Journal of Speech Sciences 8(1): 01-35. 2019.
Available at: http://revistas.iel.unicamp.br/joss

VALIDITY AND RELIABILITY OF SPEECH ARTICULATION ASSESSMENT TOOLS FOR CHILDREN AND ADOLESCENTS: A SYSTEMATIC REVIEW

UBERTI, Letícia Bitencourt[1]*
PORTALETE, Caroline Rodrigues[1]
PAGLIARIN, Karina Carlesso[1]
KESKE-SOARES, Márcia[1]

[1] Department of Speech-Language and Hearing Pathology, Federal University of Santa Maria, Santa Maria, Rio Grande do Sul, Brazil.

Abstract: *While some children with speech sound disorders successfully overcome speech errors, others continue to present with omissions, substitutions or distortions in adolescence and even adulthood. These are known as residual speech errors, and can have a significant influence on speech perception, production and phonological representation. Other causes of speech errors include articulation disorders in clients with cleft lip or palate, alterations in the lingual frenulum, genetic syndromes and neurological damage. Articulation tests are important tools for speech-language pathologists (SLPs) assessing clients with these issues. Purpose: [i]To identify and describe available instruments to evaluate speech articulation, [ii]and to verify which of these instruments have evidence to support their validity and reliability. Method: This study was conducted according to PRISMA guidelines. The literature search was performed by two SLPs via major online scientific databases including PubMed and Scopus using keywords related to assessment, articulation, assessment methods, and speech. Search results were limited to studies involving children and adolescents, published in English, Portuguese, or Spanish. Results: 51 articles were selected for full-text review. Seventy eight percent (n=39) of these studies assessed children aged 6:0 to 9:9. A total of 28% (n=14) used instruments to evaluate language and communication skills, including measures of phonological processing, working memory, phonological awareness and automatization of correct speech production. Only 11% (n=6) of studies provided evidence of validity and normative data for assessment instruments. Conclusions: Most of the studies evaluated participants using instruments developed for non-commercial research use. Researchers also combined measures of speech articulation and instruments that evaluated speech perception and other aspects of communication. Despite the absence of psychometric studies, all instruments had adequate methodological quality. Future studies should continue to collect evidence of the psychometric properties of assessment instruments in order to improve reliability.*

Keywords: Speech-Language Pathology; Speech. Articulation; Tests; Systematic Review.

*Corresponding author: Letícia Bitencourt Uberti
Department of Speech-Language and Hearing Pathology, Federal University of Santa Maria, Building 26E – Room 212. Roraima Avenue, No. 1000 – Building 26 E – Room 212. University Campus Camobi District. Santa Maria, Rio Grande do Sul, Postal Code 97105-900, Brazil.
E-mail: leticiab_uberti@hotmail.com

1 Introduction

Speech is a complex process that relies on several physiological and motor articulatory behaviors, as well as sensory perception and the development of grammatical and lexical skills (Van Severen et al., 2013). The development of speech articulation starts in early childhood and continues in the following years, reaching a peak in the preschool stage. Substitutions, omissions, and distortions of speech sounds often occur in the developmental process due to structural, motor and/or sensory conditions, though these difficulties are usually overcome over time (Goldman & Fristoe, 2000; Lee, Whitehill, Ciocca & Samman, 2002; Newmeyer, et al., 2007; Stoel-Gammon & Sosa, 2009 Ferrante, Borsel & Pereira, 2009; Rabelo, et al., 2011). However, some children over the age of four/five years continue to present with some of these speech errors, which are no longer expected to occur at this age (Ceron, 2015). Adolescents and adults may also exhibit substitutions or distortions of liquids, fricatives, and affricates, which in this case are known as residual speech errors (Preston & Koenig, 2011; Veríssimo, Van Borsel & Pereira, 2012; Preston, et al., 2018).

Difficulties in the perception, production, and/or phonological representation of speech sounds and segments which compromise intelligibility are referred to as speech sound disorders (SSD) (ASHA, 2013). Clients with a cleft lip or palate, for instance, usually have compensatory articulation disorders characterized by the production of glottal stops; pharyngeal, velar, or nasal fricatives; and pharyngeal or mid-dorsum palatal stops (Clements, 1985). Alterations in the lingual frenulum may cause phonetic disturbances characterized by distortions or alterations in articulatory speed or precision (Martinelli & Marchesan, 2015; Cuestas, Demarchi, Corválan, Razetti & Boccio, 2014). Genetic syndromes associated with bone and muscle alterations, such as Down's Syndrome, can cause stomatognathic disorders which result in articulatory difficulties (Alves, Lima, Lima & Delgado, 2016). Lastly, neurological damage can cause speech and language disorders such as dysarthria or apraxia, which affect the coordination or sequencing of speech movements (Kent & Read, 2002).

Speech production is usually assessed using figure naming, word elicitation, word repetition, narrative, or conversation tasks (Morrison & Shriberg, 1992; Klintö, Salameh, Svensson & Lohmander, 2011; Masterson, Bernhardt & Hofheinz, 2005). When evaluating articulatory disorders, speech-language pathologists (SLPs) must rely on articulation tests to evaluate speech production and verify whether and how the patient can produce specific speech sounds. These tests tend to be relatively quick and easy to administer and score (Paul, 1995). A thorough assessment of speech production is crucial to ensure the adequate rehabilitation of SSDs and reduce the negative impact of speech errors (McLeod & Verdon, 2014; Abou-Elsaad, Baz & El-Banna, 2009).

The importance of assessment instruments that provide accurate scores and parameters for speech sound production is well known. Yet SLPs often report that the instruments available to measure speech production, especially articulation, lack psychometric evidence (of validity and reliability) and have not been adequately adapted to the cultural and linguistic characteristics of non-English speaking populations (Caesar & Kohler, 2007; Jordaan, 2008; Roseberry-McKibbin, Brice & O'Hanlon, 2005; Skahan, Watson & Lof, 2016; Williams & McLeod, 2012).

Several important considerations must be made when deciding on the type of task to be used to evaluate speech in both clinical and research settings (Masterson, Bernhardt & Hofheinz, 2005). Thus, to help researchers and clinicians select the best evaluation methods for speech articulation, the present study was guided by the following questions: What methods are

used to assess speech articulation in children and adolescents? What instruments are available for this purpose? Do these instruments have adequate validity and reliability?

These issues will be discussed through a systematic literature review of studies involving articulation assessments in children and adolescents, to identify and describe the instruments and methods available to evaluate speech articulation. This study will also investigate the psychometric properties (validity and reliability) of the articulation tests used in these populations.

2 Methods

The search procedures and eligibility criteria were constructed according to PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines (Moher, Liberati, Tetzlaff, Altman & Group, 2009). The method used will be described below in the following sections: literature search, screening, eligibility, and article inclusion.

2.1 Literature Search

The first methodological stage aimed to identify studies that evaluated speech articulation in children and adolescents. To increase reliability, two researchers independently searched the PubMed and Scopus databases in June and July 2019.

The search strategy was structured around four constructs: (1) assessment; (2) articulation; (3) assessment methods; (4) speech. Each construct was represented by a series of keywords combined using the Boolean operator OR. Search terms were selected based on the frequency with which they were cited in articles about the topic of study.

The assessment construct was represented by the following set of keywords: "Evaluation" OR "Assessment" OR "Protocol" OR "Instrument" OR "Battery" OR "Task" OR "Screening" OR "Exam". The articulation construct was captured by the term "Articul*". The assessment method was represented by the terms "Repetition" OR "Imitation". Lastly, the construct of speech was captured by the term "Speech*". The search was limited to titles and abstracts in both databases.

The association between all four constructs was then investigated in each database by combining all four searches with the Boolean operator AND. Search results were restricted to publications in English, Spanish, and Portuguese. The populations studied were limited to "Child" (6 to 12 years) and "Adolescent" (13 to 18 years). No restrictions were placed on the date of publication or full-text availability of search results, as this data will be analyzed.

2.2 Screening

The second stage aimed to carry out a preliminary screening of studies retrieved in the original search to identify potentially eligible articles for full-text analysis. A spreadsheet was created to compile information on all articles retrieved from the two databases. Duplicate articles were removed. Search results were then refined based on information from the titles, abstracts, and keywords of each article.

This process involved the application of the following exclusion criteria: (1) articles in languages other than English, Spanish or Portuguese; (2) participants outside the 6- to 18-year age group; (3) no description of a method used to evaluate speech articulation (e.g. word or sentence repetition used to evaluate verbal fluency, phonological memory or working memory).

Articles were independently screened by two researchers (first and second author). Articles whose titles and abstracts were judged to be relevant according to previously described

criteria were then retrieved in full. Agreement rates between the two researchers were evaluated and any discrepancies were resolved by a third rater (fourth author), who would determine whether the article would continue to the next stage.

2.3 Elegibility

The third stage involved the assessment of full-text articles for potential inclusion in the review. The following inclusion criteria were applied: (1) experimental studies; (2) at least one group of participants consisting of children and/or adolescents aged 6 to 18 years; (3) use of assessment methods or instruments specifically aimed at evaluating speech articulation.

Studies that used verbal assessments to examine abilities other than speech articulation (e.g. fluency, morphosyntactic skills, phonological working memory, perception, vocabulary) were not included in this review.

2.4 Data Extraction

After study selection, relevant data were extracted from each article and inserted into a spreadsheet for analysis. The following information was collected from each study: (1) participant characteristics; (2) assessment methods; (3) assessment instruments; (4) psychometric properties.

Studies that mentioned the use of specific instruments to evaluate speech articulation were further classified according to the availability of evidence to support their psychometric properties. The following categories were used for this purpose: (a) evidence of validity (content, criterion, construct); (b) evidence of reliability; (c) standardization; (d) availability of normative data; (e) partial evidence of validity and reliability.

3 Results

The initial search retrieved a total of 226 abstracts (PubMed = 49; Scopus = 177). The independent search procedures used by each of the two researchers resulted in a similarity of 100% between findings. After the removal of duplicates (n=39), a total of 187 unique publications were included in the screening stage.

A total of 126 studies were excluded based on the aforementioned criteria, leaving 60 studies for full-text analysis. The researchers agreed in the assessments of 49 articles (82%) but disagreed with regards to the remaining 11 (18%). These articles were therefore screened by a third rater, who recommended the inclusion of two articles and the exclusion of the other nine, since eight did not evaluate speech articulation and one included participants older than 18 years. At the end of the screening stage, 51 articles were selected for full-text review.

All 51 articles were read by the researchers during the eligibility stage. After the exclusion of one article which did not report on an experimental study and therefore did not fulfill inclusion criteria, 50 articles were included in the final review. The first author read the full text of all selected articles. The article selection process is illustrated in the PRISMA flow diagram in Figure 1.

