was made (n = 18), and (e) lack of an available full-text version (n = 4). We sent e-mails to the authors of articles for which we could not otherwise access full-text manuscripts, but no responses had been received as of the completion of this article. Thus, a total of 33 articles were included in this study, of which five articles had a focus on CAS and 28 articles had a focus on AOS.

A larger proportion of studies focused on AOS (n = 28, 84.8%) than those concerning CAS (n = 5, 15.1%). Among the five CAS articles, there were two master's theses, two articles published in peer-reviewed journals, and one from a non–peer-reviewed journal. The identified articles focusing on AOS were published in peer-reviewed journals (17/28, 60.7%) and journals for which the review process was unspecified (11/28, 39.3%). Details are listed in Tables 2 and 3. In the five studies with a focus on children with CAS, one study was conducted in Hong Kong

Table 2. Summary of the assessment methods, diagnostic criteria, and quality assessments of diagnostic processes for the five studies concerning Chinese speakers with CAS.

No.	Study	Participants			Assessment task	Assessment method (perceptual or objective)	Diagnostic criteria and diagnostic accuracy	Diagnosis was made by	Quality
		Diagnosis	*n*	Age or range (*M*), years					
1	E. C. H. Wong (2017)	CAS	2	3.5–6.75 (5.13)	An assessment protocol was designed for the study[a]	Perceptual	Perceptual judgment of the presence of 29 clinical features identified for Chinese by a group of experts plus ASHA's three consensus features.[a,b] No diagnostic accuracy was reported.	Two SLPs with more than 10 years of clinical experience with 100% agreement	IIIa
2	Liu et al. (2019)[c]	CAS	6	4.42–5.5 (5.36)	A list of 105 three-word probes[a]	Perceptual	Presence of AHSA's three consensus features[b]; not clear how they were determined. No diagnostic accuracy was reported.	One SLP	IIIb
3	Chen (2011)	CAS	5	3.17–5.75 (3.96)	The children were diagnosed with CAS before participating in the study. No information was provided.	Perceptual	Not specified. Clinical features were reported by the corresponding SLPs before enrolment, including delayed or limited babbling, nonspeech groping, inconsistent errors, sound substitutions or distortions, low speech intelligibility, slow speech rate, and inconsistent tone errors (especially in imitation). No diagnostic accuracy was reported.	One SLP per child[d]	IIIb
4	T.-W. Wang et al. (2015)[c]	CAS	1	6.0	Not specified	N.A.	Not specified. The article described two clinical features that were observed in the child (i.e., prosodic deficits and articulatory groping). No diagnostic accuracy was reported.	One SLP[d]	IIIb
5	Y. Wang et al. (2018)	CAS	1	2.42	Clinical observation. No further information was provided.	Perceptual	Not specified	Not specified	IV

Note. CAS = childhood apraxia of speech; ASHA = American Speech-Language-Hearing Association; SLP = speech-language pathologist; N.A. = not available.

[a]More information about the task, scoring method, and diagnostic criteria is presented in Appendix A. [b]ASHA's three consensus features include inconsistent errors, deficits in co-articulation, and prosodic impairment (ASHA, 2007). [c]The article was published in a peer-reviewed journal. [d]The speech-language pathologists were not otherwise involved in the study.

Table 3. Summary of the assessment methods, diagnostic criteria, and quality assessments of diagnostic processes for 28 studies concerning Chinese speakers with AOS.

No.	Study	Participants			Assessment task/ protocol/tool	Assessment method (perceptual or quantitative)	Diagnostic criteria and diagnostic accuracy	Diagnosis was made by	Quality
		Diagnosis	*n*	Age/range (*M* ± *SD*), years					
1	J. Wang et al. (2019)[a]	AOS	52 (37 M; 15 F)	24–73 (54.67)	An assessment developed based on the ABA-2 and Wambaugh et al. (2006)[b]	Perceptual	Not specified	Two SLPs with more than 10 years of experience in the evaluation and treatment of adult with aphasia and AOS	IIIa
2	Z. Wang & Li (2013)[a]	AOS	20 (19 M; 1 F)	30–78 (51.1)	Not specified	N.A.	Not specified	Two SLPs with more than 10 years of clinical experience	IIIa
3	Chen et al. (2020)	AOS	34 (27 M; 7 F)	47–70 (57.7)	Assessment standard suggested by CRRC[b]	Perceptual	Not specified	"An SLP"	IIIb
4	Gu (2011)	AOS	13 (6 M; 7 F)	40–73 (no mean provided)	Assessment standard suggested by CRRC[b]	Perceptual	Not specified	Not specified	IV
5	Kong et al. (2004)	AOS	29 (19 M; 10 F)	45–75 (65.5)	Assessment standard suggested by CRRC[b]	Perceptual	Not specified	Not specified	IV
6	H. Li et al. (2002)[a]	AOS	5 out of 100	37–78 (58.44 ± 14.36)	Assessment standard suggested by CRRC[b]	Perceptual	Not specified	Not specified	IV
7	H. Li et al. (2003)[a]	AOS	5 out of 100	19–79 (61.51 ± 16.23)	Assessment standard suggested by CRRC[b]	Perceptual	Not specified	Not specified	IV
8	X. Pan et al. (2006)[a]	AOS	8	45–75 (65 ± 3.5)	Assessment standard suggested by CRRC[b]	Perceptual	Not specified	Not specified	IV
9	Wan et al. (2004)	AOS	17 (14 M; 3 F)	32–73 (58.0)	Assessment standard suggested by CRRC[b]	Perceptual	Not specified	Not specified	IV
10	Wu et al. (2012)[a]	AOS	16 out of 22	29–76 (51.23)	Assessment standard suggested by CRRC[b]	Perceptual	Not specified	Not specified	IV
11	You et al. (2018)	AOS	38 (29 M; 9 F)	No age range provided (58.2)	Assessment standard suggested by CRRC[b]	Perceptual	Not specified	Not specified	IV
12	Zhang et al. (2010)[a]	AOS	8 out of 9	59–74 (63.4)	Assessment standard suggested by CRRC[b]	Perceptual	Not specified	Not specified	IV
13	Lu et al. (2020)[a]	AOS	30 (23 M; 7 F)	No age range provided (51.87)	Assessment standard suggested by CRRC[b] (Z. Wang & Li, 2013) + observation of clinical features suggested by Yang & Wang (2014)[b]	Perceptual	Not specified	Not specified	IV
14	Jiang, Yang, Xiang, Tang, et al. (2015)[a]	AOS	60 (39 M; 21 F)	15–72 (50.87)	Psycholinguistic Assessment in Chinese Aphasia (PACA1.0)[b]	Perceptual	Not specified	Not specified	IV
15	Jiang, Yang, Xiang, Chang, et al. (2015)	AOS	60 (20 M; 10 F)	15–72 (51)	PACA1.0[b]	Perceptual	Not specified	Not specified	IV
16	Jiang, Yang, Chang, et al. (2015)[a]	AOS	60 (39 M; 21 F)	15–72 (50.87)	PACA1.0[b]	Perceptual	Not specified	Not specified	IV
17	Sun et al. (2019)	AOS	20 (11 M; 5 F)	35–85 (58.4)	PACA1.0[b]	Perceptual	Not specified	Not specified	IV
18	You et al. (2019)[a]	AOS	42 (32 M; 10 F)	38–85 (57.62)	ABA-2	Perceptual	Not specified	Not specified	IV
19	Deng et al. (2014)[a]	AOS	1 (F)	59	Neuropsychological tests	Not specified	Not specified	Not specified	IV
20	Q. Li (2019)[a]	AOS	80 (48 M; 32 F)	31–72 (50.7)	Based on Tan et al. (2017)[c]	N.A.	Based on Tan et al. (2017)[c]	Not specified	IV

(*table continues*)

Table 3. *(Continued).*

No.	Study	Participants			Assessment task/ protocol/tool	Assessment method (perceptual or quantitative)	Diagnostic criteria and diagnostic accuracy	Diagnosis was made by	Quality
		Diagnosis	*n*	Age/range (*M* ± *SD*), years					
21	C. Pan and Mao (2000)	AOS	1 (M)	40.0	Vowel sequence, word sequence, and imitation of words	Perceptual	Not specified	Not specified	
22	Qin (2008)[a]	AOS	1 (M)	48.0	Vowel sequence, word sequence, and imitation of words	Perceptual	Not specified	Not specified	IV
23	Gao et al. (2012)[a]	AOS	118 (60 M; 58 F)	31–76 (53.4)	Not specified	N.A.	Not specified	Not specified	IV
24	D. Li et al. (2016)[a]	AOS	6	No age range provided (60.7)	Not specified	N.A.	Not specified	Not specified	IV
25	Ren et al. (2016)	AOS	30 (20 M; 10 F)	No age range provided (66.25 ± 11.55)	Not specified	N.A.	Not specified	Not specified	IV
26	Shi (1998)	AOS	25 out of 38	18–73 (49.26)	Not specified	N.A.	Not specified	Not specified	IV
27	Wei & Li (2000)[a]	AOS	1 (M)	42.0	Not specified	N.A.	Not specified	Not specified	IV
28	Zhang et al. (2020)	AOS	32 (23 M; 9 F)	No age range provided (59.31)	Not specified	N.A.	Not specified	Not specified	IV

Note. AOS = acquired apraxia of speech; M = male; F = female; ABA-2 = Apraxia Battery for Adults–Second Edition (Dabul, 2000); SLP = speech-language pathologist; N.A. = not available; CRRC = China Rehabilitation Research Centre.

