Supplemental Material S3. Data extraction sheets.

Data extraction sheet 1: Characteristics of the studies

Author, year, county	Study characteristics					Demographic details		Clinical characteristics				
	Study design	Setting	Period	Sample size	Inclusion and exclusion criteria of participants	Age	Sex	Etiology /lesion	Severity of apraxia	Length of the disorder	Diagnostic rating	Presence of other communication disorder
Ballard et al., 2019, Australia	SCED	Home and clinic	4 weeks	5	1) Chronic AOS; 2) No prior history of speech; 3) language or reading disorder; 4) Absence of other neurological disorder; 5) Negative histories for alcohol and/or substance abuse; 6) English speakers; 7) No hearing loss; 8) Having iPad; 9) competent to navigate apps	60 to 73 years	4 M 1 F	4 L/CVA 1 R/CVA	2 mild, 2 mild-moderate 1 moderate	More than 6 months	A	Anomic or Broca's aphasia
Bislick, 2020, USA	SCED	Aphasia and Related Conditions Research Laboratory space at UCF	12.5 weeks for P1 and 9 weeks 1 day for P2	2	1) Right handed; 2) use English as a primary language; 3) minimum of high school education; 4) passed an audiometric pure tone, air conduction screening at 35 dB HL at 500, 1000, and 2000 Hz for at least one ear; 5) normal or corrected to normal visual acuity (20/20 to 20/40) as determined by a vision screen (Tumbling E eye chart); 6) score above a 23/36 on the Raven's Colored Progressive Matrices; (7) demonstrate sufficient auditory comprehension to participate in the study; 8) negative medical history of untreated depression or other psychiatric illness, degenerative neurological illnesses, chronic medical illness, or dysarthria.	46 years and 61 years	2 M	2 L/CVA	1 mild-moderate, 1 moderate-severe	134 months- 84 months	A	Aphasia
Bislick et al., 2014, USA	SCED	Home	6 weeks	1	L/CVA	38 years	1 M	1 L/CVA	1 severe	22 months	A	Aphasia
Farias et al., 2014, USA	SCED	NR	4 weeks	1	1) AOS as primary diagnosis; 2) normal hearing and vision; 3) no dysarthria	56 years	1 M	1 L/CVA	1 mild-moderate	6 months	B	Aphasia
Haley et al., 2021, USA	SCED	Home at the kitchen table	36 weeks	1	1) Chronic AOS; 2) normal hearing	60 years	1 F	1 L/CVA	1 moderate-severe	8 or 9 years	B	Non-fluent aphasia

Author, year, county	Study characteristics					Demographic details		Clinical characteristics				
	Study design	Setting	Period	Sample size	Inclusion and exclusion criteria of participants	Age	Sex	Etiology /lesion	Severity of apraxia	Length of the disorder	Diagnostic rating	Presence of other communication disorder
Hurkmans et al, 2015, Netherland	Case series	Rehabilitation center	12 to 20 weeks	5	1) age between 18 and 75; 2) speech problems due to stroke; 3) no language or articulation disorders before stroke; 4) normal or adjusted-to-normal hearing; 5) between 3- and 6-months post-stroke; 6) diagnosis of AOS on the basis of the DIAS (Feiken & Jonkers, 2012); 7) no previous SMTA treatment	47-72 years	4M 1F	5 L/CVA	3 mild 2 severe	3-6 months	B	1 global aphasia 3 Broca's aphasia 1 Wernicke's aphasia
Johnson, 2018, USA	SCED	University clinic	12 weeks	2	1) Individuals who had difficulty producing speech following a stroke or brain injury ; 2) between the ages of 18 and 85; 3) at least 1 year post onset of a left hemisphere stroke or accident; 4) a native English speaker; 5) have adequate hearing and vision.	61 and 55 years	2 M	1 L/CVA 1 bilateral embolic stroke	2 moderate-severe	19 and 28 months	B	Broca's aphasia
Johnson, Lasker, et al., 2018, USA	Case study	Home, via skype	Cycle 1-10 weeks Cycle 2-7 weeks	1	NR	52 years	1 M	1 L/CVA	1 moderate-severe	5 months	B	Broca's aphasia
Johnson, Lott, et al., 2018, USA	SCED	NR	9 weeks	2	Normal hearing and vision	61 and 68 years	2 M	2 L/CVA	1 moderate 1 moderate-severe	91 and 86 months	B	Transcortical motor aphasia Broca's aphasia
Jungblut et al., 2014, Germany	SCED	Institute	25 weeks	3	1) German speaking; 2) 18 months after the incident; 3) no premorbid history of neurological or psychiatric problems; 4) no perceptual hearing impairments and sufficient auditory comprehension to understand the instructions; 5) good capacity regarding concentration and attention and general health condition stable enough for continuous participation during the 6-month treatment period of this research study.	44-53 years	2M 1F	3 L/CVA	1 moderate 2 severe	18 months	A	1 Broca's aphasia, 2 global aphasia
Marangolo et al., 2013, Italy	Group experimental	NR	Speech therapy for 10 days	8	1) Native Italian proficiency; 2) pre-morbid right-handedness (based on the Edinburgh Handedness Questionnaire; Oldfield, 1971); 3) a single left hemispheric stroke at least 6 months prior to the investigation; 4) no acute or chronic neurological symptoms requiring medication.	37-68 years	4M 4F	8 L/CVA	NR	6 months to 6 year 2 months	B	Aphasia