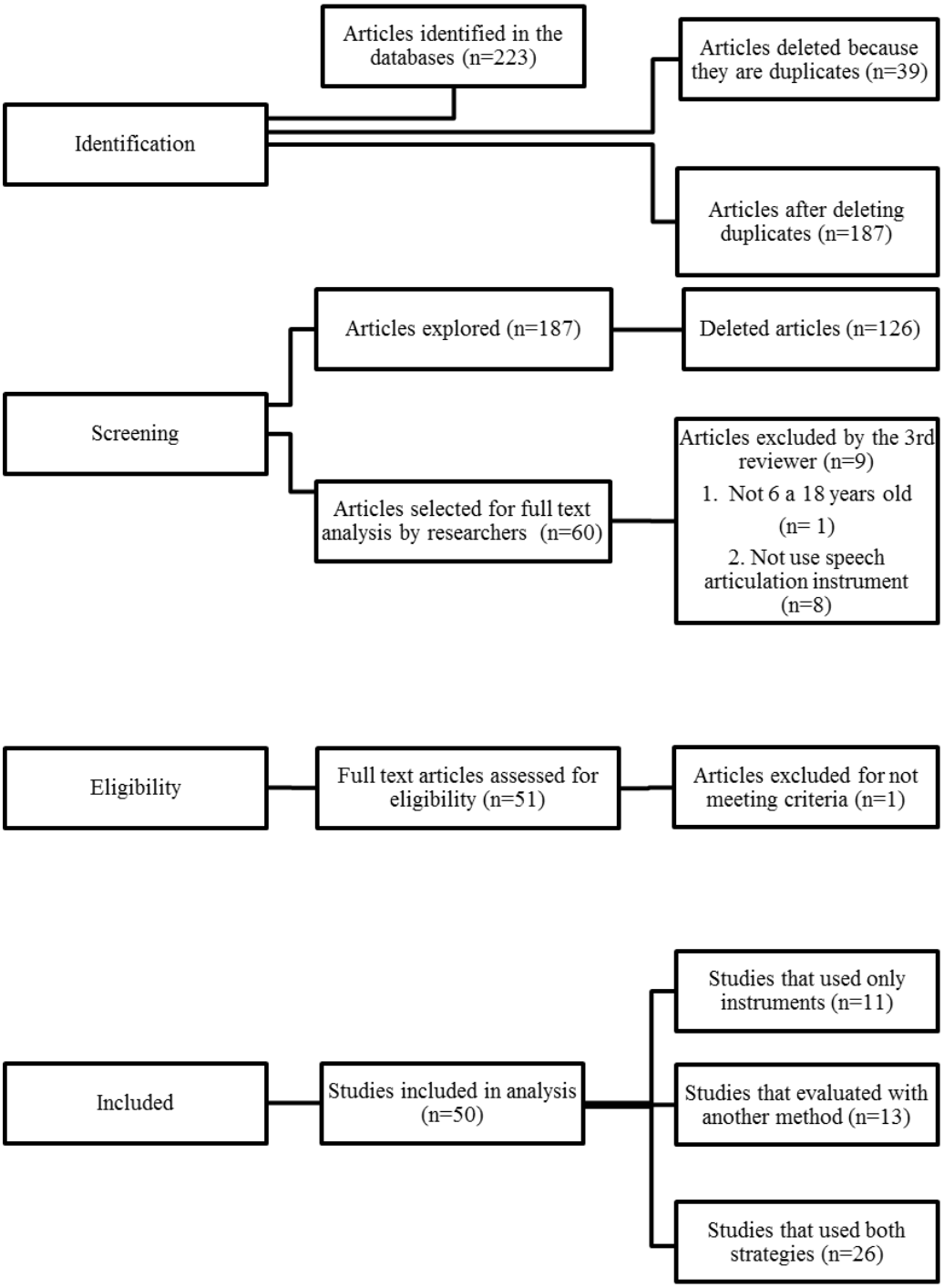

Figure 1: Study selection flow diagram

The fifty articles included in this systematic review are shown in Table 1, which includes general information about each study, including their article identification numbers, authors, year of publication and presence of SLPs in the research team. The table also describes the techniques, methods and assessment instruments used in each study, the target population, and their main conclusions.

Most of the studies analyzed in this review (n=41) were conducted and authored by at least one SLP who worked as a researcher in a speech language-pathology department and

studied communication or speech disorders. In nine studies, SLPs were responsible for evaluating research participants, but were not included as authors in the published articles. In some studies, research teams also included physicians, psychologists, occupational therapists, and/or linguists.

The oldest study included in the review dates from 1969 (Wright, Shelton & Arndt, 1969), and used the Templin-Darley Screening Test (Templin & Darley, 1960). Seventy eight percent (n=39) of the articles, however, were published in the 21st century, with 41% (n=16) of these published in the past five years.

Table 1. General information, and speech and articulation assessment methods and instruments of each study included in this systematic review

Article ID	Authors / Year	SLP researcher	Speech and Articulation Assessment		Main conclusions
			Methods	Instruments	
A1	Piazzalunga, Previtali, Pozzoli, Scarponi & Schindler 2019	Yes	Nonword repetition	Self-authored assessment instrument	Results indicated excellent test-retest, intra-rater, inter-rater reliability agreement, high internal consistency and good concurrent validity. Normative data suggested that successful performance on NWR increases with age and declines with an increase in stimuli length.
A2	ÖAlves, Ode & Strömbergsson 2019	Yes	Isolated words imitation Picture naming Connected speech elicited through task of concurrent commenting during the playback of a silent short film Free conversation	STI-CH (Lagerberg et al.,2015) LINUS (Blumenthal & Lundeborg, Hammarström, 2014) Episode of *Pingu* (*Pingu as a Chef* or *Pingu Helps around the House)* Question "What would you do if you won one million Swedish crowns?"	The model called "concurrent commenting" was a useful method for eliciting connected speech in the context of unintelligible speech. The Percentage of Intelligible and Correct Syllables was a valid measure of articulation proficiency/speech adequacy in the evaluation of speech containing unintelligible portions to which standard measures of speech accuracy are not applicable.
A3	Saletta, Goffman, Ward & Oleson 2018	Yes	Picture naming	BBToP (Bankson & Bernthal, 1990)	Children with SLI showed a distinct pattern of speech motor learning in sentence production tasks of high load condition that requires active retrieval, as they demonstrated great variability compared with their peers with TD. In the articulatory kinematic domain, children with SLI became closer to the performance of their peers with TD in imitation conditions of lower linguistic load.
A4	Lenoci & Ricci 2018	Yes	Ultrasound images collected from nonword repetition	Self-authored assessment instrument	The stutters group presented a larger coarticulation and larger variability in comparison with the control group. Speech motor control system of children who stutter is less mature in preparing and executing the speech gestures required for fluent speech.

A5	Vuolo & Goffman 2018	Yes	Picture naming 25 words named 3 times Agent imitation + action phrases (imitation task) and phrases retrieved (retrieval task) Morphosyntactic structures elicited through presentation of 44 photographs and corresponding verbal prompts	BBToP (Bankson & Bernthal, 1990) DEAP WI (Dodd et al., 2006) Puppets performing the actions: baby pops, mommy beeps, mommy bumps, puppy mops, baby puffs, and puppy wipes SPELT-P 2; SPELT-3; (Dawson et al., 2005; 2003)	All children showed increased articulatory variability in the retrieval task compared with the imitation task, but only children with SLI showed disproportionate increased articulatory variability compared with their peers with TD in the retrieval task. Language and articulatory levels interact during speech production: higher-level language processes affect lower-level speech motor control processes. In children with SLI, deficits in articulatory control are mediated by language rather than speech processes.
A6	Sjögreen, Mårtensson & Ekström 2018	No	Conversation Single word repetition Sentence repetition	ORIS (Holmberg & Bergstrom 2008) SVANTE (Lohmander et al., 2017)	The deviant production of bilabial consonants, interdental articulation and hypernasal speech are characteristics of flaccid dysarthria in congenital and childhood DM1. Also, dysarthria is more frequent and more severe in congenital DM1 compared with childhood DM1.
A7	Rvachew & Matthews 2017	Yes	Syllable repetition Production of consonants in word-initial and final position Repetition of "pattycake" Single and sequenced nonspeech movements Free speech sample 25 words named 3 times Maximum phonation duration of [a] and [mama], maximum fricative duration of [f], [s], [z], maximum monosyllabic [pa], [ta], [ka] and	SRT (Lohmeier & Shriberg, 2011) DEAP Artic and DEAP WI (Dodd et al., 2006) MPT (Thoonen et al., 1996)	This study concluded that qualitative interpretation of children's performance as described here provides useful diagnostic information, especially when used together the DEAP and the SRT, because they provide assessment data that is informative with regard to children's phonological and motor planning abilities and also determine deficits in the underlying speech processes. Motor planning abilities were revealed by the oral motor exam, maximum performance test, and prosody errors in connected speech.

			trisyllabic [pataka] repetition rate		
A8	Lohmander, Lundeborg & Persson 2017	Yes	Words production Sentence repetition Conversation/Connected speech	SVANTE (Lohmander, et al., 2017)	The normative values were useful for a wide range of patients. The test was easy to perform and can give information on articulation, consonant inventory and nasality as well as an overall rating of the velopharyngeal function and intelligibility.
A9	Nip 2017	Yes	Verbal response to stimulus pictures Single word production Sentence production Oral DDK task Speaking tasks	CELF-4 or CELF-P2 (Semel, Wiig & Secord, 2003;2004) TOCS+ (Hodge & Daniels, 2007) SIT (Yorkson et al., 2007) Diadochokinetic "buh" Syllable repetition "uhba" Sentence repetition "buy Bobby a puppy"	Speakers with CP generally showed a lower degree of spatial and temporal interarticulator coordination. In addition, speech movements embedded within a simple sentence were generally produced with greater spatiotemporal coordination as compared to syllables or DDK.
A10	Vuolo & Goffman 2016	Yes	Picture naming Production of consonants in word-initial and final position Repetition of "pattycake" Single and sequenced nonspeech movements 25 words named 3 times Agent imitation + action phrases (imitation task) and phrases retrieved (retrieval task)	BBToP CI (Bankson & Bernthal, 1990) DEAP Artic and DEAP WI (Dodd et al., 2006) Puppets performing the actions: baby pops, mommy beeps, mommy bumps, puppy mops, baby puffs, and puppy wipes	Short-term speech production practice in rote imitation disrupts articulatory control in children with and without CAS, and spatiotemporal variability is disrupted in children with CAS as they have higher degrees of articulatory variability. It also showed that speech motor control processes could not fully account for high levels of segmental variability.
A11	Tresoldi et al. 2015	Yes	Repetition	Schindler's Repetition Test (Tresoldi et al, 2015)	Schindler's test could be considered a reliable, valid and easily applicable instrument to assess speech abilities in Italian speaking children aged from 3 to 6 years old. This