[a]The article was published in a peer-reviewed journal. [b]More information about the tasks, scoring methods, and diagnostic criteria is presented in Appendix A. [c]Tan et al. (2017) did not make AOS or childhood apraxia of speech diagnoses. All of the participants involved were patients with aphasia.

(E. C. H. Wong, 2017), three were conducted in Taiwan (S.-P. Chen, 2011; Liu et al., 2019; T.-W. Wang et al., 2015), and one was conducted in Mainland China (Y. Wang et al., 2018). The Hong Kong study had two Cantonese-speaking children, whereas the participants from the rest of the studies spoke Mandarin as their main language for communication. All AOS studies were conducted in Mainland China with all of the participants speaking Mandarin. The studies included 15 children with CAS and 812 patients with AOS. Among the 15 children with CAS, there were 12 boys and three girls, with a mean age of 4.82 years (SD = 1.33). The demographics of the 812 patients with AOS were not clearly stated in all studies, so the gender distributions and means and standard deviations of age could not be calculated.

Synthesis of Results

A narrative synthesis of the data from the included studies was performed with respect to participant characteristics, assessment tasks, assessment methods, diagnostic criteria and accuracy, and information about the examiners. The main outcome of this review was the evaluation of diagnostic methods used for CAS or AOS in studies of Chinese speakers. Qualitative analysis was performed to determine the consistency or diversity of the diagnostic methods used for CAS or AOS. Quantitative analysis was performed to determine the quality of the diagnoses made in the studies. A quality assessment was developed in this review to critically evaluate the diagnostic processes reported in the literature concerning Chinese speakers with CAS or AOS. The three criteria adopted were as follows: reliable and valid tests or methods, objective data, and perceptual judgments. Diagnostic processes that involved any reliable and valid test or method were rated at the highest level (Level I). Level I was further divided into five levels, from the highest, Ia, to the lowest, Ie. A higher level of quality was assigned when more criteria were evident. For example, Ia was assigned to diagnostic processes that had three different criteria including reliable and valid tests or methods, objective data, and expert perceptual judgment. Ie was assigned to diagnostic processes that only met one criterion, reliable and valid tests or methods only. Levels Ib, Ic, and Id were ordered based on the second criterion used, objective data or criteria, which referred to any quantitative measures aimed to assess the symptoms. Diagnostic processes with objective data (Level Ib) were rated higher than those with perceptual judgment only (Levels Ic and Id), and expert perceptual judgment (Ic) was rated higher than nonexpert perceptual judgment (Level Id). Experts were defined in this review as having at least 10 years of clinical experience in assessing or treating patients with CAS or AOS. Diagnostic processes that included no reliable or valid tests or methods but with objective data or criteria were rated at Level II. Diagnostic processes that involved a combination of objective data and expert perceptual judgment of clinical features were rated as having the higher level of quality (IIa), whereas those that involved nonexpert perceptual judgment were rated at a lower level (IIb). Objective data only were rated at the lowest sublevel of Level II (IIc). Those processes that involved only perceptual judgment were rated at Level III. Studies in which diagnoses were based upon the judgments of clearly defined experts were rated higher (Level IIIa) than those based upon nonexpert perceptual judgments (Level IIIb). Articles that included unclear diagnostic processes were rated at Level IV. The levels of quality are presented in Table 1.

Results

Type of Studies

The identified CAS research in Chinese speakers ranged from understanding the genetic basis of CAS (Y. Wang et al., 2018), through identifying and evaluating the clinical features of CAS directly from children with CAS (S.-P. Chen, 2011) and indirectly via an expert panel (E. C. H. Wong, 2017), to examining treatment efficacy (Liu et al., 2019; T.-W. Wang et al., 2015; E. C. H. Wong, 2017). E. C. H. Wong (2017) demonstrated the treatment efficacy of a combined treatment approach (syllable repetition method plus modified dynamic temporal and tactile cueing) for two participants. Liu et al. (2019) investigated the efficacy of a combined treatment with integral stimulation and the minimal contrast approach. T.-W. Wang et al. (2015) reported a case study using a speech production exercise that systematically increased the complexity of stimuli via visual supports.

The identified AOS articles varied with respect to investigation types. The search identified a relatively large body of research on treatment efficacy for adult patients with AOS diagnoses (22/28, 78.6%). The diagnostic protocols described in these works were assessed in this review. There were five studies investigating language therapy for aphasia and 17 on AOS treatment. Several AOS intervention studies investigated Rosenbek's eight-step treatment (Rosenbek et al., 1973) and melodic intonation therapy (Sparks et al., 1974) in Chinese speakers (C. Chen et al., 2020; Q. Li, 2019; Qin, 2008; Wei & Li, 2000). Other approaches included action observation therapy based on mirror neuron theory (You et al., 2018, 2019; Zhang et al., 2020), transcranial direct current stimulation (J. Wang et al., 2019), and low-frequency repetitive transcranial magnetic stimulation (Ren et al., 2016). In addition, some studies investigated the efficacy of using different

supporting methods (e.g., providing mouth shapes as visual cues, simultaneous production, nonspeech oral motor exercises, and gestural cues) in speech production exercises (Gu, 2011; Lu et al., 2020; C. Pan & Mao, 2000; Sun et al., 2019). There were treatments that have not been reported in the English literature as well, such as applying Chinese medicine (Gao et al., 2012) and acupuncture with speech production exercises (Jiang, Yang, Chang, et al., 2015; Jiang, Yang, Xiang, Chang, et al., 2015; Jiang, Yang, Xiang, Tang, et al., 2015). Among the remaining six articles, two reported speech dysfunction in patients with acquired AOS after stroke (X. Pan et al., 2006) or degenerative diseases (D. Li et al., 2016). One article reported the incidence rate of aphasia after stroke (H. Li et al., 2003), one reported an assessment method (Z. Wang & Li, 2013), one reported a gene mutation in Alzheimer's disease and found AOS as one of the symptoms (Deng et al., 2014), and the last article investigated swallowing disorders in patients after stroke (H. Li et al., 2002).

CAS Assessment

Table 2 summarizes the assessment tasks, assessment methods, diagnostic criteria and accuracy, and quality of the diagnostic processes of the five studies concerning Chinese speakers with CAS. The diagnostic process varied across the five studies. Three out of five studies gave CAS diagnoses to the participants, whereas the participants of the other two studies were diagnosed with CAS before being enrolled in the studies. In the three studies with confirmation of diagnosis, E. C. H. Wong (2017) provided the details of the assessment protocol, which included assessment of language skills, articulation skills, tone identification skills, motor speech skills, and nonspeech oral motor skills, as well as observation of prosody (see Appendix A). Two speech-language pathologists (SLPs) with more than 10 years of clinical experience in assessing or treating children with CAS perceptually judged the presence of ASHA's three consensus features and 29 clinical features on a checklist. These 29 features had been identified by a group of experienced SLPs in Hong Kong (E. C. H. Wong et al., 2020) and are presented in Appendix B. Liu et al. (2019) administered a phonology test with 105 three-word probes. The CAS diagnoses were made by an SLP based on perceptual judgment of the presence of the three ASHA consensus features. Y. Wang et al. (2018) reported the *FOXP2* gene mutation in a Mandarin-speaking boy who had CAS. The speech diagnosis was made based upon clinical observation, but the clinical features observed were not specified.

No diagnostic processes were carried out in the other two studies. S.-P. Chen (2011) included five children with CAS. The diagnoses were made by different SLPs before their enrolment, and no confirmation of the diagnoses was carried out. There was also no report of the tasks or methods used for the CAS diagnoses, but the author reported that the children had delayed or limited babbling, nonspeech groping, inconsistent errors, sound substitutions or distortions, low speech intelligibility, slow speech rate, and/or inconsistent tone errors (especially in imitation). These seven clinical features were proposed to have high sensitivity and specificity for diagnosing CAS in Mandarin-speaking children. T.-W. Wang et al. (2015) involved a child with CAS who was diagnosed by an SLP who was not otherwise involved in the study. No confirmation of the diagnosis was reported. Although no diagnostic criteria were reported, the authors described two clinical features that were observed, prosodic deficits and articulatory groping.

Quality of CAS Diagnoses

The quality of CAS diagnoses was rated based on the assessment tasks, methods, and diagnostic criteria and accuracy as well as the qualifications of the examiners reported in the studies. There was no report of valid and reliable tests or objective measures for CAS diagnosis in Chinese speakers. All of the diagnoses were made based upon perceptual judgment. Therefore, no Level I or II diagnoses were rated. One study reported details of the assessment task, the criteria for CAS diagnoses, and the involvement of two experts (E. C. H. Wong, 2017). The quality of the CAS diagnoses in this study was rated as IIIa (expert perceptual judgment only). Three studies reported that the diagnoses were made by SLPs without specifying their experience (S.-P. Chen, 2011; Liu et al., 2019; T.-W. Wang et al., 2015). The quality of assessment was thus rated as Level IIIb, as nonexpert perceptual judgment was assumed. The last study (Y. Wang et al., 2018) failed to report information about the examiner and was rated as Level IV.