Author, year, county	Study characteristics					Demographic details		Clinical characteristics				
	Study design	Setting	Period	Sample size	Inclusion and exclusion criteria of participants	Age	Sex	Etiology /lesion	Severity of apraxia	Length of the disorder	Diagnostic rating	Presence of other communication disorder
Mauszycki, Nessler, et al., 2016, USA	SCED	Home	~13 weeks	2	1) Passing a pure tone hearing screening at 35 dB at 500, 1000, and 2000 Hz for at least one ear, unaided; 2) performance within normal limits on the test of nonverbal intelligence-3; 3) self-reported negative histories for alcohol and substance abuse and neurological conditions other than stroke, verified through existing medical records.	51 and 53 years	1M 1M	2 L/CVA	2 moderate	76 & 237 months	A	Broca's aphasia
Mauszycki & Wambaugh, 2020 USA	SCED	Participant's home or research laboratory	8 weeks	2	1) Normal hearing; 2) normal performance on the Test of Nonverbal Intelligence–Fourth Edition; 3) negative histories for alcohol or substance abuse, psychological disorders, and neurological conditions; 4) not receiving any other speech therapy services during their participation in this study	73 and 40 years	1 M 1 F	1 L/CVA 1 L/GSW	1 severe 1 moderate	30 and 53 months	A	Broca's aphasia
Mauszycki et al., 2016 USA	SCED	Participant's home or research laboratory	(Varies)	4	1) Normal hearing in one ear aided or unaided; 2) performance within normal limits on the test of nonverbal intelligence–4; 3) negative histories for alcohol or substance abuse, psychological disorders, and neurological conditions; 4) not receiving speech therapy for AOS during research	37 to 57 years	2 M 2 F	3 L/CVA 1 GSW	2 moderate 2 moderate-severe	12-113 months	A	Broca's aphasia
Mozeiko et al., 2019	Case study	University speech and hearing clinic	2 weeks	1	1) A native English speaker with no reported history of language disorder, substance abuse, or neurological or psychological conditions; 2) passing a pure-tone hearing screening at 35 dB in each ear	51 years	1M	1 L/CVA	1 severe	3.6 years	B	Broca's aphasia
Preston & Leaman, 2014, USA	Case study	Clinic room or participant's home	6 weeks	1	NR	59 years	1F	1 L/CVA	1 moderate-severe	14.5 months	B	Residual aphasia
Varley et al., 2016, UK	RCT	Homes	Speech 6 weeks Sham 6 weeks rest 4 weeks	50	1) Adults with chronic AOS (at least 5 months post onset of apraxic stroke); 2) unilateral left hemisphere lesion(s); 3) the absence of neurodegenerative condition; 4) premorbid competence in English; 5) sufficient auditory/visual acuity to interact with a laptop	Speech first group-28 to 91 years, sham first group-36 to 86 years	29 M 21 F	50 L/CVA	Score given out of 20. Speech first group=0–11, sham first group=0–9	At least 5 months	A	Aphasia