					test is recommended in clinical practice, as a screening test or as a first clinical assessment instrument.
A12	Murray, McCabe, Heard & Ballard 2015	Yes	25 words named 3 times 50-item picture naming 50 utterances recorded over 10 minutes (McLeod, 1997) Oral motor assessment including DDK Verbal response to stimulus pictures	DEAP WI (Dodd et al., 2006) Single-Word Test of Polysyllables (Gozzard, Baker & McCabe, 2004; 2008) Connected speech (McLeod, 1997) Oral and Speech Motor Control Protocol (Robbins & Klee, 1987) CELF-4 or CELF-P2 (Semel, Wiig & Secord, 2006; Wiig, Secord, & Semel, 2006)	CAS and non-CAS in verbal 4- to 12-year-olds in this sample could be discriminated with 91% accuracy based on four measures, following completion of a thorough OMA including DDK accuracy. These results suggested that the polysyllabic production accuracy and an OMA including DDK might be sufficient to reliably identify CAS and rule out structural abnormality or dysarthria.
A13	Icht, & Ben-David 2015	Yes	Non-word and real word DDK task performed twice	"Pataka" and "Bodeket" repetition (Icht & Ben-David, 2014).	Results showed that real word repetition was significantly faster than non-word and it also was documented a developmental pattern, as performance rates were faster for 11 year olds than for 9 or 10 year olds.
A14	Lin & Demuth 2015	No	Ultrasound images collected from words repetition	Four high-frequency, imageable, monosyllabic /CVl/ and /lVC/ words: two onset /l/s and two coda /l/s; and two controls /CVC/ words with /w/ onsets	Australian English speaking children's norms were similar to other English-speaking children. The disparity in the production/perception of children's singleton onset /l/s was linked to both physiological and phonological development.
A15	Lagerberg et al. 2015	Yes	Word repetition Picture-naming	STI-CH (self-authored assessment instrument) SVANTE (Lohmander, et al., 2005)	Results indicates that STI-CH could be an option for the assessment of intelligibility in Swedish-speaking children, and that the principles used in the development of the test could be of use in the design of intelligibility tests in languages other than Swedish.

A16	Turner et al. 2015	Only for assessment	Word and non-word repetition Maximum vowel prolongation, maximum repetition rate of monosyllables and trisyllables Motor speech planning and programming taks Oral motor tasks Receptive and expressive language tasks	Nonword Memory Test and Multisyllabic Word Repetition Task MPT (Thoonen et al., 1996) ABA-2 (Dabul, 2000) Frenchay Dysarthria Assessment-2 CELF-4 (Semel, Wiig & Secord, 2006) PPVT-4 (Dunn & Dunn, 2007) Expressive Vocabulary Test-2 (Williams, 2007) Test for Reception of Grammar - 2(Bishop, 2003)	The study demonstrated that distinctive features of dysarthria and dyspraxia are found in individuals with GRIN2A mutations, often in the setting of epilepsy-aphasia syndromes; dysarthria has not been previously recognized in these disorders.
A17	Spencer & Weber-Fox 2014	Yes	Morphosyntactic structures elicited through presentation of 44 photographs and corresponding verbal prompts Picture naming Repetition nonsense words of increasing length	SPELT-3 (Dawson, Stout, & Eyer, 2003) BBToP CI (Bankson & Bernthal, 1990) NRT (Dollaghan & Campbell, 1998)	Results suggested a possible role for phonological and speech production abilities, including consonant production in picture naming, and auditory perception, phonological working memory, speech planning and execution for novel phonological sequences (NRT), as indices of eventual stuttering persistence or recovery. Differences in linguistic and phonological proficiencies among children who stutter were observed.

A18	Krishnan et al. 2013	Yes	Imitation of simultaneous and sequential orofacial movements	Non-Word Repetition (Wagner, Torgesen & Rashotte, 1999)	Nonlinguistic oromotor skills contributed to children's NWR ability, suggesting that important aspects of language learning and consequent language deficits might be rooted in the ability to perform complex sensorimotor transformations.
A19	Wertzner et al. 2013	Yes	Phonology test Retelling of a history DDK	ABFW (Andrade, Befi-Lopes, Fernandes & Wertzner, 2004) Book "*Esconde-Esconde*" (Furnari, 1993) Motor Speech Profile® software	Although no correlation between DDK measurement and age have been observed, there was no difference between children with and without SSD, indicating that children aged between 5 and 7:11 years are still improving their motor speech development and therefore it was not possible to differentiate them during a DDK evaluation.
A20	Wren et al. 2013	Yes	Confrontation naming task Three picture description activities Non-word repetition	WOLD (Rust, 1996). Adaptation of CNRep (Gathercole & Baddeley, 1996)	This study provided population level data on a range of measures of speech production across single word, connected speech and non-word repetition sample types in 8-year-old children.
A21	Hack et al. 2012	Yes	Single word production targeting 39 consonants and clusters across word positions Elicitation of the 19 Cantonese consonants across word positions	GFTA-2 (Goldman & Fristoe, 2000) CSPT (So, 1993)	Bilingual children had significantly lower standard scores than monolingual children in GFTA-2, that is, they had near perfect phonology in Chinese, but a range of speech sound differences in English. These results suggests a need for formal phonological assessment in both languages of bilingual children.
A22	Castro & Wertzner 2012	Yes	Stimulability test	ABFW (Andrade, Befi-Lopes, Fernandes & Wertzner, 2004)	Stimulability test applied was effective on the identification of stimulable children among those presenting absent sounds. Children with SSD with absent sounds were more severe since their PCC-R were lower. Results suggested that the most part of the children with absent sounds are stimulable but might not be stimulable depending on the syllable structure or the complex articulatory gestures involved at the production.

A23	Peter 2012	Yes	Receptive vocabulary tasks using single words Sentence imitation Single word production targeting 39 consonants and clusters across word positions Same word productions from the GFTA-2	PPVT-3 (Dunn & Dunn, 1997) Recalling Sentences subtest from the CELF-P or CELF-4 (Wiig, Secord & Semel, 1992; Semel, Wiig & Secord, 2006) GFTA-2 (Goldman & Fristoe, 2000) KLPA-2 (Khan & Lewis 2002)	Results showed that syllable durations were significantly correlated with vowel durations and hand clap intervals, sentence imitation was correlated with all three timed movement measures, and clustering on syllable repetition durations produced three clusters that also differed in sentence imitation scores.
A24	Nip 2012	Yes	Verbal response to stimulus pictures Oral DDK task Speaking tasks	CELF-4 (Semel, Wiig &Secord, 2003) Diadochokinetic "buh" Syllable repetition "uhba" Sentence repetition "buy Bobby a puppy"	Despite increased movement speed, speaking rate was slower for individuals with CP and is affected by task demands. One potential reason for this decrease in speaking rate might be that individuals with CP have reduced force control.
A25	Moss & Grigos 2012	Yes	20 items for total motor control 42 items for oromotor control 16 items for sequencing and two complementary areas Connected speech and language Single word production targeting 39 consonants and clusters across word positions Receptive and expressive language tasks Productions of labial sounds in one-, two-, and three-syllable words to complete a cloze sentence or respond to a "who" question cued by the picture probe	VMPAC (Hayden & Square, 1999) GFTA-2 (Goldman & Fristoe, 2000) TELD-3 (Hresko, Reid & Hammill, 1999) Short story with three characters: "Pop", "Puppet" and "Puppypop," represented by two-dimensional picture probes (Smith et al., 2010)	Children with CAS were differentiated by higher values of lip opening, that is, they had more difficulty in generating stable movement plans. These results provided evidence that aspects of coordination might differentiate children with CAS from those with articulation/phonological impairments, because variability in individual articulator movements and in the overall consistency of the movement goal differed in these groups.

A26	Newton 2012	No	Sentences with two-word environments for potential assimilation or elision	Sentence Repetition Test (Newton, 1999) EPG	Conversation was an appropriate context for a children's phonology assessment. All of the children produced instances of word boundary behaviors reported in adult speech, as well as some which are considered to be atypical.
A27	Pizolato Fernandes & Gavião 2011	Yes	Five pictures for sequential naming, containing all of the phonemes of the Brazilian Portuguese language in several positions	AFC (Yavas, 1991)	There was no association between TMD and speech disorders. Occlusal alterations might be factors of influence, allowing distortions and frontal lisp in phonemes /s/ and /z/, and inadequate tongue position in /t/, /d/, /n/, and /l/.
A28	Ho & Wilmut 2010	Yes	Non-verbal movements (mouth opening and closing) 35 single syllable words Plosive CV nonsense syllable sequences and nonsense words Sentence repetition children	Self-authored articulatory speed performance assessment	Children with DCD who did not display overt speech and language problems, tended towards an atypical pattern of lip movement during complex speech tasks. Results suggested that oro-motor control in children with DCD is an area worthy of examination in understanding the full motor phenotype of DCD.
A29	Sasisekaran et al. 2010	Yes	Verbal response to stimulus pictures Articulatory structures and Movements evaluation Four non-words repetition: "mabfaisheib", "mabsheitaidoib", "mabspoukweefleib", and "mabskrisploistroob"	CELF-3 (Semel, Wiig & Secord, 1995) OSMSE-R (St Louis & Ruscello, 1987) NRT (Dollaghan & Campbell, 1998)	This study demonstrated that the short-term changes were associated with underlying changes in the neural circuitry that controls speech movements, reflecting changes in synaptic densities and weights. The authors concluded that changes in motor commands to muscles occur in a systematic way to reach an optimal coordinative pattern in both children and adults.
A30	Moore, Tompkins & Dollaghan 2010	Yes	16 non-words, ranging in length from 1 to 4 syllables Consonant sounds from the Early-8 and Middle-8 sound classes Consonant sounds from Late-8 sound classes	NRT (Dollaghan & Campbell, 1998) L8NRT (Dollaghan & Campbell, 2003)	Results showed that inter- and intra-rater reliability of the L8NRT were high; split-half reliability was significant and comparable to that of the NRT; average L8NRT scores were significantly lower than NRT scores overall, and at all nonword lengths but the shortest. The study concluded that the psychometric properties of the L8NRT were acceptable, but an error analysis suggested ways in which the task might be improved to better control perceptual demands and articulatory feature overlap.