AOS Assessment

Table 3 summarizes the assessment tasks, methods, diagnostic criteria and accuracy, information about the examiners, and quality assessments of the diagnostic processes of 28 studies concerning Chinese speakers with AOS. A plurality (10/28, 35.7%) used the assessment tasks suggested by the CRRC. There have been several versions of the CRRC assessment standard. The earliest version identified in the literature was proposed by S. Li and Shirasaka (1994; in a book that was referenced by the included studies, but the book itself was not included in this review), but no details of the tasks could be found. Wei and Li (2000) and S. Li (2008; another book that was not included in this review but was referenced by an identified study) reported two highly similar tasks. Another

variation of the CRRC assessment standard was reported by Z. Wang and Li (2013), in which the use of a set of Chinese materials "referred from" (Z. Wang & Li, 2013, English abstract, p. 70) the MSE (Wertz et al., 1991) was proposed. Detailed information about the variations on the CRRC assessment standard is provided in Appendix A. These methods consist of using perceptual judgments of participants' performance. No diagnostic criteria or accuracy have been reported for these tasks. Specifically, two of the 10 studies (H. Li et al., 2002, 2003) cited S. Li and Shirasaka (1994), two studies (Wan et al., 2004; Zhang et al., 2010) cited Wei and Li, one study (Wu et al., 2012) cited S. Li (2008), and five studies (C. Chen et al., 2020; Gu, 2011; Kong et al., 2004; X. Pan et al., 2006; You et al., 2018) mentioned the use of the CRRC assessment standard without specifying the version. Unfortunately, none of the studies reported any diagnostic criteria used or any information about diagnostic accuracy.

Among the 28 identified studies of patients with AOS, one study (Lu et al., 2020; 1/28, 3.6%) used a combination of the CRRC assessment standard (Z. Wang & Li, 2013) and additional observations of the clinical features suggested by Yang and Wang (2014), in which eight clinical features were adapted from Wambaugh et al. (2006). These included (a) slow speech rate; (b) prolongations; (c) increased pauses; (d) intrusive schwa; (e) inappropriate prosody, especially lexical stress and equal stress; (f) sound distortions; (g) distorted substitutions; and (h) consistent articulatory errors. No diagnostic criteria, cutoff scores, or diagnostic accuracy were reported by Lu et al. (2020) with respect to the diagnostic processes used.

Four studies (Jiang, Yang, Chang, et al., 2015; Jiang, Yang, Xiang, Chang, et al., 2015; Jiang, Yang, Xiang, Tang, et al., 2015; Sun et al., 2019; 4/28, 14.3%) reported the use of the speech motor planning module of the Psycholinguistic Assessment in Chinese Aphasia (PACA1.0), which is a computer program for Chinese speakers with aphasia. Examiners perceptually judge the participant's performance and enter scores into the computer. However, no validity has been reported for this test module, nor a cutoff point for the diagnosis of AOS. Detailed information about the content and scoring details of the PACA1.0 module were found in the Home of Rehabilitation Therapist (2020) and are presented in Appendix A. No diagnostic criteria or accuracy cutoff levels required for a diagnosis of AOS were reported for the module in these four studies or in other literature.

You et al. (2019) reported using the ABA-2 (Dabul, 2000; 1/28, 3.6%), which is a standardized assessment tool for English speakers with AOS. Unfortunately, no information about its content, validity, or reliability were reported with respect to the application of this measure to Chinese speakers, and the authors did not provide any cutoff score for the diagnosis of AOS.

Among the remaining 12 studies, one study (J. Wang et al., 2019) used an assessment that was developed based on the ABA-2 (Dabul, 2000) and Wambaugh et al. (2006). Information about this method is presented in Appendix A. However, no diagnostic criteria or accuracy were reported for this study. Deng et al. (2014) reported the use of a neuropsychological test, but no details about the content or the criteria for making AOS diagnoses were reported. Q. Li (2019) used the tasks and criteria suggested by Tan et al. (2017), who did not report any AOS diagnoses in their study. Two studies (C. Pan & Mao, 2000; Qin, 2008) reported the use of vowel sequences, word sequences, and imitations of words as assessment tasks without reporting the diagnostic criteria, diagnostic accuracy, or information about the examiners. Seven other studies (Gao et al., 2012; D. Li et al., 2016; Ren et al., 2016; Shi, 1998; Z. Wang & Li, 2013; Wei & Li, 2000; Zhang et al., 2020) did not specify their assessment tasks, methods, diagnostic criteria and accuracy, or information about the examiners.

Quality of AOS Diagnoses

The quality of the AOS diagnoses made in the 28 studies was rated based on the assessment tasks, assessment methods, and diagnostic criteria and accuracy as well as the qualifications of the examiners. There were no Level I or Level II diagnoses as there were no reports of valid and reliable tests or objective measures for the AOS diagnoses of Chinese speakers. Because all of the specified assessment tasks involved only perceptual judgments of correctness and/or ratings, the studies were rated as Level III. Three studies reported information about who made the diagnoses. Two studies specified the experience of the examiners. J. Wang et al. (2019) reported that the AOS diagnoses of the participants were made by two SLPs with more than 10 years of experience in the assessment and treatment of adults with aphasia and AOS, whereas Z. Wang and Li (2013) involved two SLPs with more than 10 years of clinical experience in rating the assessment performance of participants. The AOS diagnoses were assumed to be made by these two SLPs. Although no diagnostic criteria were reported in either of these studies, the quality of the diagnoses of the studies was rated as IIIa, as expert perceptual judgment was involved. C. Chen et al. (2020) reported that the diagnoses of AOS were made by an SLP, without stating their experience. This study was rated as Level IIIb, as nonexpert perceptual judgment was assumed. There were seven studies that failed to report any information about the assessment tasks, assessment methods, and diagnostic criteria and accuracy or the qualifications of the examiners and were

rated as Level IV (unclear). There were 18 studies that reported assessment tasks but failed to report diagnostic criteria, diagnostic accuracy, or the qualifications of the examiners. These studies were also rated as Level IV (unclear).

Discussion

The purpose of this study was to review the diagnostic methods that have been used in research on AOS and CAS in Chinese speakers and to examine the quality of these diagnostic methods, in terms of the assessment tasks, assessment methods, and diagnostic criteria and accuracy as well as the qualifications of the examiners.

CAS in Chinese

There were only five studies investigating CAS in Chinese speakers. These five studies varied with respect to their assessment tasks. There were no reports of standardized tests or objective measures used for diagnosing Chinese speakers with CAS. Although further studies are strongly warranted in both English and Chinese, there are available tests in English that provide valuable information on speech motor planning and programming skills (e.g., DEMSS: Strand & McCauley, 2019; DEAP: Dodd et al., 2006), acoustic measures used for English speakers with CAS (e.g., PM; Shriberg et al., 2017a), and maximum repetition rate data for speakers of Western languages producing monosyllables and trisyllabic sequences in MPT (Thoonen et al., 1999). Thus, there is a discrepancy between the available tools for diagnosing Chinese speakers with CAS and those used in clinical practice internationally.

Detailed information about the assessment tasks was reported in some of the identified studies. Although these tasks are not standardized tests with reported diagnostic accuracy, some of them are aligned with the English literature, such as the imitation of polysyllabic words and the DDK tasks used by E. C. H. Wong (2017). Moreover, a potential diagnostic tool, Wong's checklist of 29 clinical features (E. C. H. Wong, 2017), is also available for Chinese speakers with CAS. Although the relatively objective diagnostic approach proposed by E. C. H. Wong et al. (2021) and E. C. H. Wong et al. (2022) was not reviewed in this study, these studies have provided a reference for diagnosing CAS in Cantonese speakers. The alignment between the Chinese studies of CAS and the international standard is evident. However, there is still no available diagnostic tool for Chinese speakers with CAS as there have been no reports of the sensitivity or specificity of these assessment tasks or this checklist for making CAS diagnoses.

Two studies reported their diagnostic criteria for CAS in detail. Liu et al. (2019) and E. C. H. Wong (2017) reported the use of ASHA's three consensus features in their diagnostic criteria. Given that these studies were independent of each other and were conducted in Taiwan and Hong Kong, respectively, the use of ASHA's three consensus features seems to be the only consensus in assisting with the diagnosis of CAS in Chinese speakers. Despite the documented validity of the three consensus features in English-speaking children with CAS (Chenausky et al., 2020), it is still unclear if these features are valid for differentiating CAS from SSDs, even in English. However, regarding the manner of observation and interpretation, many of those English CAS studies that have used ASHA's three consensus features as their only diagnostic criteria have provided working definitions of each feature and of their criteria (e.g., Grigos & Case, 2018; Thomas et al., 2016). In contrast, E. C. H. Wong (2017) suggested documenting each feature at least once in the initial assessment, but no clear working definitions were provided. Thus, it is not clear how the three features were defined and observed in that study.

Despite the presence of detailed information about assessment tasks and clear descriptions of diagnostic criteria, the only assessment method involved in the Chinese CAS studies was perceptual judgment. Therefore, the quality rating relied in this case, also, on the experience levels of the examiners. The results suggested that the current highest standard for CAS diagnosis in Chinese speakers is still expert perceptual judgment based on a list of clinical features.

AOS in Chinese

Among the 28 identified studies concerning Chinese-speaking adults, a majority of articles (89.3%) failed to clearly report the manner in which participants were diagnosed with AOS. Regarding assessment methods, there were no objective measures identified. All of the reported assessment tasks required only perceptual judgment or ratings of performance. Given the use of objective measures to supplement expert perceptual judgment in making AOS diagnoses in English speakers, such as PVIdur and WSD, there is a discrepancy between the available tools for Chinese speakers with AOS and those available for clinical practice internationally.