Author, year, county	Study characteristics					Demographic details		Clinical characteristics				
	Study design	Setting	Period	Sample size	Inclusion and exclusion criteria of participants	Age	Sex	Etiology /lesion	Severity of apraxia	Length of the disorder	Diagnostic rating	Presence of other communication disorder
Wambaugh et al., 2016, USA	SCED	Participants home or research laboratory	P1, 2-6 weeks, P3, 4-13 weeks	4	1) Chronic AOS and Broca's aphasia; 2) left hemisphere stroke; 3) absence of other neurological disorder; 4) negative histories for alcohol and/or substance abuse; 5) English speakers; 6) no hearing loss	37 to 83 years	2 M 2 F	4 L/CVA	1 mild 3 mild-moderate	17-259 months	A	Broca's aphasia
Wambaugh, Nessl, et al., 2014, USA	SCED	Participant's home, research laboratory, or university clinic	~13 weeks	6	1) Normal hearing at least one ear, aided or unaided; 2) demonstrating performance within normal limits on the test of nonverbal intelligence–4; 3) negative histories for alcohol or substance abuse, psychological disorders, and neurological conditions other than stroke; 4) not receiving any other speech/language therapy during the course of the study	46 to 71 years	5 M 1 F	6 L/CVA	2 moderate 3 moderate-severe 1 severe	28-87months	A	Broca's aphasia, dysarthria
Wambaugh et al., 2013, USA	SCED	Participants home or research laboratory	Intense 4 weeks Traditional 5.5 weeks For all Follow up 4 weeks	4	1) Chronic AOS and Broca's aphasia; 2) single episode of stroke; 3) absence of other neurological disorders and psychological disorders; 4) negative histories for alcohol and/or substance abuse; 5) English speakers; 6) no hearing loss; not receiving speech therapy other than this	34 to 53 years	3 M 1 F	4 L/CVA	1 mild-moderate, 2 moderate 1 moderate-severe	1.5-20 years	A	Broca's aphasia
Wambaugh et al., 2017, USA	SCED + Group experimental	Participant's residence, research laboratory, or university clinic	2 phases P5, P15 10sessions per phase others 20 sessions per phase	20	1) Native speakers of English; 2) normal hearing at least one ear or, when aided; 3) negative histories for alcohol and/ or substance abuse and neurological conditions other than stroke; 4) not receiving other speech therapy during this study; 5) performing within normal limits on the Test of Nonverbal Intelligence–Fourth Edition	29 to 83 years	13 M 7 F	19 CVA 1 GSW	Range from mild to severe	6-259 months	A	Broca's aphasia, anomic aphasia, dysarthria
Wambaugh et al., 2021, USA	SCED	Participant's home, research laboratory, or university clinic	2 phases P5, P15 10 sessions per phase others 20 sessions per phase	20	1) Native speakers of English; 2) normal hearing at least one ear or, when aided; 3) negative histories for alcohol and/ or substance abuse and neurological conditions other than stroke; 4) not receiving other speech therapy during this study; 5) performing within normal limits on the Test of Nonverbal Intelligence–Fourth Edition	29 to 83 years	13 M 7 F	19 CVA 1 GSW	Range from mild severe	6-259 months	A	Broca's aphasia or anomic aphasia

Author, year, county	Study characteristics					Demographic details		Clinical characteristics				
	Study design	Setting	Period	Sample size	Inclusion and exclusion criteria of participants	Age	Sex	Etiology /lesion	Severity of apraxia	Length of the disorder	Diagnostic rating	Presence of other communication disorder
Wambaugh, Wright, Boss, et al., 2018, USA	SCED	Participants home or research laboratory	Intense- 3 weeks No treatment- 2 weeks Traditional - 9 weeks	5	1) Chronic AOS and aphasia; 2) absence of other neurological disorders; 3) negative histories for alcohol and/or substance abuse; 4) English speakers; 5) no hearing loss; not receiving speech therapy other than this	44 to 64 years	5 M	5 L/CVA	2 moderate, 2 moderate-severe, 1 severe	14-228 months	A	Broca's aphasia or conduction aphasia
Wambaugh, Wright, Mauszycki, et al., 2018, USA	SCED	Homes or the VA lab	Ps-1,3,4 40-45 hours P2 – 60-68 hours	4	1) Native English speakers with self-reported negative histories for alcohol or substance abuse and neurological conditions other than the condition that resulted in aphasia; 2) normal hearing for at least one ear, unaided; 3) none of the participants received any other speech/language therapy during this investigation	39 to 69 years	2 M 2 F	3 L/CVA 1 GSW	2 mild-moderate 2 moderate-severe	26-103 months	A	Broca's aphasia
Wambaugh, Wright, et al., 2014, USA	SCED	Homes, a clinic setting, or a social rehabilitation day program	P1-~9 weeks, P2,3,4-~13 weeks	4	1) Native English speakers with self-reported negative histories for alcohol or substance abuse and neurological conditions other than the condition that resulted in aphasia; 2) pass a pure tone hearing screen at least 1 ear unaided	36 to 72 years	4 M	4 L/CVA	1 mild–moderate, 2 moderate, 1 moderate-severe	12-255 months	A	Anomic and Broca's aphasia
Wambaugh et al., 2020, USA	SCED	Participant's residence or research laboratory	Intense 3 week Traditional 9 weeks	12	1) Native speakers of English; 2) negative histories for alcohol/substance abuse and neurological conditions other than stroke; 3) passing a pure tone hearing screening for at least one ear, unaided within normal limits on the Test of Nonverbal Intelligence– Fourth Edition Exclusion criteria: 1) no AOS symptoms or AOS sound errors too infrequent to develop sufficient numbers of treatment stimuli; 2) history of neurological condition other than stroke; 3) psychopathology;4) a score ≥ 10 on the Geriatric Depression Scale–Short Form; 5) treatment for AOS in the 2 months preceding application of treatment in this study; 6) dysarthria as defined by Duffy (2013); 7) premorbid history of speech/language problems as adults;8) participation in any other speech/language therapy during the time of the planned investigation;9) ability/willingness to adhere to the intense and non-intense treatment schedules.	43 to 81 years	8 M 4 F	12 L/CVA	Score of ASRS ranged between 16-30	9-163 months	A	Anomic and Broca's aphasia
Zumbansen et al, 2014, Canada	Case series	NR	6 weeks	3	1) Native French-speaking, right-handed men with aphasia; 2) had experienced a single ischemic unilateral left hemisphere cerebrovascular accident more than 1 year prior to their involvement in the study and had been	48-57 years	3M	3 L/CVA	NR	20-24 months	C	Broca's aphasia