A31	Castro & Wertzner 2009	Yes	Imitation of 63 syllables with the /l/, /ɾ/, and /ʎ/, as well as seven oral vowels (with the possibility of providing sensory cues)	Self-authored stimulability test for liquid sounds in Brazilian Portuguese.	Results showed that the use of sensory cues seemed to facilitate sound stimulability, making it possible for the children with PD to accurately produce the sounds modeled. Sensory cues were required, especially for the sounds /ɾ/ and /ʎ/, and seemed to be effective production facilitators.
A32	Baylis, Munson & Moller 2008	Yes	Conversation Verbal and nonverbal intelligence skills tasks Creation of 79 non-words using the phonemes /h, w, p, b, m, n, g, k, t, d, f/ combined with vowels to create sequences of CVC and CVCVC	Connected speech sample K-BIT matrices and vocabulary sections (Kaufman, 1990) Phonological Learning Ability: The Implicit Priming Task. (Fisher et al., 2001) Novel phonetic inventory task of the Phonetic Inventory (Single-Word Phonetic Accuracy)	Children with VCFS had poorer articulation skills compared with children with cleft palate or VPD. Articulation difficulties seen in the children with VCFS did not appear to be associated with speech perception skills or the ability to learn new phonological representations.
A33	Archibald & Alloway 2008	No	Completion of a sentence about a picture that is designed to elicit particular grammatical structures; Repetition of auditorally presented sentences of increasing grammatical complexity Single word production targeting 39 consonants and clusters across word positions 40 non-word repetitions	Word Structure and Recalling Sentences subtests of CELF-UK3 (Semel et al., 1995) GFTA-2 (Goldman & Fristoe, 2000) CNRep (Gathercole and Baddeley 1996)	Both SLI and DCD groups had a slower articulation rate than that of TD children of the same age. Also, the DCD group was less impaired on this measure than the SLI group, which was a more severely impaired group overall as reflected by their receptive language deficits and lower raw scores on the non-word repetition test. This study suggested that even in children with SLI and no observable articulation/phonological deficits, subtle motor speech deficits might still be present.

			Listening to a story read aloud, and then retelling it Repetition of the following words individually, as fast as possible, five times: "elephant", "newspaper", "telephone", "banana", and "bicycle"	Story retelling Articulation rate	
A34	Newmeyer et al. 2007	No	Oral Movement Simple Phonemic and Syllabic Level Complex Phonemic and Syllabic Level Spontaneous Length and Complexity	KSPT (Kaufman, 1995)	The presence of abnormal oral-motor imitation skills was associated with sub-average fine motor performance. In children with SSD, the presence of abnormal oral-motor imitation skills was found to be correlated with the presence of fine motor deficits. These results suggested a common neurophysiologic link between the speech planning and fine motor movements.
A35	Walker & Archibald 2006	Yes	Articulation Repeating sentences Responding to directives and answering questions Describing actions Sequencing events Receptive vocabulary tasks using single words Sounds-in-Words subtest Stimulability subtest Sounds-in-Sentences subtest Spontaneous speech	Fluharty Preschool Speech and Language Screening Test (Fluharty 1978) PPVT (Dunn & Dunn 1981) GFTA (Goldman & Fristoe, 1972) Story-telling task, a description of "Goldilocks and the Three Bears", in response to the presentation of a picture book of the story	This study provided unique information about the development of articulation rate in preschool children and of certain variables that interact with it. Diadocokinesia and variability did not show significant difference according to age. Also, speech and repetition tasks were significantly faster than imitation task. Results suggested certain trends in the development of speaking rate and underscore both its complexity and the range of individual differences.

			Imitation task	Repetition of the utterances "I took a toy" and "I saw a cow and gave him some hay"	
			Automatic speech task	Recitation of the nursery rhyme "Humpty Dumpty"	
			Repetition task	The following directions were given: "What is your name? Right! Now I want you to say 'My name is [inserting child's name]' five times".	
A36	Munson, Edwards & Beckman 2005	Yes	Non-word Repetition Test Sounds-in-Words subtest Stimulability subtest Sounds-in-Sentences subtest	KSPT (Kaufman, 1995) GFTA (Goldman & Fristoe, 1972)	Children with PD were less accurate in general, but showed no greater disadvantage for low-frequency sequences than their age pairs. Adults generalizations differed from child's phonology generalizations: they did not usually result in production errors unless the system is stressed. These results indicated that phonological acquisition involves not only the development of well-practiced articulatory and acoustic–auditory representations but also the emergence of a symbolic representation.
A37	Wang, Kent, Duffy, Thomas & Weismer 2004	Yes	Syllable-AMR tasks 3 minutes of conversational speech (narratives) Acoustic analysis	Diadochokinetic /pa/, /ta/, /ka/, /da/, and /as/ Connected speech sample Kay Elemetrics Model 5500 digital Sona-Graph®	Quantitative and qualitative acoustic analyses of the AMR task were both feasible and informative for the study of TBI-induced dysarthria patients who vary widely in their intelligibility or severity ratings. Acoustic analyses provided specific information on motor speech limitations in individuals with TBI, and these analyses could potentially be related to other speaking tasks such as sentence recitation or conversation.

A38	Prathanee, Thanaviratananich & Pongjanyakul 2003	No	Oral DDK tasks (lip, tongue, and lip and tongue functions)	DDK Visi Pitch® software	A basis for assessment of children with neuromotor deficits or oral structure impairment was provided for both Thai and South East Asian children with similar physiological development and languages with similar sounds. Thai children differed from those of English-speaking children.
A39	Watkins, Dronkers & Vargha-Khadem 2002	No	Repetition of 40 words and 40 non-words Naming of 36 line drawings Generation of many words as possible in 2 min Inflectional and derivational morphological production through a picture presentation to complete a pair of sentences.	Word and Non-word Repetition Test (Gathercole and Baddeley, 1989) Naming test (Oldfield and Wingfield, 1965) Phonemic and Semantic Verbal Fluency Tasks Morphological Production Test (Vargha-Khaden et al., 1991)	Results showed that the articulation test score successfully discriminated between affected and unaffected family members: affected family members and patients with aphasia had remarkably similar profiles of impairment on the testes administered. However, patients with aphasia had enjoyed a normal course of cognitive development and language experience.
A40	Thoonen et al. 1999	No	Maximum phonation duration of [a] and [mama], maximum fricative duration of [f], [s], [z], maximum monosyllabic [pa], [ta], [ka] and trisyllabic [pataka] repetition rate.	MPT (Thoonen, et al., 1996)	Sensitivity and specificity values ranging from 89 to 100% were obtained. It could be concluded that the diagnostic procedure yields quantitative measures of the degree to which dysarthria or apraxia plays a role in the development and maintenance of speech disorders in children.
A41	Lewis & Freebairn 1998	No	Sounds-in-Words subtest Stimulability subtest Sounds-in-Sentences subtest. 44 words elicited by the GFTA on the Sounds-in-Words subtest Oral motor assessment including DDK	GFTA (Goldman & Fristoe, 1972) KLPA (Khan & Lewis, 1986) Oral Speech and Motor Control Protocol (Robbins & Klee, 1987)	Results demonstrated that speech production, as measured by these tasks, continues to improve into adulthood. Ffamily members who reported histories of childhood speech and language problems performed more poorly on these challenging articulatory tasks than did individuals without such a history. These tasks were positively

			Rapid repetition of 20 multisyllabic words and 15 multisyllabic nonsense words Verbal response to stimulus pictures	Multisyllabic Word List (Catts, 1986) CELF-3 (Semel, Wiig & Secord, 1995)	correlated with reading, spelling, and language achievement measures, thus suggesting a relationship between spoken and written language.
A42	Thoonen et al. 1997	Yes	10 min of spontaneous speech Imitation of eight short sentences Imitation of 30 multi-syllabic real words and 36 two- and three syllable nonsense words Imitation of the oral presentation by the examiner	Connected speech sample Self-authored assessment instruments: Phonetic assessment	Children with CAS produced similar types of consonant errors, which corroborated the elicitation method as a valid procedure for evaluation. This study formed a basis for the construction of a test for developmental apraxia of speech. The results emphasized the importance of a standardized procedure and the analysis of a comprehensive set of speech characteristics that allows for the assessment of a speech profile.
A43	Thoonen et al. 1994	No	Maximum performance with respect to respiration, voicing, and articulation 10 min of spontaneous speech Imitation of 30 multi-syllabic real words and 36 two- and three syllable nonsense words Imitation of eight short sentences	MPT (Wit et al., 1993) Connected speech sample Self-authored assessment instruments: Phonetic assessment	The characteristics of feature retention detected in this study for the DVD group were in accordance with previous descriptions of DVD. The main result of the present study which was not expected on a priori grounds-was the striking qualitative similarity in the patterns of feature retention, feature-value retention, feature-value preference, and assimilation for the DVD and control groups.
A44	Wit et al. 1994	Yes	Sustaining /a/, /z/, /f/, /s/, and repeating /ma/ Frequency range from lowest to highest, expressed in semitones Repeating monosyllabic sequences as /papa.../, /tata.../, /kaka.../ as quickly as possible	Non-invasive MPT: MSP FFR MMR	The performance of the PSC on all MPT tasks was poorer than that of their peers with normal speech. In contrast, the TSC performed within the normal limits on MSP and FFR, but their MMR was extremely slow. Authors concluded that the three MPT could be used to uncover subtle differences in the manifestations of the pathophysiological condition spastic dysarthria in children.

A45	Sturner et al. 1993	Yes	Repetition of 15 sentences containing a total of 104 morphemes and 30 phonemes Speech production through drawings of common objects or events	SRST (Sturner, et al., 1993) AAPS (Fudala, 1974)	Results indicated that an elicited-imitation task can predict the combined outcome of receptive and expressive language problems, as well as articulation problems. The study demonstrated that the use of sentence-repetition screening tasks could be a very efficient strategy for screening for both language and articulation problems in kindergarten children.
A46	Hardcastle, Morgan Barry & Clark 1987	Yes	Repetition of four word lists consisting of 43 single-word items (with all the lingual consonants of English in a variety of vowel and consonant-cluster environments) using flashcards that contained both the written word and a pictorial representation	EAT (Anthony et al., 1971) EPG	EPG provided relevant diagnostic information in that all 4 experimental subjects showed patterns that differed from the control subjects in both spatial configuration and variability. The nature of their distorted patterns allowed a tentative diagnosis of 2 of the children as verbal dyspraxic. Instrumental analysis was able to provide more details on the precise nature of their speech activity.
A47	Cermak, Ward & Ward 1986	Yes	Identifying items in 50 line drawings Speech sample of 50 connected words (answering and naming) Receptive vocabulary tasks using single words	Templin-Darley Articulation Screening Test (Templin & Darley, 1964) Standard questions about address, classroom, and family, and numbers and colors naming. PPVT (Dunn & Dunn, 1981)	Children with articulation problems had more problems with motor coordination and slight neurological signs than the TD children. The results of this study allowed several contributions for occupational therapists: the relationship between articulation problems and motor coordination suggested that OT screening for motor coordination disorders should be considered in children with articulation problems; OT using sensory integration procedures had been found to improve language in children with sensory integration problems.
A48	Lapko & Bankson 1975	Yes	Stimulability of /s/ Articulation of /s/	Carter-Buck Nonsense-Syllable Imitation Test (Carter & Buck, 1958) The McDonald Deep and Screening Tests of Articulation (McDonald, 1964)	A significant correlation between the child's ability to discriminate his own production of /s/ (internal or self-monitoring) and the consistency of misarticulation of /s/ was obtained, as well as one between the consistency of misarticulation of /s/ and the stirnulability of /s/. No statistically significant correlations were found between the other variables. A low non-significant correlation was