There were three assessment tasks or protocols identified in the studies. They were the assessment standards suggested by the CRRC (S. Li, 2008; Z. Wang & Li, 2013; Wei & Li, 2000), the PACA1.0, and an assessment based on the ABA-2 (J. Wang et al., 2019). However, there were no reports of the diagnostic accuracy of these tasks or protocols. Therefore, they are not yet considered to be tests with psychometric properties but rather as informal assessment tasks or protocols. Apart from assessment tests or tasks, observations of clinical features were

reported. For example, Lu et al. (2020) included patients with AOS who had difficulty producing speech sounds, had limited phonetic inventories, produced rigid language, or were mute. However, the sensitivity or specificity of these features in Chinese speakers with AOS is unknown. Therefore, no valid diagnostic tool is currently available for Chinese speakers with AOS.

Although assessment tasks were described or referenced in some studies, no diagnostic criteria or accuracy levels were reported in the original articles or the individual studies. Thus, it is unknown how the assessment data were interpreted and how the AOS diagnoses were given based on the assessment results. Therefore, the quality assessment of the diagnoses relied heavily on the reported experience levels of the examiners. There were only three studies that reported this information about the examiners, of which two studies reported the involvement of experts who had more than 10 years of clinical experience and one study reported the involvement of an unspecified SLP, who was assumed to be a nonexpert. It is important to note that the criteria used by the experts and SLPs in the studies were unknown. These results suggest that the current highest standard of AOS diagnosis of Chinese speakers is relying on expert opinions with unknown criteria.

There may be some reasons for the lack of diagnostic criteria identified for Chinese speakers with AOS. Given that there is no empirical diagnostic marker for AOS reported for either Chinese or English speakers, there is no evidence-based method for interpreting the assessment data, resulting in the absence of diagnostic criteria. It is possible that some diagnostic criteria do exist but were not reported in the articles, as most of the studies (22/28, 78.6%) focused on treatment. The descriptions of the diagnostic criteria may also have been omitted due to the stringent word limits of Chinese journals. Moreover, the type of information reported in the articles may have been influenced by precedent. A large proportion of the AOS studies adopted the assessment methods suggested by the CRRC, which is a leading speech-language pathology research center in China; the authors are pioneers of AOS research in Chinese. The proposed CRRC assessment tasks and the associated article may be considered to be gold standards for both AOS assessment and journal publication. Because no set of diagnostic criteria was reported in the original article, there may have been a misinterpretation that it is not necessary to report diagnostic criteria.

Although no diagnostic criteria were reported in any identified studies, the use of specific clinical features appears to have varied within the AOS diagnostic processes. For example, You et al. (2019) reported that the scoring system of the CRRC assessment tasks requires observation of "articulatory groping behavior" and "vowel errors" and that the presence of the former feature results in a lower score than the presence of the latter. This scoring system implies that "articulatory groping behavior" is considered to be a more severe feature than "vowel errors" in adults with AOS. Without any independent evidence of the relationships among these clinical features, it is unclear whether the differential weightings of these clinical features are valid.

Future Investigations

CAS

CAS in Chinese speakers requires further investigation. First, a valid and reliable tool and objective measures for Chinese speakers with CAS are needed for both research and clinical practice. In this review, two assessment protocols have been identified, from Liu et al. (2019) and E. C. H. Wong (2017). These protocols could be further developed as assessment batteries for Cantonese and Mandarin speakers with CAS, respectively. In addition, E. C. H. Wong et al. (2021) examined lexical tone production skills in Cantonese speakers with CAS by using the tone sequencing task (TST). The authors reported preliminary findings that tone accuracy and consistency, as well as the acoustic durations of different tone sequences produced by children with CAS, were significantly different from those with speech and language impairment or typical development. The authors suggest that the TST may be a potential tool that includes both qualitative and quantitative measures for Cantonese speakers with CAS. However, it is important to note that this study reported on only three children with CAS.

Available assessment tasks or tests for English speakers with CAS may provide a basis for further understanding the nature of CAS in Chinese speakers and may contribute to the development of a valid and reliable test or quantitative measures. For example, temporally based acoustic measures aimed at yielding quantitative data, such as the MPT (Diepeveen et al., 2019), the PM (Shriberg et al., 2017a), the Pairwise Variability Index (Ballard et al., 2012), and WSD measures (Haley et al., 2012), allow investigation of speech duration features in Chinese speakers with CAS. Application of these measures to Chinese speakers with CAS may indicate whether Chinese and English CAS are manifested similarly as a temporally affected speech disorder, which has been assumed but never explored empirically. However, caution is advised as the individual measures have their own limitations. For example, MPT performance varies across speakers of different ages and genders (Karlsson & Hartelius, 2021) and within healthy speakers (Ziegler et al., 2019). Such variation may also be found in children. Also, MPT has been suggested to be accurate for differentiating CAS from dysarthria, but it is unclear whether it can differentiate children with CAS from those who are typically developing

or those with other types of SSDs, such as phonological disorders. The PVIdur is only sensitive in multisyllabic words with strong–weak stress patterns; the WSD measure is relatively stable, but it has not been explored in children with CAS; and the PM still needs large-sample investigations to confirm its validity.

Nonacoustic measures such as the SRT (Shriberg et al., 2012) and the Single-Word Test of Polysyllables (Gozzard et al., 2006) convert perceptual judgments to a quantitative scale. These tasks, which have been empirically studied in English speakers, have shown validity for measuring the underlying deficits of CAS (i.e., speech motor planning and/or programming impairment) and may be applicable to Chinese speakers with CAS. Lastly, the DEMSS is a test that has sufficient psychometric properties, and it provides a method for using dynamic assessment to identify children with CAS. However, use of the DEMSS with Chinese speakers should be attempted with caution because of the prosodic differences between English and Chinese.

Importantly, it has been suggested that researchers or clinicians should not rely on any single measure when a CAS diagnosis is required (Preston et al., 2021). Clinical checklists (e.g., Strand 10-point checklist; Shriberg et al., 2011) may still play a role in guiding the selection of assessment tasks. Thus, an assessment that both aims to actively observe different clinical features and collects a variety of objective data is recommended. Murray et al. (2015) have provided a relatively objective approach to serve this purpose. In addition to the use of a valid checklist, the authors also suggested using a polysyllabic word production task (i.e., the Single-Word Test of Polysyllables; Gozzard et al., 2006) to calculate the percentage of stress matches, an oral motor examination (Robbins & Klee, 1987) to quantify oral motor movements for both speech and nonspeech functions, and DDK tasks to evaluate speech motor planning and programming skills. With the advancement of empirical evidence, the gold standard for making a CAS diagnosis is shifting from relying heavily on expert perceptual judgment of clinical features (Maas et al., 2012; Murray et al., 2015) to a combination of objective measures and expert perceptual judgment. This change in the gold standard provides a more objective approach to the diagnosis of CAS.

Direct investigation of CAS in Chinese speakers is also recommended. The identified studies have provided some important information that may help guide future investigations of CAS clinical features in Chinese. S.-P. Chen (2011) showed that seven clinical features, including "limited phonetic inventory," "vowel errors," "inconsistent errors," "difficulty in sequencing phonemes and syllables," "groping behaviours," "inappropriate prosody," and "reduced diadochokinetic rate," could be used for assessing CAS in Mandarin-speaking children given their high sensitivity (80%) and specificity (100%). E. C. H. Wong (2017) and E. C. H. Wong et al. (2020) compared the clinical features found in English and Cantonese speakers with CAS and proposed a list of 29 clinical features for clinical observation in Cantonese. The results of these studies have provided potential directions for the diagnosis of CAS in Chinese speakers.

AOS

The results of this review suggest that there is no available psychometrically evaluated tool for AOS in Chinese speakers. The current diagnostic standard consists of relying on expert perceptual judgment without a clear set of criteria. Future investigations are therefore needed, particularly for developing a valid and reliable assessment tool or quantitative measures. Potential assessment tasks were identified in this review. The different variations on the assessment methods suggested by the CRRC (S. Li, 2008; Z. Wang & Li, 2013; Wei & Li, 2000) are potential candidates because they were the most popular methods identified in the literature and have gained attention from both researchers and clinicians. The latest version has also been suggested to be suitable for Chinese speakers (Z. Wang & Li, 2013). The assessment methods used by Lu et al. (2020) and J. Wang et al. (2019) were designed based on existing resources for English speakers with AOS, the ABA-2 (Dabul, 2000), and Wambaugh et al.'s (2006) study. However, it is important to note that some elements of the ABA-2 and some features proposed by Wambaugh et al., such as "consistent articulatory errors," are no longer considered appropriate for AOS diagnosis due to the advancement of research evidence. Caution should be used when adopting these measures in future investigations of AOS. In addition, some other methods that have been studied and used in English speakers with AOS may be applicable to Chinese speakers with AOS. For example, the ASRS (Strand et al., 2014) may be a potentially useful diagnostic tool to perceptually rate the presence and severity of AOS characteristics in Chinese speakers. Because most of the ASRS items are nonlinguistic clinical features (e.g., sound distortions, slow speech rate, articulatory groping, and increased sound errors with increased utterance length or articulatory complexity), it is believed that the ASRS may be applicable to Chinese speakers with AOS, with the assumption that motor speech disorders across languages share the same underlying deficits but with different clinical manifestations (Whitehill, 2010; E. C. H. Wong et al., 2020). Also, adaptation of the ASRS to Chinese-speaking populations may provide a practical diagnostic criterion, which is currently absent from the literature.