Author, year, county	Study characteristics					Demographic details		Clinical characteristics				
	Study design	Setting	Period	Sample size	Inclusion and exclusion criteria of participants	Age	Sex	Etiology /lesion	Severity of apraxia	Length of the disorder	Diagnostic rating	Presence of other communication disorder
					through the standard public rehabilitation services, which commonly discharge aphasic patients when their language improvements reach a plateau; 3) had not received any speech-language therapy; 4) had not experienced neurological or psychiatric problems before the stroke; 5) an examination by a certified audiologist attested that they had no hearing deficit.							

Data extraction sheet 2: Characteristics of the intervention

Study	Intervention characteristics									
	Type of intervention	Materials & measurement tools used	Procedure	People involved	Mode of delivery	No. of sessions	Schedule	Duration of one session	Duration of follow up	Intensity or dose
Ballard et al., 2019,	Tablet-based word training app with ASR software	iPad with the app installed, computer	Each session involved cycling once through the 20 pictures, following the on screen prompts for supportive cues and requests for naming. Each trial for a given word involved a series of steps/screens including (a) a presentation of the picture with the written question "Do you recognize the picture?" and yes/no response buttons; (b) an opportunity to record a naming attempt by tapping the record icon on the screen and producing a response within the 5 s recording window; (c) a semantic loop to facilitate naming, including a category cue (four response buttons to select the object's category, e.g., fruit vs. three foils) and a use cue (four response buttons to select a common use for the object, e.g., cooking vs. three foils); and (d) a phonological loop, including an orthographic cue (four response buttons to select the object's first letter, e.g., P vs. three foils), a prerecorded auditory cue of the initial consonant and vowel (e.g., "ap" for apple), and a prerecorded whole word model for imitation. After every response selection, the participant is again given the opportunity to record a naming attempt. Hence, for any given word trial, one word production would be recorded if the attempt was judged as correct, or multiple productions would be recorded until a "correct" decision was delivered or the participant selected the "skip" function to move on to the next word.	SLPs	Face-to-face and through an app	80	1 session per day for 4 times a week or in any configuration that suited participants' schedule	1 hour	1 month	Traditional/ non-intense
Bislick, 2020	Phonomotor Treatment	Communicative Participation Item Bank	Stage 1 begins with sounds in isolation, and Stage 2 combines sounds to make more complex productions. Treatment tasks including motor description, speech production, speech perception, and, if appropriate, grapheme to phoneme correspondence are implemented at each difficulty level	SLP, student research assistant	Individual, face-to-face	P1 Phase 1-24, Phase 2-14; P2 Phase 1-22, Phase 2-6	3 days a week for a 1-hour session each day	1 hour	P1-10 weeks P2-8 weeks	NR