			Internal and external auditory discrimination of the /s/ sound	Farquhar-Bankson In-depth Test of Auditory Discrimination (Farquhar, 1961)	found between the stimulability of /s/ and internal discrimination abilities.
A49	Yoss & Darley 1974	Yes	Expressive and receptive language tasks: 1. Response to auditory stimulation with words containing specified single consonant sounds 2. Single consonant words imitation 3. Two- and three-consonant words imitation 4. Sentence repetition 7. Ability to move the tongue independently of the jaw and lips in the syllable /la/ Imitation of three two-item sequences and three three-item sequences, which were demonstrated for each subject by the examiner 2 min sample of spontaneous contextual speech (180 words) was elicited from each subject to provide a 60-word corpus Repetition of 13 nonsense words and 13 real words in CVC form	Utah Test of Language Development (Mecham, Jex, and Jones, 1967) Part 1-4 and 7 of the PSTA (Van Riper and Erickson, 1968) IVOM and SVOM Phoneme Production in Spontaneous Contextual Speech Phoneme Production of Real and Nonsense Words	Highly significant differences were found between TD subjects and subjects with defective articulation. A rationale was established for division of the defective articulation group on the basis of their performance on isolated volitional oral movement tasks. Combinations of variables that emerged as statistically significant differentiating predictors between these two subgroups of subjects with defective articulation were neurologic ratings, two- and three-feature errors, distortions, prolongations and repetitions, additions, one-place errors, and omissions.

A50	Wright, Shelton & Arndt 1969	Yes	Receptive vocabulary tasks using single words Speech sample of 50 connected words	PPVT (Dunn & Dunn, 1981) Templin-Darley Screening Test (Templin and Darley, 1960)	Subjects made articulation improvement on the imitative task. They made less improvement on the reading task and much less improvement on the talking task. Problems concerning automatization were more pressing than comparison of acquisition methods.

BBToP = Bankson-Bernthal Test of Phonology; BB-ToP CI = BBToP Consonant Inventory; DEAP = Diagnostic Evaluation of Articulation and Phonology; DEAP Artic = DEAP Articulation subtest; DEAP WI = DEAP Word Inconsistency subtest; ABA-2: Apraxia Battery for Adults 2nd edition; STI-CH: Swedish Test of Intelligibility for Children; GFTA-2: Goldman-Fristoe Test of Articulation 2nd edition; CSPT: Segmental Test of Cantonese; WOLD: Weschler Objective Language Dimensions; PPVT = Peabody Picture Vocabulary Test; KLPA-2 Khan-Lewis Phonological Analysis 2nd edition; VMPAC: Verbal Motor Production Assessment for Children; TELD-3: Test of Early Language Development-3; AFC: Phonological Assessment of the Child; NRT: Non-word Repetition test; OSMSE-R: Oral Speech Mechanism Screening Evaluation-Revised; KBIT: Kaufman-Brief Intelligence Test KSPTC: Kaufman Speech Praxis Test for Children; BPVS: British Picture Vocabulary Scales; M-ABC: Movement Assessment Battery for Children; Maximum Sound Prolongation (MSP); Fundamental Frequency Range (FFR); Maximum Repetition Rate (MRR); Arizona Articulation Proficiency Scale (AAPS);EAT: Edinburgh Articulation Test; IVOM: Isolated Volitional Oral Movements; SVOM: Sequenced Volitional Oral Movements; EPG: Elecropalatography; PSTA: Predictive Screening Test of Articulation.

The studies were analyzed with a focus on four topics: 1) participant characteristics (information on the population assessed by the studies); 2) speech articulation assessment methods; 3) speech articulation assessment instruments; and 4) psychometric properties of speech articulation assessment instruments.

3.1 Population assessed

The majority (78%; n=39) of studies investigated children aged 6 years to 9 years 9 months, although 54% (n=27) of studies also included clients older than 9 years.

A total of 62% (n=31) of studies assessed articulatory disorders caused by neurological damage, with 35% (n=11) focusing on childhood apraxia of speech and dysarthria (A5, A10, A7, A12, A25, A37, A40, A43, A44, A49), 3% (n=1) on structural anomalies (cleft lip/palate)(A32) and 61% (n=19) on phonological disorders (A2, A5, A10, A11, A12, A15, A19, A22, A23, A25, A26, A31, A34, A36, A40, A42, A43, A46, A47).

In 10% (n=5) of studies, speech articulation was evaluated in individuals with specific language disorders (A3, A26, A31, A33, A46), while 4% (n=2) focused on individuals with stutters (A4, A17), 3% (n=1) looked at myotonic dystrophy type 1 (A6), 4% (n=2) investigated patients with cerebral palsy (A9, A24) and 8% (n=4) examined individuals with other clinical conditions (A16, A27, A28, A39). Additionally, 28% (n=14) of investigations included only typically developing speakers (A1, A8, A13, A14, A18, A20, A21, A29, A30, A35, A38, A41, A45, A48, A50).

In addition to English-speaking populations, studies also examined Swedish (A6, A8, A15), Italian (A1, A4, A11), Chinese (A21), Brazilian Portuguese (A19, A22, A27, A31), Hebrew (A13) and Thai (A38) speakers.

3.2 Speech articulation assessment methods

Several methods were used to assess articulation in the studies reviewed. The most commonly reported methods included word and nonword repetition, naming, enunciation, conversation, and diadochokinetic speech tasks.

In 46% (n=23) of studies, the researchers used assessment protocols developed by their own research groups to assess the repetition or enunciation of syllables, words and sentences (A2, A4, A5, A6, A7, A12, A14, A19, A21, A22, A26, A28, A31, A32, A33, A35, A37, A39, A42, A43, A44, A47, A49). The repetition of word and non-word lists was also used in 32% (n=16) of the studies reviewed (A1, A4, A16, A17, A18, A20, A29, A30, A32, A36, A39, A41, A42, A43, A48, A49). Most studies described the development of the assessment procedures used, from the selection of visual stimuli for naming tasks to the composition of syllable, word and sentence lists for verbal repetition tests

Another method used in these studies was semi-structured elicitation, with 30% (n=15) of articles collecting speech samples from image descriptions, film interpretations, and conversations about topics of interest (A2, A6, A7, A12, A19, A32, A33, A35, A37, A39, A42, A43, A44, A47, A49). These speech samples were phonetically transcribed to determine the Percentage of Consonants Correct (PCC; Shriberg & Kwiatkowski, 1982).

In 28% (n=14) of articles, speech articulation was assessed using two traditional diadochokinetic tasks: the alternating motion rate (AMR) (/pa/, /ta/ and /ka/) and the sequential motion rate (SMR) (/pa.ta.'ka/) (A7, A9, A13, A16, A19, A24, A25, A28, A33, A35, A37, A38, A40, A49). Some of these studies also used additional stimuli, such as "buh", "uhba", "Buy Bobby a puppy" (A9, A24), and "Pop", "Puppet" and "Puppypop" (A25).

In addition to assessing speech articulation, 26% (n=13) of studies also investigated oral language development, including phonology; cognitive processes, such as working memory and phonological awareness; and the automatization of speech production (A5, A9, A12, A16, A20, A23, A24, A25, A29, A33, A34, A36, A41). The instruments used to assess these functions included versions of the Clinical Evaluation of Language Functions (CELF-P, CELF-3, CELF-4; Wiig, Secord & Semel, 1992; Semel, Wiig & Secord, 1995; 2003) and the Structured Photographic Expressive Language Test (SPELT-P, SPELT-3; Dawson, Stout & Eyer, 2003; Dawson, Eyer & Fonkalsrud, 2005).

The methods used to collect speech data, such as repetition, naming and spontaneous speech tasks, did not change over the years. However, the approaches used to describe participants' speech profile, and the applications of the tests used, differed between studies. The variables assessed by these studies included speech articulation (A11), nasality and articulation (A8), intelligibility (A15), non-word repetition (A1, A30) and articulation rate (A40).

3.3 Speech articulation assessment instruments

A total of 64% (n=32) of studies used instruments or tests to evaluate speech articulation. In this review, speech articulation assessment instruments were analyzed based on their target age group, the language for which they were developed, the contents of the tests, the length of assessment, the method of collecting speech data, as well as the availability of psychometric evidence. This data are shown in the Table 2.

Table 2. Characteristics of the speech articulation assessment instruments

Instrument	Age (years)	Language	Test content	Duration (minutes)	Data collection	Psychometric properties				
						Validity	Reliability	Standardiz. ation	Normative data	Partial validity & reliability
Non-Word Repetition Test	3 – 7	Italian	List with 46 nonwords	2 – 4	Single nonwords production with different structures and sizes	x	x		x	
The Swedish Articulation and Nasality Test (SVANT)	3 – 19	Swedish	74 images	n/a	Single-word production, sentence repetition, and connected speech				x	
Schindler's repetition test	3 – 10:8	Italian	List with 30 words	2	Single-word production through repetition	x	x		x	
Swedish Test of Intelligibility for Children (STI-CH)	4:6 – 8:3	Swedish	List with 60 words	n/a	Word repetition	x	x			
Linus	3 – 6	Swedish	Slide show with 107 figures	n/a	Single-word production				x	
Syllable Repetition Task (SRT)	3 – 17	English	List with 18 nonwords	n/a	Nonword repetition	x	x			
Test of Children's Speech+ (TOCS+)	3 – 7	English	Software	10 – 15	Single words and sentences production	x	x			
Speech Intelligibility Test (SIT)	n/a	English	Software	n/a	Single-words and sentences production		x			
Diagnostic Evaluation of Articulation and Phonology (DEAP)	3 – 8:11	English	Figures for appointment	screen: 5 others:15	Single-word production, proof of stimulability, and proof of inconsistency	x	x	x		

Khan-Lewis Phonological Analysis—2 (KLPA-2)	2 – 21:11	English	53 target words elicited by GFTA-2 (Sounds-in-Words)	10 – 30	Single-word production	x	x	x		
Goldman-Fristoe Test of Articulation (GFTA e GFTA-2)	2 – 21:11	English	53 cards with pictures	5 – 15	Single-word production Connected speech Word repetition for stimulability test	x	x	x	x	
Phonology Test of the Child Language Test (ABFW)	3 – 12	Brazilian Portuguese	34 figures and 39 words	n/a	Single-words production in picture naming and word repetition				x	
Apraxia Battery for Adults, Second Edition (ABA-2)	9+	English	Battery with 6 subtests with pictures and wordlist	20	Diadokokinetic rate, word output, latency and expression time for polysyllable words, and repeated-test testing	x	x			
Verbal Motor Production Assessment for Children (VMPAC)	3 – 12	English	Battery with 82 items (16 for sequencing skills)	n/a	Sequence production of duplicate and triplicate phonemes					x
Maximum Performance Tasks (MPT)	6 – 10	English	Maximum Phonation Duration (MPD) and Maximum Repeat Rate (MRR)	n/a	Extension of [a], [f], [s], [z] Repetition of [pa], [ta], [ka] Repetition of [pataka]	x				
Kaufman Speech Praxis Test for Children (KSPT)	2 – 6	English	List of vowels, diphthongs, consonants, duplicate syllables and words	5 – 15	Stimulus repetition	x		x	x	
Sentence Repetition Screening Test (SRST)	4 – 8	English	15 sentences with 30 phonemes	3	Sentence repetition	x				