In addition, quantitative measures may provide an objective approach for AOS diagnoses. Although several temporally based acoustic measures, such as PVIdur and

WSD, are available for English speakers with AOS, adaptation of these measures to Cantonese speakers requires extra caution due to the linguistic differences between English and Chinese (especially stress-timed vs. syllable-timed rhythm). Further investigations of these measures in Chinese speakers are urgently needed. Because Chinese is a syllable-timed language with lexical tones, acoustic measures of pitch variation may provide insight for Chinese AOS diagnosis. E. C. H. Wong et al. (2021) has provided a method and preliminary findings for a lexical tone acoustic measure for identifying CAS in Cantonese speakers. Further investigation into the use of similar acoustic measures as quantitative methods for the diagnosis of AOS in Chinese is suggested. In summary, more investigations of applications of existing empirical methods to Chinese speakers with AOS should be conducted before actual clinical implementation.

It is striking that very few of these studies were intended to investigate the clinical features of AOS in Chinese speakers. There were only two studies that had a focus on identifying AOS symptoms (D. Li et al., 2016; X. Pan et al., 2006). Furthermore, the main subject of these two studies was Chinese patients with aphasia and primary progressive aphasia. AOS was present in some of the participants as a coexisting disorder. Neither article described the features of AOS in detail. Thus, the clinical features of AOS in Chinese patients remain unclear. Moreover, a recent survey study investigating the clinical practice of AOS diagnosis in Cantonese speakers in Hong Kong showed that SLPs in Hong Kong make AOS diagnoses based on clinical features reported in English speakers with AOS (T. O. K. Wong et al., 2021). The study also reported a lack of consensus on the selection of assessment tasks and diagnostic criteria among the clinicians. Without a clear understanding of the clinical features of AOS in Chinese speakers, it is almost impossible to achieve a valid and accurate diagnosis. There is a need for direct investigations into the clinical features of Chinese speakers with AOS, as well as development of a valid and reliable assessment tool.

Cross-Linguistic Investigation

Although direct investigations of CAS and AOS in Chinese speakers are suggested, cross-linguistic investigations of both CAS and AOS are even more urgently needed. First, as mentioned, the claim that motor speech disorders share the same underlying deficits but with different clinical manifestations should be empirically tested in speakers with CAS and AOS.

Examination of the language universality and specificity of the features can be conducted in a cross-linguistic approach. Temporal deficits in English speakers with CAS result in lexical stress errors, whereas difficulty with pitch-variation skills is proposed to affect Chinese speakers with CAS and to result in lexical tone errors (E. C. H. Wong et al., 2021). Investigations of Chinese–English bilingual speakers may provide a deeper understanding of how the same underlying deficits may be manifested in different languages. Similarly, other clinical features, including vowel errors, inconsistent errors, and syllable segregation, may also be explored in this way. The results of such investigations will enrich our understanding of CAS and AOS in different languages.

Limitations

This study was limited in several ways. First, this is the first ever study aiming to review diagnostic standards in Chinese speakers with CAS or AOS, especially including articles written in Chinese. However, in the absence of official correspondences between Chinese and English terminology, the translations from Chinese to English were mainly based on the knowledge of the first two authors. It is possible that this process introduced some bias. Second, due to the lack of replies to the authors' requests for full-text articles, some articles have not been reviewed. The included articles may not represent all of the evidence from the literature.

Conclusions

This study is the first ever review with the aim of identifying and evaluating diagnostic methods for identifying CAS or AOS in Chinese speakers. There were 33 articles identified from the English and Chinese literature, with 15.1% concerning children with CAS and 84.8% concerning patients with AOS. A variety of assessment tasks and diagnostic criteria were used among the studies. Unfortunately, the results of the qualitative assessment revealed no high-quality approaches to diagnosing CAS or AOS in Chinese speakers in these studies due to the lack of valid and reliable tests or methods and quantitative measures. The current standard for diagnosing CAS or AOS in Chinese speakers is expert perceptual judgment. However, the existing studies have provided bases for future development of better assessment tools. Further investigations should focus on the development of valid and reliable assessment tools, the cross-linguistic investigation of the genetic and neural bases as well as the clinical features of CAS or AOS in Chinese and English speakers, and the eventual evaluation of treatment efficacy for these populations.

Author Contributions

Eddy Chun Ho Wong: Conceptualization (Lead), Data curation (Equal), Formal analysis (Equal), Methodology (Equal), Writing – original draft (Lead), Writing – review &

editing (Lead). **Min Ney Wong:** Conceptualization (Supporting), Data curation (Equal), Formal analysis (Equal), Methodology (Supporting), Supervision (Lead), Writing – original draft (Supporting), Writing – review & editing (Equal). **Shelley L. Velleman:** Data curation (Equal), Formal analysis (Equal), Methodology (Supporting), Supervision (Supporting), Writing – original draft (Supporting), Writing – review & editing (Supporting).

Acknowledgments

The authors are sincerely grateful to Yiting Chen for her assistance in searching and managing articles.

References

Allison, K. M., Cordella, C., Iuzzini-Seigel, J., & Green, J. R. (2020). Differential diagnosis of apraxia of speech in children and adults: A scoping review. *Journal of Speech, Language, and Hearing Research, 63*(9), 2952–2994. https://doi.org/10.1044/2020_JSLHR-20-00061

American Speech-Language-Hearing Association. (2007). *Childhood apraxia of speech* [Technical report]. https://www.asha.org/policy/tr2007-00278/

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. https://doi.org/10.1016/j.neuropsychologia.2015.12.010

Ballard, K. J., Djaja, D., Arciuli, J., James, D. G., & van Doorn, J. (2012). Developmental trajectory for production of prosody: Lexical stress contrastivity in children ages 3 to 7 years and in adults. *Journal of Speech, Language, and Hearing Research, 55*(6), 1822–1835. https://doi.org/10.1044/1092-4388(2012/11-0257)

Bartle-Meyer, C. J., Goozée, J. V., & Murdoch, B. E. (2009). Kinematic investigation of lingual movement in words of increasing length in acquired apraxia of speech. *Clinical Linguistics & Phonetics, 23*(2), 93–121. https://doi.org/10.1080/02699200802564284

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. https://doi.org/10.1080/02699050802649654

Benway, N. R., & Preston, J. L. (2020). Differences between school-age children with apraxia of speech and other speech sound disorders on multisyllable repetition. *Perspectives of the ASHA Special Interest Groups, 5*(4), 794–808. https://doi.org/10.1044/2020_PERSP-19-00086

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. https://doi.org/10.1044/2016_JSLHR-S-14-0261

Chen, C., Xie, Y., Wu, C., Gui, P., Wu, H., & Zhang, B. (2020). Effect of melodic intonation therapy on speech apraxia in early stroke patients with non-fluent aphasia. *Journal of Clinical and Experimental Medicine, 19*(05), 502–505. https://doi.org/10.3969/j.issn.1671-4695.2020.05.016

Chen, S.-P. (2011). *Investigation of clinical features of preschool children with childhood apraxia of speech* [Master's thesis, National Kaohsiung Normal University]. https://hdl.handle.net/11296/r796w5

Chenausky, K. V., Brignell, A., Morgan, A., Gagne, D., Norton, A., Tager-Flusberg, H., Schlaug, G., Shield, A., & Green, J. R. (2020). Factor analysis of signs of childhood apraxia of speech. *Journal of Communication Disorders, 87,* 106033. https://doi.org/10.1016/j.jcomdis.2020.106033

Cheung, P. S. P., Ng, A., & Tos, C. K. S. (2006). *Hong Kong Cantonese articulation test*. Language Information Sciences Research Centre and City University of Hong Kong.

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), S1562–S1572. https://doi.org/10.1044/1092-4388(2012/11-0323)

Dabul, B. L. (2000). *Apraxia Battery for Adults–Second Edition*. Pro-Ed.

Deng, B., Lian, Y., Wang, X., Zeng, F., Jiao, B., Wang, Y.-R., Liang, C.-R., Liu, Y.-H., Bu, X.-L., Yao, X.-Q., Zhu, C., Shen, L., Zhou, H.-D., Zhang, T., & Wang, Y.-J. (2014). Identification of a novel mutation in the presenilin 1 gene in a Chinese Alzheimer's disease family. *Neurotoxicity Research, 26*(3), 211–215. https://doi.org/10.1007/s12640-014-9462-3

Diepeveen, S., van Haaften, L., Terband, H., de Swart, B., & Maassen, B. (2019). A standardized protocol for maximum repetition rate assessment in children. *Folia Phoniatrica et Logopaedica, 71*(5–6), 238–250. https://doi.org/10.1159/000500305

Dodd, B., Hua, Z., Crosbie, S., Holm, A., & Ozanne, A. (2006). *Diagnostic Evaluation of Articulation and Phonology (DEAP)*. The Psychological Corporation.

Duffy, J. R. (2013). *Motor speech disorders: Substrates, differential diagnosis, and management* (3rd ed.). Pakhaksa.

Duffy, J. R. (2020). *Motor speech disorders: Substrates, differential diagnosis, and management* (4th ed.). Elsevier Health Sciences.