Study	Intervention characteristics									
	Type of intervention	Materials & measurement tools used	Procedure	People involved	Mode of delivery	No. of sessions	Schedule	Duration of one session	Duration of follow up	Intensity or dose
Bislick et al., 2014	Phonomotor Treatment	SAQOL-39 ASHA FACS HD camcorder	Practicing sounds in isolation (Stage one), to single-syllable sound combinations (CV, VC, and CVC; Stage two) and beyond to longer syllable combinations (Stage three). Three sounds in isolation were introduced to the participant upon treatment initiation (described in detail in the next section). A new sound was introduced after every 4 to 8 hours of treatment (stimuli selection are discussed below). Each treatment session began with a blocked practice schedule. In other words, the participant was asked to produce sounds in isolation in consecutive blocks (AAAAA, BBBBB, CCCCC). Once the participant was able to successfully produce sounds in isolation, approximately five times consecutively, treatment moved on to random practice.	SLPs, a student with Bachelor's degrees in linguistics	Individual, face-to-face	48	6 weeks, 4 days a week, for 2 hours a day, for 48 hours	1 hour	2 months	NR
Farias et al., 2014	Implicit phoneme manipulation.	Recorder	Encourage the participant to covertly manipulate phonemes to create a new word, and to associate this newly formed word to a target picture among 4 choices.	SLPs	Individual, face-to-face	12	3 times a week for 1–1.5 hours	1-1.5 hours	6 week	NR
Haley et al., 2021	ActionSC treatment.	Tablet computer equipped with a custom app, written instructions for login and navigation, a notepad for study related communication purposes A paper log for tracking practice goals.	3 conversation topics were selected. Record video cues for the 10 phrases that were treated in the first conversation topic about movies were executed within the app by tapping one of the three video cue place holders. LB and the clinician reviewed her practice log and discussed how the practice had unfolded since they last saw each other. Next, LB guided the clinician to challenges and accomplishments by demonstrating phrases and cues she had worked on.	2 clinicians	Face-to-face and self-administering with app	31	2-3 times per day	43 minutes actively using the app on session dates and 24 minutes actively using the app each day she practiced on non-session days.	3 and 4 weeks	NR
Hurkmans et al, 2015	Speech–Music Therapy for Aphasia	Video and audio recorders	Each SMTA session started with warming up of the voice for approximately 2 min. Then, the speech therapy line of treatment (i.e., phoneme, word and sentence level) was followed, depending on the degree of the speech problem and target objectives. Therefore, this line of treatment was variable per participant. In contrast, all participants	SLPs	Individual, face-to-face	24	Two SMTA sessions per week	30 minutes	3 months	NR

Study	Intervention characteristics									
	Type of intervention	Materials & measurement tools used	Procedure	People involved	Mode of delivery	No. of sessions	Schedule	Duration of one session	Duration of follow up	Intensity or dose
			followed the same MT line of treatment (i.e., singing, rhythmical chanting and speaking). Each target item was trained within this structure. The content of the treatment (i.e., level of the speech therapy line of treatment and the use of various musical elements) and the selection of target items were decided by the speech therapist and music therapist and was not standardized by the investigators of this study. Each target item was practiced until the participant was able to produce it fluently without the therapist's help.							
Johnson, 2018	Motor learning guided treatment	11 point multi-dimensional rating scale, modified survey of communicative effectiveness Panasonic V750 video recorder	In Treatment Cycle 1, 15 of the phrases were programed into their system with a single target item stored under a single button. When the participant pressed a specific area on the system, identified by the written target item, a target utterance was 'spoken' aloud by the device for the participant to practice independently at home on non-therapy days. The participants were provided written instructions to follow for the self-controlled practice of the targets at home using the MLG steps (minus augmented feedback). In place of the clinician's modeled productions the instructions were to press the button on the device to hear the target phrase. The number of productions and pause time between productions remained the same. Participants were asked to record the amount of time spent practicing daily on a paper calendar. Additionally, the number of 'hits' per phrase was recorded on the speech generating device.	SLPs	Face-to-face and self-controlled home practice	-	Twice a week	60 minutes	10 months post-treatment	High and low dose
Johnson, Lasker, et al., 2018	Motor learning guided treatment	Speech generating device, 11-point multidimensional rating scale	MLG protocol was used (appendix) Each treatment cycle used three conditions of practice: high dose (clinician trained + self-controlled practice), low dose (clinician trained only) and untreated (untreated). Once the high dose set reached criterion for mastery, training and self-controlled practice of these phrases concluded and the treatment focus turned to training with the low dose set (no self-controlled practice) only. Once the low dose set reached criterion for mastery, training of this set of phrases concluded. Finally, treatment of the untreated phrases began until criterion for mastery was met.	SLP Spouse	Via skype	To reach mastery, Cycle 1, 9, 12 and 7 Cycle 2, 4, 6 and 7 sessions for high dose, low dose and untrained respectively.	Twice a week by therapist, Self-controlled practice varies	Self-controlled practiced on high dose, Cycle 1 , 10-50 minutes Cycle 2, 20-60 minutes	1- & 6-months post cycle 1, 3 months post cycle 2	High and low dose
Johnson, Lott, et al., 2018	Motor learning guided treatment	Panasonic HC V750 video recorder 11 point	Stage 1: Written presentation of the stimuli accompanied by a clinician model. After the model, the participant produced the stimuli followed by a blank screen for 4 seconds. The participant's productions and pause time was	Clinicians	Individual, face-to-face	18	2 times a week for 9 weeks	30 minutes	10 months	NR