Cantonese Segmental Phonology Test (CSPT)	2 – 6:6	Cantonese	List with 31 words and 1 picture	6	Single-word production and connected speech	x	x		
Phonological Assessment of the Child (AFC)	3+	Brazilian Portuguese	5 thematic figures with 125 stimulus items	n/a	Production of single-words by appointment or delayed imitation				
Bankson-Bernthal Test of Phonology (BBToP)	3 – 9	English	80 pictures	10 – 15	Naming	x	x	x	
Multisyllabic Word List	12 – 16	English	List with 45 multisyllabic words	n/a	Single-word production				
Fluharty Preschool Speech and Language Screening Test	2 – 6	English	15 figures, 10 sentences	n/a	Single-word production and sentence repetition	x			
Arizona Articulation Proficiency Scale (AAPS)	3+	English	Picture test cards and sentence test	10 – 15	Single-word production, spontaneous speech, and sentence repetition	x	x	x	x
Edinburgh Articulation Test (EAT)	3 – 6	British English	Book with 41 images	6 – 15	Single-word production			x	
The McDonald Deep and Screening Tests of Articulation	4 – 7	English	Screen: 90 two-syllable words Deep: screening test + picture & sentence form + 25 one-syllable words	n/a	Single-word production	x		x	x
Predictive Screening Test of Articulation	6 – 7	English	List 38 words, 1 sentence, 2 sounds, 3 syllables	5 – 10	Words, sentences, syllables and sounds production	x			
Templin-Darley Articulation Screening Test	3 – 8	English	19 cards with 50 figures	n/a	Single-word production or word repetition or sentence production	x	x		x

n/a: not available

3.4 Psychometric properties of speech articulation assessment instruments

Two of the speech articulation instruments used by the studies reviewed had well-established psychometric properties. The Goldman-Fristoe Test of Articulation (GFTA, GFTA-2; Goldman & Fristoe, 1972; 2000) and Arizona Articulation Proficiency Scale (AAPS; Fudala, 1974) are standardized and norm-referenced tests with ample evidence of validity and reliability, supporting their psychometric integrity and the reliability of their results.

Normative data were also available for the SVANTE–The Swedish Articulation and Nasality Test (Lohmander, Lundeborg & Persson, 2017), Linus (Blumenthal & Lundeborg Hammarström, 2014) and Child Vocabulary Test (Teste de Linguagem Infantil - ABFW; Andrade, et al., 2004), though there is still a need for studies of the validity and reliability of these instruments.

4 Discussion

This review aimed to provide information on the most adequate methods for the study and evaluation of speech articulation in children and adolescents. The majority of studies reviewed used specific assessment instruments to evaluate speech articulation. However, some investigations also relied on the repetition or enunciation of words, syllables, and sentences to assess articulation skills.

Most investigations used assessment batteries constructed by the researchers themselves to evaluate multiple skills, including speech articulation but also orofacial movement, expressive vocabulary, and auditory discrimination.

To diagnose SSDs, SLPs used a variety of assessment methods, including connected speech samples, naming, stimulability, and intelligibility classifications. However, the most commonly used task in this regard was the production of single words containing a given target sound (with some assessments also including vowels, consonant clusters, and tones).

Studies conducted in the United States and Australia found that most SLPs always used single word articulation tests to evaluate the frequency and accuracy of consonant production (Mc Leod & Masso, 2018; Skahan, Watson & Lof, 2016; McLeod & Verdon, 2014).

However, from the perspective of some SLPs who participated in these studies, some assessments – including measures of speech articulation – may have a negative impact on clinical practice due to time and resource constraints (McLeod & Verdon, 2014). It is essential that before administering any instrument – including those used to evaluate speech articulation - clinicians familiarize themselves with aspects of test administration such as duration, resources required, and the target population. Familiarity with the psychometric properties of the test when these are available is also of great importance for the selection, administration, and interpretation of assessment instruments (McLeod, 2012; McCauley & Swisher, 1984).

The imitation tasks used to evaluate articulation only capture the stages of speech processing between phonological coding and articulatory production (Dell, Martin & Schwartz, 2007). As such, robust language processing techniques may not be required to complete these tasks (Baddeley, Hitch & Allen, 2009; Dell, Martin & Schwartz, 2007). Spontaneous naming, on the other hand, places a greater demand on cognitive and language abilities, including perceptual, semantic and grammatical processes (Kurland, Reber & Stokes, 2014), as well as phonological processing, articulatory planning, and execution.

However, questions remain regarding the use of imitation versus the spontaneous production of target words for articulation assessment. While some studies have found that consonant production accuracy is greater in imitation tasks (DuBois & Bernthal, 1978; Johnson

& Somers, 1978), others have not identified any significant differences in this variable between methods of assessment (Powell, 1997; Andrews & Fey, 1986).

Non-word lists are often used in the literature to evaluate other skills in addition to articulation since these tasks also require skills such as speech perception, phonological coding, phonological assembly, lexical knowledge, and motor planning (Coady & Evans, 2008; Gathercole, 2006). Diadochokinesis rates also provide a measure of oromotor skills that contribute to diagnosis and the differentiation between functional articulation disorders and conditions caused by neurological impairment (Cohen & Waters, 1999; Bradford, Murdoch, Thompson & Stokes, 1997).

To provide an accurate diagnosis, articulation assessments must include word repetition and naming, as well as sentence production and/or imitation and connected speech, since speech production can improve under controlled conditions (McLeod & Masso, 2018) because children with SSD can experience problems at the level of lexeme retrieval, phonological encoding, articulo-motor planning, and programming, and/or execution (van Haaften, et al., 2019).

Regarding the psychometric evidence of speech articulation assessment instruments, the GFTA and GFTA-2 were the most widely used and accepted by researchers. Both are also widely used in clinical and research settings within speech-language pathology. Since most studies used the speech assessment instruments to divide samples into clinical and control groups, the reliability of the instruments could not be examine.

5 Conclusion

This systematic review aimed to identify and describe the methods and instruments used to assess of speech articulation, and report on their psychometric properties, especially their validity and reliability.

According to the information provided in the articles reviewed, most studies were authored by SLPs, while in others, the role of the SLPs was limited to the administration of speech assessment instruments. Research methods did not differ between studies with and without a SLP as an author. Research methods also showed no changes over the years.

Most studies focused on children up to nine years old, and a lack of speech assessment studies in adolescent populations was observed. Over half the studies used speech articulation instruments to evaluate people with SSDs caused by neurological, structural or functional/phonological conditions.

As for the instruments and methods used to evaluate speech articulation, most studies relied on the repetition of words and nonwords, naming, enunciation, conversation, and diadochokinetic speech tasks, and also self-authored instruments to assess speech articulation.

All of the instruments used in these research studies had adequate methodological quality. However, only the Goldman-Fristoe Test of Articulation (GFTA, GFTA-2; Goldman & Fristoe, 1972; 2000) has undergone a full psychometric evaluation, including assessments of validity and reliability, standardization, and norm-referencing.

This constitutes a limitation in phonological assessment in Brazilian Portuguese, with a negative impact on diagnosis and treatment planning for children with SSDs, especially in the case of inexperienced examiners. The development of assessment protocols with robust psychometric properties is crucial for clinical practice and research in speech pathology. In addition to psychometric studies, information technology may also contribute to phonological assessment. The use of computer software for phonological assessment may be both more

appealing to children and advantageous for practitioners, as it makes for faster and simpler test administration.

This review strongly recommends the use of instruments with established psychometric properties and confirmed applicability to the population of interest in the assessment of speech articulation and all other aspects of language. This is crucial to ensure the adequate diagnosis and treatment planning for children, adolescents and other people with SSDs. Additionally, it is important that multiple assessment methods be used to evaluate speech in both clinical and research settings in order to provide a more comprehensive evaluation of speech production.

REFERENCES

1. Van Severen L, Gillis JJ, Molemans I, Van Den Berg R, De Maeyer S, Gillis S. *The relation between order of acquisition, segmental frequency and function: the case of word-initial consonants in Dutch.* J Child Lang. Sep;40(4):703-40, 2013.
2. Goldman R, Fristoe M. *Goldman-Fristoe Test of Articulation 2 - Test manual.* Circle Pines, MN: American Guidance Services Inc: 2000.
3. Stoel-Gammon C, Sosa AV. *Phonological Development.* In E. Hoff & M. Shatz (Eds.), Blackwell Handbook of Language Development (pp. 238–254). United Kingdom: Blackwell Publishers Ltd, 2009.
4. Newmeyer AJ, Grether S, Grasha C, White J, Akers R, Aylward C, Ishikawa K, Degrauw T. *Fine motor function and oral-motor imitation skills in preschool-age children with speech-sound disorders.* Clin Pediatr (Phila). Sep;46(7):604-11, 2007.
5. Ferrante C, Borsel JV, Pereira MMDB. *Análise dos processos fonológicos em crianças com desenvolvimento fonológico normal.* Revista da Sociedade Brasileira de Fonoaudiologia, v. 14, n. 1, p. 36-40, 2009.
6. Rabelo ATV, Alves CRL, Goulart LMHF, Friche AAL, Lemos SMA, Campos FR, Friche CP. *Alterações de fala em escolares na cidade de Belo Horizonte.* J. Soc. Bras. Fonoaudiol, 23(4): 344-350, 2011.
7. Lee AS, Whitehill TL, Ciocca V, Samman N. *Acoustic and perceptual analysis of the sibilant sound /s/before and after orthognathic surgery.* Journal of Oral and Maxillofacial Surgery, 60, 364–372, 2002.
8. Ceron MI. *Instrumento de Avaliação Fonológica (INFONO): Desenvolvimento e estudos psicométricos.* Universidade Federal de Santa Maria, Santa Maria; 2015.
9. Preston JL, Koenig, LL. *Phonetic variability in residual speech sound disorders: Exploration of subtypes.* Top Lang Disord, 31(2):168–184, 2011.
10. Veríssimo A, Van Borsel J, Pereira MB. *Residual /s/ and /r/ distortions: The perspective of the speaker.* J International Journal of Speech-Language Pathology, 14(2):183–186, 2012.
11. Preston JL, McAllister T, Phillips E, Boyce S, Tiede M, Kim JS, Whalen DH. *Treatment for Residual Rhotic Errors With High- and Low-Frequency Ultrasound Visual Feedback: A Single-Case Experimental Design.* J Speech Lang Hear Res, 61(8):1875-1892, 2018.
12. American Speech and Hearing Association (ASHA). *Speech Sound Disorders-Articulation and Phonology*. 2013.
13. Morrison JA, Shriberg LD. *Articulation testing versus conversational speech sampling.* J Speech Lang Hear Res, 35:259–273, 1992.
14. Klintö K, Salameh EK, Svensson H, Lohmander A. *The impact of speech material on speech judgement in children with and without cleft palate.* Int J Lang Commun Disord, 46(3), 348–360, 2011.
15. Masterson JJ, Bernhardt BH, Hofheinz MK. *A comparison of single words and conversational speech in phonological evaluation.* Am J Speech Lang Pathol, 14(3), 229–241, 2005.
16. Paul R. *Language Disorders from Infancy through Adolescence: Assessment and Intervention*. St Louis, Mosby, pp 288– 290, 1995.
17. Clements GN. *The Geometry of Phonological Features.* In: Phonology Yearbook, p. 225-252, 1985.
18. Martinelli RLC, Marchesan IQ. *Aspectos da Fala nas alterações de frênulo lingual.* In: Berretin-Félix Giédre et al Org. A Fala nos diversos contextos da Motricidade Orofacial. Editora Pulso São José dos Campos, 2015.