Galletta, E. E., & Goral, M. (2018). Response time inconsistencies in object and action naming in anomic aphasia. *American Journal of Speech-Language Pathology, 27*(1S), 477–484. https://doi.org/10.1044/2017_AJSLP-16-0168

Gao, H., Zhang, L., Li, Y., Lu, F., & Yu, S. (2012). Clinical observation of using Yiqi Huatan decoction treatment for apraxia of speech after cerebral infarction. *Chinese Journal of Postgraduates of Medicine, 9*(35), 39–40.

Gozzard, H., Baker, E., & McCabe, P. (2006). Children's productions of polysyllables. *ACQuiring Knowledge in Speech, Language and Hearing, 8,* 113–116.

Grigos, M. I., & Case, J. (2018). Changes in movement transitions across a practice period in childhood apraxia of speech. *Clinical Linguistics & Phonetics, 32*(7), 661–687. https://doi.org/10.1080/02699206.2017.1419378

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. https://doi.org/10.1044/2015_JSLHR-S-13-0221

Gu, H. (2011). Effect of remediating speech and oral apraxia on language rehabilitation in aphasia. *Chinese Community Doctors, 13*(12), 157.

Gubiani, M. B., Pagliarin, K. C., McCauley, R. J., & Keske-Soares, M. (2021). Dynamic evaluation of motor speech skill: Adaptation for Brazilian Portuguese. *Journal of Communication Disorders, 93,* 106–114. https://doi.org/10.1016/j.jcomdis.2021.106114

Haley, K. L., Cunningham, K. T., Jacks, A., Richardson, J. D., Harmon, T., & Turkeltaub, P. E. (2021). Repeated word

production is inconsistent in both aphasia and apraxia of speech. *Aphasiology, 35*(4), 518–538. https://doi.org/10.1080/02687038.2020.1727837

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. https://doi.org/10.1080/02699206.2018.1550813

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. https://doi.org/10.1044/1092-4388(2012/11-0318)

Home of Rehabilitation Therapist. (2020). *Assessment of rehabilitation—Assessment techniques of aphasia*. http://med.china.com.cn/content/pid/173975/tid/1026/

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. https://doi.org/10.1044/2016_jslhr-s-15-0184

Jiang, Y., Yang, Y., Chang, E., Xiang, R., Ju, H., Zhang, Y., Zuo, B., & Zhang, Q. (2015). Treatment observation of the combined scalp electroacupuncture and Rosenbek's eight step continuum for treating patients with apraxia of speech after stroke. *Chinese Journal of Rehabilitation Medicine, 30*(4), 366–368.

Jiang, Y., Yang, Y., Xiang, R., Chang, E., Zhang, Y., Zuo, B., & Zhang, Q. (2015). Clinical study of post-stroke speech apraxia treated with scalp electric acupuncture under anatomic orientation and rehabilitation training. *Chinese Acupuncture & Moxibustion, 35*(7), 661–664.

Jiang, Y., Yang, Y., Xiang, R., Tang, S., Ju, H., Huo, Y., Xiao, L., Ji, C., & Chang, E. (2015). Scalp electroacupuncture combined with speech training for apraxia of speech with oral and facial apraxia in stroke patients. *Zhong Guo Kang Fu, 30*(4), 247–249.

Karlsson, F., & Hartelius, L. (2021). On the primary influences of age on articulation and phonation in maximum performance tasks. *Language, 6*(4), 174. https://doi.org/10.3390/languages6040174

Kong, Y., Qiu, W., Hu, X., Wan, G., Zhu, H., & Liu, X. (2004). Characteristics of basal ganglion aphasia after stroke and the effect of pertinenced language training. *Chinese Journal of Clinical Rehabilitation, 8*(25), 5216–5217.

Lee, K. Y. S. (2012). *Hong Kong Cantonese tone identification test*. Department of Otorhinolaryngology, Head and Neck Surgery, The Chinese University of Hong Kong.

Lee, K. Y. S., Lee, L. W., & Cheung, P. S. P. (2009). *Hong Kong Cantonese Receptive Vocabulary Test. Institute of Human Communicative Research*. The Chinese University of Hong Kong.

Li, D., Zhao, L., Jin, H., Zhang, M., Guo, D., Yu, Y., Wu, L., Tang, Y., Li, F., Zhou, A., Han, Y., & Jia, J. (2016). Profiles of cognitive and language impairment of logopenic and nonfluent variant of primary progressive aphasia. *Zhonghua Yi Xue Za Zhi, 96*(29), 2356–2359. https://doi.org/10.3760/cma.j.issn.0376-2491.2016.29.015

Li, H., Liu, Y., Ren, L., Guo, F., & Liu, N. (2003). The incidence and related factors of speech dysfunction of stroke patients. *Chinese Journal of Rehabilitation Medicine, 18*(8), 26–28.

Li, H., Wu, B., Zhang, T., Zhang, Y., & Wang, Z. (2002). Incidence and correlative factors of dysphagia in stroke patients. *Chinese Journal of Physical Medicine and Rehabilitation, 24*(5), 26–28.

Li, Q. (2019). Application effect of music therapy on adults with speech apraxia. *Chinese Journal of Convalescent Medicine, 28*(11), 1134–1136.

Li, S. (2008). *Speech therapy*. Tianjin University Press.

Li, S., & Shirasaka, Y. (1994). *Rehabilitation of hearing and language*. China Rehabilitation Research Centre and Japan International Cooperation Agency China Office.

Liu, J.-C., Liao, S.-F., Chang, M.-Y., & Liu, C.-H. (2019). Efficacy of combined treatment with integral stimulation and the minimal contrast approach for childhood apraxia of speech: A preliminary study. *The Changhua Journal of Medicine, 17*(1), 8–14. https://doi.org/10.6501/CJM.201903_17(1).0002

Lu, Y., Lei, B., Ran, J., & Liu, L. (2020). Observation on clinical efficacy of hand gesture guidance applied to apraxia of speech and buccofacial apraxia. *Chinese Scientific Journal of Hearing and Speech Rehabilitation, 18*(1), 57–61.

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

McNeil, M. R., Ballard, K. J., Duffy, J. R., & Wambaugh, J. L. (2017). Apraxia of speech theory, assessment, differential diagnosis, and treatment: Past, present, and future. In P. van Lieshout, B. Maassen & H. Terband (Eds.), *Speech motor control in normal and disordered speech: Future developments in theory and methodology* (pp. 195–221). ASHA Press.

Melle, N., & Gallego, C. (2012). Differential diagnosis between apraxia and dysarthria based on acoustic analysis. *Spanish Journal of Psychology, 15*(2), 495–504. https://doi.org/10.5209/rev_sjop.2012.v15.n2.38860

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. https://doi.org/10.1044/2014_JSLHR-S-12-0358

Pan, C., & Mao, S. (2000). A case report of rehabilitation training for apraxia of speech. *Modern Rehabilitation, 4*(6), 851.

Pan, X., Kong, Y., Chen, P., Zeng, H., Han, Q., Zhu, H., Fang, S., Deng, X., & Zhang, D. (2006). Characteristics of basal ganglion aphasia after stroke. *Journal of Apoplexy and Nervous Diseases, 23*(3), 341–343.

Pilkington, E., Keidel, J., Kendrick, L. T., Saddy, J. D., Sage, K., & Robson, H. (2017). Sources of phoneme errors in repetition: Perseverative, neologistic, and lesion patterns in jargon aphasia. *Frontiers in Human Neuroscience, 11,* 225. https://doi.org/10.3389/fnhum.2017.00225

Preston, J. L., Benway, N. R., Leece, M. C., & Caballero, N. F. (2021). Concurrent validity between two sound sequencing tasks used to identify childhood apraxia of speech in school-age children. *American Journal of Speech-Language Pathology, 30*(3S), 1580–1588. https://doi.org/10.1044/2020_AJSLP-20-00108

Qin, J. (2008). Effect of character communication on apraxia of speech: A case report. *Chinese Journal of Rehabilitation Theory and Practice, 14*(5), 485–486.

Ren, Y., Guo, G., & Li, Q. (2016). Application of low frequency repetitive transcranial magnetic stimulation on apraxia of speech plus Broca's aphasia after cerebral infarction. *Clinical Medicine, 36*(2), 42–43.

Rex, S., Hansson, K., Strand, E., & McAllister, A. (2021). Performance of Swedish children on a dynamic motor speech assessment. *International Journal of Speech-Language Pathology,* 453–464. https://doi.org/10.1080/17549507.2020.1862300

Reynell, J., & Huntely, M. (1987). *Reynell Developmental Language Scales—Cantonese (Hong Kong version)*. NFER-Nelson.

Robbins, J., & Klee, T. (1987). Clinical assessment of oropharyngeal motor development in young children. *Journal of Speech and Hearing Disorders, 52*(3), 271–277. https://doi.org/10.1044/jshd.5203.271

Rosenbek, J. C., Lemme, M. L., Ahern, M. B., Harris, E. H., & Wertz, R. T. (1973). A treatment for apraxia of speech in adults. *Journal of Speech and Hearing Disorders, 38*(4), 462–472. https://doi.org/10.1044/jshd.3804.462

Rvachew, S., & Matthews, T. (2017). Using the syllable repetition task to reveal underlying speech processes in childhood apraxia of speech: A tutorial. *Canadian Journal of Speech-Language Pathology & Audiology, 41*(1), 106–126.

Shi, C. (1998). Treatment of speech rehabilitation after stroke. *Modern Rehabilitation, 2*(12), 3–5.