Study	Intervention characteristics									
	Type of intervention	Materials & measurement tools used	Procedure	People involved	Mode of delivery	No. of sessions	Schedule	Duration of one session	Duration of follow up	Intensity or dose
		multidimensional rating scale	repeated 3 times. After the third production, the clinician provided a modeled production of the stimuli followed by knowledge of results feedback. This was repeated for all five of the stimuli. Stage 2: the process was the same with the exception of no initial clinician model Stage 3 of the treatment protocol, the process was the same as Stage 2; only the pause time between productions was increased to 10 seconds							
Jungblut et al., 2014, Germany	Rhythmic-melodic voice training	MR scanner and MR-compatible headphones	Procedure is given elsewhere	Singer, percussionist, SLPs	Individual, face-to-face	50	Twice a week	60 minutes	NR	NR
Marangolo et al., 2013	Transcranial direct current stimulation + speech-language therapy	A battery driven Eldith (NeuroConn GmbH, Germany) Programmable Direct Current Stimulator with a pair of surface-soaked sponge electrodes (5 9 7 cm).	Step 1: The clinician auditorily presented the whole stimulus and asked the patient to repeat it. If the patient correctly repeated the stimulus, the clinician would present another stimulus but if he or she made errors the clinician would move on to the next step. Step 2: The clinician auditorily presented the stimulus with a pause between syllables, prolonged the vowel sound, exaggerated the articulatory gestures and asked the patient to do the same. Step 3: As in step 2, the clinician auditorily presented the stimulus, again with a pause between syllables, prolonged the vowel sound, exaggerated the articulatory gestures and asked the patient to do the same. Step 4: The clinician auditorily presented one syllable at a time, prolonged the vowel sound, exaggerated the articulatory gestures and asked the patient to do the same.	SLPs	Individual, face-to-face	10	Every weekday for 10 days	NR	1 week	NR
Mauszycki, Nessler, et al., 2016	Melodic intonation therapy	Audio recorder	Appendix 3 and appendix 4	SLPs	Individual, face-to-face	20 per each item set	3 times a week	45-60 minutes	2, 4 & 6 weeks	NR
Mauszycki & Wambaugh, 2020	Electropalatography treatment and Sound Production Treatment	Audio recorder	Step 1: The speech language pathologist (SLP) provided a verbal-visual model of the word or phrase and requested a repetition. Step 2: The SLP used printed letters or words to identify the speech sound in error, instructed the participant to focus on the target sound, provided another verbal model, and requested a repetition. Step 3: The SLP used integral stimulation "watch me, listen to me, say it with me" to elicit simultaneous production with a maximum of three attempts to elicit a correct production. Step 4: The SLP provided articulatory cueing appropriate for the speech sound production error and then repeated the procedures used in the previous step. Articulatory	SLP	Individual, face-to-face	Phase 1-8, phase 2-8, phase 3-8	3 times a week	50-60 minutes	2,6,10 weeks	NR

Study	Intervention characteristics									
	Type of intervention	Materials & measurement tools used	Procedure	People involved	Mode of delivery	No. of sessions	Schedule	Duration of one session	Duration of follow up	Intensity or dose
			cueing took the form of verbal and visual modeling and instruction concerning place, manner, and/or voicing; cues were specific to the articulatory error produced. Step 5: The SLP presented the next item. EPG The screen displayed side by side real time tongue to palate contact for the clinician and participant. During the treatment steps, the participant was able to observe the screen (i.e., during the clinician's model of the target word or phrase). Participants were provided with knowledge of results							
Mauszycki, Wright, et al., 2016	Articulatory kinematic treatment in conjunction with visual biofeedback via electropalatography	CDC method	Appendix B	SLPs	Individual, face-to-face	Mean of 8 sessions per phase range of 47-91 sessions	3 times a week	35 minutes	2,4, 8 weeks	NR
Mozeiko et al, 2019	Sound Production Treatment	Audio recorder	In the first step of this response-contingent treatment, the SLP says the word and requests a repetition of the target word. If correct, the participant is asked to repeat the word five more times and then the process is repeated with the next item. If incorrect, the SLP presents a minimal pair word. If production of the minimal pair is correct, the SLP then shows the participant the letter representing the target sound and again request a repetition of the word. If correct, the participant is asked to repeat it five times and then moves to the next item. If not, additional cues are provided including integral stimulation is used in which the participant is requested to "watch me, listen to me, and say it with me" and then articulatory placement cues. If, after all cues, the participant still cannot produce the target word, he is moved to the next item.	SLPs	Individual, face-to-face	10	From 9:00 AM-12:45 PM, weekdays, for 2 weeks resulting in a total of 30 treatment hours, after accounting for probes and one 15-min break	3 hour	4 & 10 weeks	Intensive
Preston & Leaman, 2014	Ultrasound visual feedback	Portable Seemore PI 7.5 MHz ultrasound probe	Ultrasound VBFB was paired with verbal articulatory placement instructions and clinician feedback. Six sessions were provided to the treatment of prevocalic rhotics and then for six sessions to postvocalic rhotics	SLP	Individual, face-to-face	12	2 sessions per week for 6 weeks	60 minutes	1 week	NR
Varley et al, 2016	Self-administered computer therapy	Laptop	Naming and repetition accuracy were measured as correct/ incorrect	SLPs	Self-administration	Varied	Once or twice a day for 6	At least 20 minutes	8 weeks	Determined by