19. Cuestas G, Demarchi V, Corválan MPM, Razetti J, Boccio C. *Tratamiento quirúrgico del frenillo lingual corto en niños.* Arch Argent Pediatr, 112(6):567-70, 2014.
20. Alves GAS, Lima ILB, Lima JAS, Delgado ICD. *Aspectos da fala na Síndrome de Down.* In: DELGADO, I.C. (Org.). Contribuições da Fonoaudiologia na Síndrome de Down. Ribeirão Preto: Book Toy, 2016.
21. Kent RD, Read C. *The Acoustic Analysis of Speech.* 2. ed. Madison, Wisconsin: Singular Thomson Learning, 2002.
22. McLeod S, Verdon S. *A Review of 30 Speech Assessments in 19 Languages Other Than English.* American Journal of Speech-Language Pathology, Vol. 23, 708–723, 2014.
23. Abou-Elsaad T, Baz H, El-Banna M. *Developing an articulation test for Arabic-speaking school-age children.* Folia Phoniatr Logop, 61(5):275-82, 2009.
24. Caesar LG, Kohler PD. *The state of school-based bilingual assessment: Actual practice versus recommended guidelines.* Language, Speech, and Hearing Services in Schools, 38, 190–200, 2007.
25. Jordaan H. *Clinical intervention for bilingual children: An international survey.* Folia Phoniatrica et Logopaedica, 60, 97–105, 2008.
26. Roseberry-McKibbin C, Brice A, O'Hanlon L. *Serving English language learners in public school settings:Anational survey.* Language, Speech, and Hearing Services in Schools, 36, 48–61, 2005.
27. Skahan SM, Watson M, Lof GL. *Speech-language pathologists' assessment practices for children with suspected speech sound disorders: Results of a national survey*. Am J Speech Lang Pathol., 16, 246–259, 2016.
28. Williams C, McLeod S. *Speech-language pathologists' assessment and intervention practices with multilingual children.* Int J Speech Lang Pathol, 14, 292–305, 2012.
29. Moher D, Liberati A, Tetzlaff J, Altman D, Group P. *Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement.* Journal of Clinical Epidemiology, 62: 1006-1012, 2009.
30. Piazzalunga S, Previtali L, Pozzoli R, Scarponi L, Schindler A. *An articulatory-based disyllabic and trisyllabic Non-Word Repetition test: reliability and validity in Italian 3- to 7-year-old children.* Clin Linguist Phon, 33(5):437-456, 2019.
31. Öster Cattu Alves M, Ode C, Strömbergsson S. *Dealing with the unknown - addressing challenges in evaluating unintelligible speech.* Clin Linguist Phon, 1-16, 2019.
32. Blumenthal C, Lundeborg Hammarström I. *LINUS preliminärmanual från*. Sweden: Dept. of Neuroscience and Locomotion/Speech Pathology, Linköping University, 2014.
33. Lagerberg TB, Hartelius L, Johnels JÅ, Ahlman AK, Börjesson A, Persson C. *Swedish Test of Intelligibility for Children (STI-CH)--validity and reliability of a computer-mediated single word intelligibility test for children.* Clin Linguist Phon., 29(3):201-15, 2015.
34. Saletta M, Goffman L, Ward C, Oleson JJ. *Influence of Language Load on Speech Motor Skill in Children With Specific Language Impairment.* J Speech Lang Hear Res, 675-689, 2018.
35. Bankson NW, Bernthal JE. *Bankson-Bernthal Test of Phonology.* Austin, TX: Pro-Ed, 1990.
36. Lenoci G, Ricci I. *An ultrasound investigation of the speech motor skills of stuttering Italian children.* Clinical Linguistics & Phonetics, 32. 1-19, 2018.
37. Vuolo J, Goffman L. *Language Skill Mediates the Relationship Between Language Load and Articulatory Variability in Children With Language and Speech Sound Disorders.* J Speech Lang Hear Res, 61(12):3010-3022, 2018.
38. Dodd B, Hua Z, Crosbie S, Holm A, Ozanne A. *Diagnostic Evaluation of Articulation and Phonology (DEAP).* San Antonio, TX: Psychological Corporation, 2006.
39. Dawson J, Stout C, Eyer J. *Structured Photographic Expressive Language Test-3.* Dekalb, IL: Janelle Publications, 2003.
40. Dawson J, Eyer JA, Fonkalsrud J. *Structured Photographic Expressive Language Test—Preschool: Second Edition.* DeKalb, IL: Janelle Publications; 2005.
41. Sjögreen L, Mårtensson Å, Ekström AB. *Speech characteristics in the congenital and childhood-onset forms of myotonic dystrophy type 1.* Int J Lang Commun Disord, 53(3):576-583, 2018.
42. Lohmander A, Lundeborg I, Persson C. *SVANTE – The Swedish Articulation and Nasality Test – Normative data and a minimum standard set for cross-linguistic comparison*. Clinical Linguistics & Phonetics, 31:2, 137-154, 2017.
43. Rvachew S, Matthews T. *Using the Syllable Repetition Task to Reveal Underlying Speech Processes in Childhood Apraxia of Speech: A Tutorial.* Canadian Journal of Speech-Language Pathology and Audiology, 41(1), 106-126, 2017.

44. Lohmeier HL, Shriberg LD. *Tech Rep No 17. The Phonology Project*, Waisman Center, University of Wisconsin-Madison. Reference Data for the Syllable Repetition Task,2011.
45. Thoonen G, Maassen B, Gabreels F, Schreuder R. *Validity of maximum performance tasks to diagnose motor speech disorders in children*. Clinical Linguistics & Phonetics, 13:1, 1-23,1999.
46. Nip IS. *Interarticulator coordination in children with and without cerebral palsy.* Dev Neurorehabil, 20(1):1-13, 2017.
47. Hodge MM, Daniels J. *Computer software.* Edmonton, AB: University of Alberta. TOCS+ Intelligibility measures, 2007.
48. Yorkson KM, Beukelman DR, Hakel M, Dorsey M. *Computer software.* Lincoln, NE: Institute for Rehabilitation Science and Engineering at Madonna Rehabilitation Hospital, Speech Intelligibility Test, 2017.
49. Wiig EH, Secord W, Semel E. *Clinical evaluation of language fundamentals—Preschool.* Toronto, Canada: Psychological Corporation, 1992.
50. Semel E, Wiig EH, Secord WA. *Clinical evaluation of language fundamentals, fourth edition (CELF-4).* Toronto, Canada: The Psychological Corporation/A Harcourt Assessment Company, 2003.
51. Semel E, Wiig EH, Secord WA. *The Clinical Evaluation of Language Fundamentals—3* (San Antonio: Psychological Corporation), 1995.
52. Vuolo J, Goffman L. *An exploratory study of the influence of load and practice on segmental and articulatory variability in children with speech sound disorders.* Clinical linguistics & phonetics, 31. 1-20, 2016.
53. Tresoldi M, Ambrogi F, Favero E, Colombo A, Barillari MR, Velardi P, Schindler A. *Reliability, validity and normative data of a quick repetition test for Italian children.* Int J Pediatr Otorhinolaryngol, 79(6):888-94, 2015.
54. Murray E, McCabe P, Heard R, Ballard KJ. *Differential diagnosis of children with suspected childhood apraxia of speech.* J Speech Lang Hear Res, 58(1):43-60, 2015.
55. Icht M, Ben-David BM. *Oral-diadochokinetic rates for Hebrew-speaking school-age children: real words vs. non-words repetition.* Clin Linguist Phon, 29(2):102-14, 2015.
56. Lin S, Demuth K. *Children's acquisition of English onset and coda /l/: articulatory evidence.* J Speech Lang Hear Res, 58(1):13-27, 2015.
57. Turner SJ, Mayes AK, Verhoeven A, Mandelstam SA, Morgan AT, Scheffer IE. *GRIN2A: an aptly named gene for speech dysfunction.* Neurology, 10;84(6):586-93,2015.
58. Dabul BL. *Apraxia Battery for Adults*, 2nd ed. Austin: PRO-ED; 2000.
59. Spencer C, Weber-Fox C. *Preschool speech articulation and nonword repetition abilities may help predict eventual recovery or persistence of stuttering.* J Fluency Disord, 41:32-46, 2014.
60. Dollaghan C, Campbell TF. *Nonword repetition and child language impairment.* Journal of Speech, Language, and Hearing Research, 41, 1136-1146, 1998.
61. Krishnan S, Alcock KJ, Mercure E, Leech R, Barker E, Karmiloff-Smith A, Dick F: *Articulating novel words: children's oromotor skills predict nonword repetition abilities.* J Speech Lang Hear Res, 56(6):1800-12, 2013.
62. Wagner RK, Torgesen JK, Rashotte CA. *Comprehensive Test of Phonological Processing (CTOPP).* Austin, TX: PRO-ED, 1999.
63. Wertzner HF, Pagan-Neves LO, Alves RR, Barrozo TF. *Implications of diadochokinesia in children with speech sound disorder.* CoDAS, 25(1): 52-58, 2013.
64. Andrade CRF, Befi-Lopes DM, Fernandes FDM, Wertzner HF. *ABFW: teste de linguagem infantil nas áreas de fonologia, vocabulário, fluência e pragmática.* 2a ed. Barueri- SP: Pró-Fono; 2004.
65. Wren Y, McLeod S, White P, Miller LL, Roulstone S. *Speech characteristics of 8-year-old children: findings from a prospective population study.* J Commun Disord, 46(1):53-69,2013.
66. Gathercole SE, Baddeley AD. *The Children's Test of Nonword Repetition.* London: Psychological Corporation, 1996.
67. Hack J, Marinova-Todd SH, May Bernhardt B. *Speech assessment of Chinese-English bilingual children: accent versus developmental level.* Int J Speech Lang Pathol,14(6):509-19, 2012.
68. So LKH. *Cantonese Segmental Phonology Test (research version).* Hong Kong: Department of Speech and Hearing Sciences, Hong Kong University, 1993.
69. Castro MM, Wertzner HF. *Stimulability: auxiliary measure in the identification of difficulty in speech sounds production.* J. Soc. Bras. Fonoaudiol, 24(1):49-56, 2012.
70. Peter B. *Oral and hand movement speeds are associated with expressive language ability in children with speech sound disorder.* J Psycholinguist Res, 41(6):455–474, 2012.