Shriberg, L. D., Aram, D. M., & Kwiatkowski, J. (1997). Developmental apraxia of speech. *Journal of Speech, Language, and Hearing Research, 40*(2), 286–312. https://doi.org/10.1044/jslhr.4002.286

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. https://doi.org/10.3109/02699206.2012.655841

Shriberg, L. D., Potter, N. L., & Strand, E. A. (2011). Prevalence and phenotype of childhood apraxia of speech in youth with galactosemia. *Journal of Speech, Language, and Hearing Research, 54*(2), 487–519. https://doi.org/10.1044/1092-4388(2010/10-0068)

Shriberg, L. D., Strand, E. A., Fourakis, M., Jakielski, K. J., Hall, S. D., Karlsson, H. B., Mabie, H. L., McSweeny, 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. https://doi.org/10.1044/2016_jslhr-s-15-0296

Shriberg, L. D., Strand, E. A., Fourakis, M., Jakielski, K. J., Hall, S. D., Karlsson, H. B., Mabie, H. L., McSweeny, 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. https://doi.org/10.1044/2016_jslhr-s-15-0297

Sparks, R., Helm, N., & Albert, M. (1974). Aphasia rehabilitation resulting from melodic intonation therapy. *Cortex, 10*(4), 303–316. https://doi.org/10.1016/S0010-9452(74)80024-9

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. https://doi.org/10.1016/j.jcomdis.2014.06.008

Strand, E. A., & McCauley, R. J. (2019). *Dynamic evaluation of motor speech skills*. Brookes.

Staiger, A., Finger-Berg, W., Aichert, I., & Ziegler, W. (2012). Error variability in apraxia of speech: A matter of controversy. *Journal of Speech, Language, and Hearing Research, 55*(5), S1544–S1561. https://doi.org/10.1044/1092-4388(2012/11-0319)

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. https://doi.org/10.1044/1092-4388(2012/12-0094

Sun, Y., Chen, Z., & Zhu, Y. (2019). Clinical exploration of a combined muscular patch therapy and speech training for patients with speech and oral apraxia after stroke. *Home Medicine, 18*(1), 55.

Tan, Y., Qu, Q., Liu, H., Liu, B., She, Y., & Yi, X. (2017). Curative effective observation of music therapy and Schuell stimulation method on patients with Broca aphasia after stroke. *Guiding Journal of Traditional Chinese Medicine and Pharmacy, 23*(2), 90–93.

The Ethnologue 200. (2021). *What are the top 200 most spoken languages?* SIL International. http://www.ethnologue.com

Thomas, D. C., McCabe, P., Ballard, K. J., & Lincoln, M. (2016). Telehealth delivery of Rapid Syllable Transitions (ReST) treatment for childhood apraxia of speech. *International Journal of Language & Communication Disorders, 51*(6), 654–671. https://doi.org/10.1111/1460-6984.12238

Thoonen, G., Maassen, B., Gabreels, F., & Schreuder, R. (1999). Validity of maximum performance tasks to diagnose motor speech disorders in children. *Clinical Linguistics & Phonetics, 13*(1), 1–23. https://doi.org/10.1080/026992099299211

To, C. K. S., Cheung, P. S. P., & McLeod, S. (2013). A population study of children's acquisition of Hong Kong Cantonese consonants, vowels, and tones. *Journal of Speech, Language, and Hearing Research, 56*(1), 103–122. https://doi.org/10.1044/1092-4388(2012/11-0080)

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. https://doi.org/10.7326/m18-0850

Utianski, R. L., Duffy, J. R., Clark, H. M., Strand, E. A., Boland, S. M., Machulda, M. M., Whitwell, J. L., & Josephs, K. A. (2018). Clinical progression in four cases of primary progressive apraxia of speech. *American Journal of Speech-Language Pathology, 27*(4), 1303–1318. https://doi.org/10.1044/2018_AJSLP-17-0227

Wambaugh, J. L., Bailey, D. J., Mauszycki, S. C., & Bunker, L. D. (2019). Interrater reliability and concurrent validity for the apraxia of speech rating scale 3.0: Application with persons with acquired apraxia of speech and aphasia. *American Journal of Speech-Language Pathology, 28*(2S), 895–904. https://doi.org/10.1044/2018_AJSLP-MSC18-18-0099

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), 15–33.

Wan, G., Qiu, W., Hu, X., Dou, Z., Li, J., Li, K., & Chen, X. (2004). Evaluation of global aphasia with comprehensive speech therapy. *Chinese Journal of Clinical Rehabilitation, 8*(16), 3028–3029.

Wang, J., Wu, D., Cheng, Y., Song, W., Yuan, Y., Zhang, X., Zhang, D., Zhang, T., Wang, Z., Tang, J., & Yin, L. (2019). Effects of transcranial direct current stimulation on apraxia of speech and cortical activation in patients with stroke: A randomized sham-controlled study. *American Journal of Speech-Language Pathology, 28*(4), 1625–1637. https://doi.org/10.1044/2019_ajslp-19-0069

Wang, T.-W., Yang, C.-K., & Chung, L.-C. (2015). Using visual cues teaching a child with childhood apraxia of speech: Preliminary finding. *Dong Hwa Special Education, 54,* 32–41.

Wang, Y., Wang, H.-j., Qian, Y.-y., Li, G., Zhang, P., Zhou, W.-h., Peng, X.-m., & Wu, B.-b. (2018). Childhood apraxia of speech due to FOXP2 gene mutation: A case report and literature review. *Chinese Journal of Evidence Based Pediatrics, 13*(04), 290–294.

Wang, Z., & Li, S. (2013). Speech assessment for Chinese speakers with apraxia of speech. *Chinese Journal of Rehabilitation Theory and Practice, 19*(01), 70–71.

Wei, D., & Li, S. (2000). Using Rosenbek's eight step continuum to treat apraxia of speech: One study case. *Chinese Journal of Rehabilitation Theory and Practice, 6*(2), 24–25.

Wertz, R. T., Lapointe, L. L., & Rosenbek, J. C. (1984). *Apraxia of speech in adults: The disorder and its management*. Allyn & Bacon.

Wertz, R. T., Lapointe, L. L., & Rosenbek, J. C. (1991). *Apraxia of speech in adults: The disorder and its management*. Singular.

Whitehill, T. L. (2010). Studies of Chinese speakers with dysarthria: Informing theoretical models. *Folia Phoniatrica et Logopaedica, 62*(3), 92–96. https://doi.org/10.1159/000287206

Wong, E. C. H. (2017). *Development of a treatment protocol for Cantonese-speaking children with childhood apraxia of speech: Case study* [Master's thesis, The Chinese University of Hong Kong].

Wong, E. C. H., Lee, K. Y. S., & Tong, M. C. F. (2020). The applicability of the clinical features of English childhood apraxia of speech to Cantonese: A modified Delphi survey. *American Journal of Speech-Language Pathology, 29*(2), 652–663. https://doi.org/10.1044/2019_AJSLP-19-00118

Wong, E. C. H., Velleman, S. L., Tong, M. C. F., & Lee, K. Y. S. (2021). Pitch variation in children with childhood apraxia of speech: Preliminary findings. *American Journal of Speech-Language Pathology, 30*(3S), 1511–1524. https://doi.org/10.1044/2021_AJSLP-20-00150

Wong, E. C. H., Wong, M. N., Velleman, S., Tong, M. C. F., & Lee, K. Y. S. (2022). Lexical tone perception and production in Cantonese-speaking children with childhood apraxia of speech: A pilot study. *Clinical Linguistics & Phonetics*. Advance online publication. https://doi.org/10.1080/02699206.2022.2074310

Wong, T. O. K., Wong, E. C. H., & Wong, M. N. (2021). *Assessment and diagnosis of apraxia of speech in Cantonese speakers – A survey on clinical practice.* Paper presented at the 18th Biennial Conference of the International Clinical Phonetics and Linguistics Association, Glasgow, United Kingdom.

Wu, H., Wan, G., Kang, Z., Xie, C., Chen, S., & Qiu, W. (2012). Characteristics of moderate and severe subcortex aphasia and its language training. *Chinese Journal of Rehabilitation Theory and Practice, 18*(10), 954–959.

Yang, H., & Wang, S. (2014). Diagnosis and treatment updates of apraxia of speech. *Chinese Journal of Rehabilitation Medicine, 29*(9), 893–897.

You, L., Wang, Y., Chen, W., Zhang, S., Rao, J., Liu, L., & Shan, C. (2019). The effectiveness of action observation therapy based on mirror neuron theory in Chinese patients with apraxia of speech after stroke. *European Neurology, 81*(5–6), 278–286. https://doi.org/10.1159/000503960

You, L., Wang, Y., Tian, L., Zhu, H., Rao, J., Gao, R., & Liu, L. (2018). The effects of action observation therapy based on mirror neuron theory on apraxia of speech after stroke. *Journal of Clinical Neurology, 31*(6), 429–432.

Zhang, S., Wang, M., Gao, Y., & Zhao, M. (2020). Treatment observation of action observation therapy based on mirror neuron theory on apraxia of speech after stroke. *Our Health, 8*(8), 156.

Zhang, S., Xu, J., & Liu, Y. (2016). *Research on the fMRI study of Chinese speech region of interest based on DIVA model*. Paper presented at the 2016 Chinese Control and Decision Conference (CCDC), Yinchuan, China.

Zhang, X., Yang, X., & Wang, X. (2010). Speech training for global aphasia: 9 case report. *Chinese Journal of Rehabilitation Theory and Practice, 16*(12), 1186–1187.