Study	Intervention characteristics									
	Type of intervention	Materials & measurement tools used	Procedure	People involved	Mode of delivery	No. of sessions	Schedule	Duration of one session	Duration of follow up	Intensity or dose
							weeks for each phase Determined by participant			participant
Wambaugh et al, 2016	Sound Production Treatment	CDC method	SPT is administered in the form of a response contingent hierarchy and includes the treatment ingredients of modeling/repetition, orthographic cueing, integral stimulation (''watch me, listen to me, say it with me''), articulatory cueing, feedback and repeated practice. SPT-B. For SPT-B, treatment was applied to each of the five words in a sub-set before treatment was applied with the other sub-set. Within sub-set, the words were presented in random order, but with blocking on each word. Specifically, the target word was submitted to the treatment hierarchy; upon a correct response, the same target word was resubmitted to the hierarchy. This process was repeated for a 5–6 minute period. Then, treatment was applied to the next target word for a 5–6 minute period and so on, until all five target words within the sub-set received treatment. The process was then repeated with the second sub-set of items. SPT-R. For SPT-R, grouping by sub-set was not used in presentation of treatment stimuli. All 10 treatment items (five words for each target) were presented in random order during a treatment trial.	SLPs	Individual, face-to-face	P1,P2-20 sessions P3,P4-40 sessions	3 times per week	50-60 minutes	2, 6 and 10 weeks following the second treatment phase	NR
Wambaugh, Nessl, et al., 2014	Sound Production Treatment	Audio recorder	SPT is a response contingent hierarchy that includes the following steps: 1. The SLP provides a verbal model of the target word and requests a repetition. 2. The SLP indicates the printed letter(s) representing the target sound, instructs the participant to attend to this sound, provides another model, and requests a repetition. 3. The SLP says "watch me, listen to me, say it with me" (i.e., integral stimulation) and attempts simultaneous production for a maximum of three times. 4. The SLP provides articulatory placement cues appropriate to the sound production error and then repeats the procedures used in the previous step. 5. The SLP presents the next item.	SLPs	Individual, face-to-face	40	3 times per week	50-60 minutes	2,6,10 weeks	NR

Study	Intervention characteristics										
	Type of intervention	**Materials & measurement tools used**	**Procedure**	**People involved**	**Mode of delivery**	**No. of sessions**	**Schedule**	**Duration of one session**	**Duration of follow up**	**Intensity or dose**
Wambaugh et al., 2013	Sound Production Treatment	NR	Each participant received treatment for four lists of experimental words with the following four permutations of treatment applied sequentially: (a) intense–random, (b) intense–blocked, (c) traditional– random, and (d) traditional–blocked.	SLPs	Individual, face-to-face	Each participant Intensive phase 4 sessions Traditional phase- 16 sessions	Intensive- 4 hours per day for 4 days per week for 1 week. Traditional-1 hour per day for 3 days per week for 5.5 weeks	50-60 minutes		Intensive and traditional/ non-intense
Wambaugh et al., 2017	Sound Production Treatment	NR	Applied using a response contingent hierarchy as follows: 1. The speech language pathologist (SLP) produces a verbal model of the word or phrase and requests a repetition. When monosyllabic words are the treatment targets, sub steps are used for the purposes of contrastive practice upon an incorrect production (e.g., Wambaugh & Mauszycki, 2010; Wambaugh & Nessler, 2004). When multisyllabic words are the target, then contrastive practice is not used and the next step is attempted. 2. The SLP uses printed letters or words to indicate the sound in error, directs the participant to attend to the target sound, provides another verbal model, and requests a repetition. 3. The SLP uses integral stimulation—"watch me, listen to me, say it with me"—and attempts simultaneous production until a correct production is achieved, with a maximum of three attempts. When a phrase is used, the entire phrase is attempted. 4. The SLP provides articulatory cueing appropriate for the sound production error and then repeats the procedures used in the previous step. Only the target word is practiced (not the entire phrase). 5. The SLP presents the next item.	SLPs	Individual, face-to-face	P5, P15-10 sessions per phase others-20 sessions per phase	3 times per week	50–60 minutes	2,6, 10 weeks	NR
Wambaugh et al., 2021	Sound Production Treatment	NR	Applied using a response contingent hierarchy as follows: 1. The speech language pathologist (SLP) produces a verbal model of the word or phrase and requests a repetition. When monosyllabic words are the treatment targets, sub steps are used for the purposes of contrastive practice upon an incorrect production (e.g., Wambaugh & Mauszycki, 2010; Wambaugh & Nessler, 2004). When multisyllabic words are the target, then contrastive practice is not used, and the next step is attempted. 2. The SLP uses	SLPs	Individual, face-to-face	P5, P15 10 sessions per phase others 20 sessions	3 times per week	50–60 minutes	2,6, 10 weeks	Dosage