71. Goldman R, Fristoe M. *Goldman-Fristoe Test of Articulation.* Circle Pines, Minn.: American Guidance Service, 1972.
72. Khan L, Lewis N. *Khan-Lewis phonological analysis*, 2nd edn. Circle Pines: American Guidance Service, 2002.
73. Nip IS. *Kinematic Characteristics of Speaking Rate in Individuals with Cerebral Palsy: A Preliminary Study.* J Med Speech Lang Pathol, 20(4):88-94, 2013.
74. Moss A, Grigos MI. *Interarticulatory Coordination of the Lips and Jaw in Childhood Apraxia of Speech.* J Med Speech Lang Pathol, 20(4):127-132, 2012.
75. Hayden D, Square P. *Verbal Motor Production Assessment for Children.* San Antonio: The Psychological Corporation; 1999.
76. Newton C. *Between-word processes in children with speech difficulties: insights from a usage-based approach to phonology.* Clin Linguist Phon, 26(8):712-27, 2012.
77. Pizolato RA, Fernandes FSF, Gavião MBD. *Speech evaluation in children with temporomandibular disorders.* J. Appl. Oral Sci, 19(5): 493-499, 2011.
78. Yavas MS, ed. *Phonological disorders in children: theory, research and practice.* London: Routledge; 1991.
79. Ho AK, Wilmut K. *Speech and oro-motor function in children with developmental coordination disorder: a pilot study.* Hum Mov Sci, 29(4):605-14, 2010.
80. Sasisekaran J, Smith A, Sadagopan N, Weber-Fox C. *Nonword repetition in children and adults: effects on movement coordination.* Dev Sci, 13(3):521-532, 2010.
81. Moore MW, Tompkins CA, Dollaghan CA. *Manipulating articulatory demands in non-word repetition: a 'late-8' non-word repetition task.* Clin Linguist Phon, 24(12):997-1008, 2010.
82. Dollaghan C, Campbell T. *A Late-8 Nonword Repetition Task.* Unpublished Manuscript, 2003.
83. Castro MM, Wertzner HF. *Influence of sensory cues on the stimulability for liquid sounds in Brazilian Portuguese-speaking children.* Folia Phoniatr Logop, 61(5):283-7, 2009.
84. Baylis AL, Munson B, Moller KT. *Factors affecting articulation skills in children with velocardiofacial syndrome and children with cleft palate or velopharyngeal dysfunction: a preliminary report.* Cleft Palate Craniofac J, 45(2):193–207, 2008.
85. Archibald LM, Alloway TP. *Comparing language profiles: children with specific language impairment and developmental coordination disorder.* Int J Lang Commun Disord, 43(2):165-80, 2008.
86. Kaufman N. *Kaufman Speech Praxis Test for Children.* Detroit: Wayne State University Press; 1995.
87. Walker JF, Archibald LM. *Articulation rate in preschool children: a 3-year longitudinal study.* Int J Lang Commun Disord, 41(5):541-65, 2006.
88. Fluharty NB. *Fluharty Preschool Speech and Language Screening Test* (Hingham, MA: Teaching Resources), 1978.
89. Munson B, Edwards J, Beckman ME. *Relationships between nonword repetition accuracy and other measures of linguistic development in children with phonological disorders.* J Speech Lang Hear Res, 48(1):61-78, 2005.
90. Wang YT, Kent RD, Duffy JR, Thomas JE, Weismer G. *Alternating motion rate as an index of speech motor disorder in traumatic brain injury*. Clin Linguist Phon, 18(1):57-84, 2004.
91. Prathanee B, Thanaviratananich S, Pongjanyakul A: *Oral diadochokinetic rates for normal Thai children.* Int J Lang Commun Disord, 38(4):417-28, 2003.
92. Watkins KE, Dronkers NF, Vargha-Khadem F. *Behavioural analysis of an inherited speech and language disorder: comparison with acquired aphasia.* Brain, 125(Pt 3):452-64, 2002.
93. Lewis BA, Freebairn L. *Speech production skills of nuclear family members of children with phonology disorders.* Lang Speech, 41 (Pt 1):45-61, 1998.
94. Catts, H. *Speech production/phonological deficits in reading disordered children.* Journal of Learning Disabilities, 19, 503-508, 1986.
95. Thoonen G, Maassen B, Gabreëls F, Schreuder R, de Swart B. *Towards a standardised assessment procedure for developmental apraxia of speech.* Eur J Disord Commun, 32(1):37-60, 1996.
96. Thoonen G, Maassen B, Gabreëls F, Schreuder R. *Feature analysis of singleton consonant errors in developmental verbal dyspraxia (DVD).* J Speech Hear Res, 37(6):1424-40, 1994.
97. Wit J, Maassen B, Gabreëls F, Thoonen G, de Swart B. *Traumatic versus perinatally acquired dysarthria: assessment by means of speech-like maximum performance tasks.* Dev Med Child Neurol, 36(3):221-9, 1994.
98. Sturner RA, Kunze L, Funk SG, Green JA. *Elicited imitation: its effectiveness for speech and language screening.* Dev Med Child Neurol, 35(8):715-26, 1993.

99. Fudala JB. *Arizona Articulation Proficiency Scale: Revised.* Los Angeles: Western Psychological Services, 1974.
100. Hardcastle WJ, Barry RA, Clark CJ. *An instrumental phonetic study of lingual activity in articulation-disordered children.* J Speech Hear Res, 30(2):171-84, 1987.
101. Anthony A, Bogle D, Ingram TTS, Mcisaac MW. *The Edinburgh Articulation Test.* Edinburgh: E. & S. Livingston, 1971.
102. Cermak SA, Ward EA, Ward LM. *The relationship between articulation disorders and motor coordination in children.* Am J Occup Ther, 40(8):546-50, 1986.
103. Templin MC, Darley FL. *Templin-Darley Tests of Articulation*. Iowa City: Bureau of Educational Research and Service, Division of Extension and University Services, 1960.
104. Lapko LL, Bankson NW. *Relationship between auditory discrimination, articulation stimulability, and consistency of misarticulation.* Percept Mot Skills. 1975 Feb;40(1):171-7.
105. McDonald ET. *Articulation testing and treatment: a sensory-motor approach.* Pittsburgh: Stanwix House, 1964.
106. Yoss KA, Darley FL. *Developmental apraxia of speech in children with defective articulation.* J Speech Hear Res, 17(3):399–416, 1974.
107. Van Riper C, Erickson RL. *The Van Riper Predictive Screening Test of Articulation*. Kalamazoo, Michigan: Western Michigan University, 1968.
108. Wright V, Shelton RL, Arndt WB. *A task for evaluation of articulation change. 3. Imitative task scores compared with scores for more spontaneous tasks.* J Speech Hear Res, 12(4):875–884, 1969.
109. Shriberg LD, Kwiatkowski J. *Phonological disorders III: a procedure for assessing severity of involvement.* J. Speech Hear. Dis., v. 47, n. 3, p. 256-270, 1982.
110. McLeod S, Masso S. Screening Children's Speech. *The Impact of Imitated Elicitation and Word Position.* Language, Speech, and Hearing Services in Schools, 28;50(1):71-82, 2018.
111. Skahan SM, Watson M, Lof GL. *Speech-language pathologists' assessment practices for children with suspected speech sound disorders: Results of a national survey.* American Journal of Speech-Language Pathology, 16(3), 246–259, 2007.
112. McLeod S. *Translation to practice: Creating sampling tools to assess multilingual children's speech.* In S. McLeod & B. A. Goldstein (Eds.), Multilingual aspects of speech sound disorders in children (pp. 144–153). Bristol, United Kingdom: Multilingual Matters, 2012.
113. McCauley RJ, Swisher L. *Psychometric review of language and articulation tests for preschool children.* Journal of Speech and Hearing Disorders, 49(1), 34–42, 1984.
114. Dell GS, Martin N, Schwartz MF. *A case-series test of the interactive two-step model of lexical access: Predicting word repetition from picture naming.* Journal of Memory and Language, 56(4), 490–520, 2007.
115. Baddeley AD, Hitch GJ, Allen RJ. Working memory and binding in sentence recall. Journal of Memory and Language, 61(3), 438–456, 2009.
116. Kurland J, Reber A, Stokes P. *Beyond picture naming: norms and patient data for a verb-generation task.* Am J Speech Lang Pathol, 23(2):259-70, 2014.
117. DuBois EM, Bernthal JE. *A comparison of three methods for obtaining articulatory responses.* Journal of Hearing Disorders, 43(3), 295–305, 1978.
118. Johnson S, Somers H. *Spontaneous and imitated responses in articulation testing.* International Journal of Language & Communication Disorders, 13, 107–116, 1978.
119. Powell TW. *Assessing consonant cluster production under imitative and more spontaneous conditions.* Perceptual and Motor Skills, 84, 1134, 1997.
120. Andrews N, Fey ME. *Analysis of the speech of phonologically impaired children in two sampling conditions.* Language, Speech and Hearing Services in Schools 1986; 17(3), 187–198.
121. Coady JA, Evans JL. *Uses and interpretations of non-word repetition tasks in children with and without Specific Language Impairments (SLI).* International Journal of Language & Communication Disorders, 43, 1–40 , 2008.
122. Cohen WMM, Waters D. *Measuring speech motor skills in normally developing and phonologically disordered pre-school children.* In Proceedings of the XIVth International Congress of Phonetic Sciences. pp. 789–792, 1999.
123. Bradford A, Murdoch B, Thompson E, Stokes P. *Lip and tongue function in children with developmental speech disorders: a preliminary investigation.* Clinical Linguistics and Phonetics, 363±387, 1997.
124. Van Haaften L, Diepeveen S, Van den Engel-Hoek L, Jonker M, de Swart B, Maassen B. *The Psychometric Evaluation of a Speech Production Test Battery for Children: The Reliability and*

Validity of the Computer Articulation Instrument. J Speech Lang Hear Res. 62(7):2141-2170, 2019.