Ziegler, W., Schölderle, T., Brendel, B., Amsellem, J., & Staiger, A. (2019). Higher–faster–farther: Maximum performance tests in the assessment of neurogenic speech impairment. *Folia Phoniatrica et Logopaedica, 71*(5–6), 261–274. https://doi.org/10.1159/000495784

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. https://doi.org/10.1044/2017_JSLHR-S-16-0106

Appendix A (p. 1 of 3)

Information of the Assessment Methods Identified From the Studies

Method	Diagnosis	Tasks	Scoring methods	Diagnostic criteria and accuracy
Assessment protocol used by E. C. H. Wong (2017)	CAS	• Standardized language assessment (i.e., RDLS and/or HKCRVT) • Standardized articulation test (i.e., HKCAT) • Standardized tone identification test (i.e., CANTIT) • Motor speech assessment, including (a) imitation of polysyllabic words, (b) diadochokinetic tasks, (c) increasing length tasks, (d) oral motor examination, and (e) observation of prosody	Perceptual transcription of the clients' productions and judgment	Perceptual judgment of the presence of all 29 clinical features identified for Cantonese speakers with CAS (E. C. H. Wong et al., 2020) and of ASHA's three consensus features (ASHA, 2007). The features should be observed at least once each in the initial assessment. No diagnostic accuracy was reported for this method.
A list of 105 three-word probes used by Liu et al. (2019)	CAS	• List of 105 three-word probes, which included 292 opportunities for the production of all Chinese consonants in all positions of words	Perceptual transcription and judgment of the clients' productions	Presence of ASHA's three consensus features. No diagnostic accuracy was reported for this method.
Assessment standard suggested by CRRC (S. Li & Shirasaka, 1994)[a]	AOS	• No details of the tasks could be found	No details could be found	No criteria or diagnostic accuracy reported
Assessment standard suggested by CRRC (Wei & Li, 2000)	AOS	• Imitation of [a–u–i] and [i–u–a], 5 times each • Word sequence imitation ([papa], [mama], and [titi] meaning "father," "mother," and "younger brother" respectively), 5 times each • Phrase imitation ([pa] [ta], an onomatopoeia, [ɛi] [ʂoʊ] meaning "wash hands," [ni] [mən] [ta] [tɕʰjoʊ] meaning "you play ball," and [pu] [tʰu] [pʰu] [tʰaʊ] [pʰi] meaning "do not spitgrape skins"), 5 times each	Imitation requires rating of vowel sequences, vowel errors, and groping behaviors. Word sequence imitation requires rating of word sequences, phoneme sequences, and groping behaviors. Phrase imitation requires rating of correct response, phoneme errors, and groping behaviors. However, no details of the scoring methods for these items were reported.	No criteria or diagnostic accuracy reported

(*table continues*)

Appendix A (p. 2 of 3)

Information of the Assessment Methods Identified From the Studies

Method	Diagnosis	Tasks	Scoring methods	Diagnostic criteria and accuracy
Assessment standard suggested by CRRC (S. Li, 2008)[a]	AOS	• Imitation of [a-u-i] and [i-u-a], 5 times each • Word sequence imitation ([papa], [mama], and [titi] meaning "father," "mother," and "younger brother" respectively), 5 times each • Phrase imitation ([pa] [ta] [ɛi] [ʂoʊ] meaning "pa ta wash hands," [ni] [mən] [ta] [tɕʰjoʊ] meaning "you play ball," and [pu] [tʰu] [pʰu] [tʰaʊ] [pʰi] meaning "do not spit grape skins"), 5 times each	All of the items require rating of vowel sequences, vowel errors, and groping behaviors. However, no details of the scoring methods for these items were reported.	No criteria or diagnostic accuracy reported
Assessment standard suggested by CRRC (Z. Wang & Li, 2013)	AOS	• Triphthong imitation ([aui] and [iua]), 5 times each • Single-syllable imitation ([pa], [ta], and [ka]), 1 time each • Single-syllable imitation ([pa], [ta], and [ka]) 5 times per syllable • [pataka] imitation once • [pataka] imitation 5 times • Phrase imitation ([pa] [ta] [ɛi] [ʂoʊ] meaning "pa da wash hands," [ni] [mən] [ta] [tɕʰjoʊ] meaning "you play ball," and [pu] [tʰu] [pʰu] [tʰaʊ] [pʰi] meaning "do not spit grape skins"), 1 time each	Perceptual rating using a 7-point scale (1 for fluent imitation and completing the task; 2 for less-fluent imitation, but completion of the task with a small number of sound distortions and/or metathesis; 3 for slow initiation of imitation, reduced fluency, slowed speech rate, and completion of more than 60% of the task with some phonetic distortions and/or metathesis; 4 for slow initiation for imitation, reduced fluency, slowed speech rate, and completion of less than 40% of the task with a significant number of phonetic distortions and/or metathesis; 5 for completion of less than 20% of the task with a significant number of phonetic and/or syllable distortions and metathesis; 6 for producing only a small number of syllables; and 7 for no verbal production or stereotypical speech only)	No criteria or diagnostic accuracy reported
Psycholinguistic Assessment in Chinese Aphasia (PACA1.0)[b]	AOS	• Nonspeech oral motor examination including 20 items[c] • Counting task (from 1 to 10) • Simultaneous counting with the examiner (from 1 to 10) • Pitch variation task (eight pitches) • Simultaneously varying pitch with the examiner (eight pitches) • Sound imitation tasks ([a], [o], [e], [y], [u], [ʊ], [p], [pʰ], [m], [f], [t], [tʰ], [n], [l], [k], [kʰ], [j], [tɕ], [tɕʰ], and [ɕ]) • Sound sequence imitation tasks ([a]-[o]-[e], [y]-[u]-[ʊ], [p]-[pʰ]-[m]-[f], [t]-[tʰ]-[n]-[l], [k]-[kʰ]-[h], and [tɕ]-[tɕʰ]-[ɕ])	Perceptual rating of each item (0 for incorrect/inappropriate response; 1 for correct/appropriate response)	No criteria or diagnostic accuracy reported

(*table continues*)

Appendix A (p. 3 of 3)

Information of the Assessment Methods Identified From the Studies

Method	Diagnosis	Tasks	Scoring methods	Diagnostic criteria and accuracy
Observation of clinical features proposed by Yang & Wang (2014)	AOS	• Observation of clinical features including (a) slow speech rate; (b) prolongations; (c) increased pauses; (d) intrusive schwa; (e) inappropriate prosody, especially lexical stress and equal stress; (f) sound distortions; (g) distorted substitutions; and (h) consistent articulatory errors	Perceptual judgment of presence or absence of the features	No criteria or diagnostic accuracy reported
An assessment developed based on the ABA-2 (Dabul, 2000) and Wambaugh et al. (2006), used by J. Wang et al. (2019)	AOS	• Imitation of face, tongue, and lip movements (20 scores) • Repetition of 10 monosyllabic and 10 disyllabic words • Counting from 1 to 10 • Repetition of 20 Chinese phonetic alphabet character names that include single and final initials with different places and specific articulations	Perceptual judgement of correctness	No criteria or diagnostic accuracy reported

Note. AOS = apraxia of speech; CANTIT = Cantonese Tone Identification Test (Lee, 2012); CAS = childhood apraxia of speech; CRRC = Chinese Rehabilitation Research Centre; HKCAT = Hong Kong Cantonese Articulation Test (Cheung et al., 2006); HKCRVT = Hong Kong Cantonese Receptive Vocabulary Test (Lee et al., 2009); RSLD = Reynell Developmental Language Scales-Cantonese (Hong Kong Version; Reynell & Huntely, 1987).

[a]Reported from books that were not included in this review but were referenced by the studies identified. [b]Detailed information about the content and the scoring methods for the PACA1.0 module were found in the Home of Rehabilitation Therapist (2020). [c]No information about the items in the nonspeech oral motor examination could be found in the literature.

Appendix B

Wong's 29 Clinical Features That Have Gained Consensus From a Group of Experts for the Differential Diagnosis of CAS in Cantonese Speakers

Area	Items
Nonspeech motor behaviors	Oral apraxia
	Groping behaviours
	Poor imitation of oral movement
	Discrepancies between volitional and automatic oral gestures
	Problem in alternating oro-motor movement
Speech motor behaviors	Poor speech intelligibility
	More unintelligible speech in connected speech
	Limited phonetic inventory
	Late acquisition of consonants and vowel acquisition
	Display vowel errors including distortions and deletion
	Difficulty in diphthong production
	Increase errors with increasing syllable complexity
	Increase difficulty with increasing length of utterance
	Use of relatively simple syllable structures predominately
	Incomplete phonotactic inventory
	Syllable addition or deletion
	Inconsistent errors
	Groping
	Weak syllable transition/difficulties in coarticulation
	Slow progress in learning new consonants or vowels
	Syllable sequencing errors
	Poor DDK performance (abnormal alternating motion rate [AMR] and sequential motion rate [SMR])
	Decrease in accuracy with increasing number of syllables in syllables repetition tasks
	Poor speech sound generalization
Prosodic characteristics	Difficulties in speech rate control
	Syllable segregation
Cognitive characteristics	Better cognitive skills than expressive language skills
Visual and hearing characteristics	Normal
Others	Poor treatment progress

Note. CAS = childhood apraxia of speech; DDK = diadochokinesis.