Study	Intervention characteristics									
	Type of intervention	Materials & measurement tools used	Procedure	People involved	Mode of delivery	No. of sessions	Schedule	Duration of one session	Duration of follow up	Intensity or dose
			printed letters or words to indicate the sound in error, directs the participant to attend to the target sound, provides another verbal model, and requests a repetition. 3. The SLP uses integral stimulation—"watch me, listen to me, say it with me"—and attempts simultaneous production until a correct production is achieved, with a maximum of three attempts. When a phrase is used, the entire phrase is attempted. 4. The SLP provides articulatory cueing appropriate for the sound production error and then repeats the procedures used in the previous step. Only the target word is practiced (not the entire phrase). 5. The SLP presents the next item.			per phase				
Wambaugh, Wright, Boss, et al., 2018	Sound Production Treatment	NR	Step 1: The SLP provided a verbal-visual model of the word or phrase and requested a repetition. When monosyllabic words were the treatment targets, sub steps of Step 1 were used for the purposes of contrastive practice on an incorrect production. If multisyllabic words were the target, then contrastive practice was not used, and the next step was attempted. Step 2: The SLP used printed letters/words to identify the sound in error, instructed the participant to focus on the target sound, provided another verbal model, and requested a repetition. Step 3: The SLP used integral stimulation—"watch me, listen to me, say it with me"—to elicit simultaneous production with a maximum of three attempts to elicit a correct production. Step 4: The SLP provided articulatory cueing appropriate for the sound production error and then repeated the procedures used in the previous step. Articulatory cueing took the form of verbal and visual modeling and instruction concerning the place and/or manner and/or voicing; cues were specific to the articulatory error produced. Step 5: The SLP presented the next item.	SLPs	Individual, face-to-face	Intensive-27 sessions Traditional-27 sessions	SPT-I 3 sessions per day 3 days per week for 3 weeks SPT-T 1 session per day for 3 times per week for 9 weeks	50-60 minutes	2 and 8 week	Intense and traditional/ non-intense
Wambaugh, Wright, Mauszycki, et al., 2018	CAAST	Audio recorder	CAAST treatment protocol with modifications. Modifications include (1) eight picture stimuli were used per session instead of 10 to allow for more time for SPT and for generalization practice; (2) SPT was completed one additional time with each elaborated utterance. In the initial study (Wambaugh et al., 2014), a second application of SPT with each elaborated utterance was conducted only when time permitted (which tended to be infrequently; and (3) one picture was presented a second time each session for the purpose of practicing generalization	SLPs	Individual, face-to-face	P1,3,4-40 P2-60	3 times per week	60-75 minutes	2 & 6weeks	NR

Study	Intervention characteristics									
	Type of intervention	Materials & measurement tools used	Procedure	People involved	Mode of delivery	No. of sessions	Schedule	Duration of one session	Duration of follow up	Intensity or dose
Wambaugh, Wright, et al., 2014	CAAST	NR	Step 1: Therapist-Presents action picture and says, "Tell me anything about this picture; what does it remind you of? what's happening?" Participant-No response. Therapist-"You could say something like man spills ... or ... drops a cup." Participant-"Spill." Step 2: Therapist-"Spill, great." Referring to the sentence frame, asks participant to indicate where to write "spill." Step 3: Therapist-"What does the man spill?" Participant-"Milk." Step 4: Therapist-"Milk, good, spill milk." Referring to the sentence frame, asks the participant where to write "milk." Step 5A: Therapist-"Repeat after me ... spill milk." Participant-"Pill milk." Step 5B: Therapist-"Good try, but not quite correct. Let's concentrate on this sound (underlines the "s" on the sentence frame) and try again ... spill milk." Participant-"Spill milk." Therapist-"That's right. Now, let's say it three more times." Step 6: Therapist-Removes the picture and imposes a 5 s delay (e.g., "Wait and then I'm going to ask you to say it again"). Participant-"Milk." Therapist-"Good try, but not quite. Watch me and try it with me ... spill milk."	SLPs	Individual, face-to-face	P1-28 P2,3,4-40	3 times per week	60-75 minutes	2 and 6 weeks	NR
Wambaugh et al., 2020	Sound Production Treatment	NR	The SPT hierarchy was applied to each treatment item as often as possible in a treatment session. The 30 treatment items were presented in a blocked manner; all 10 items for one sound target were presented (in random order), then all 10 items of another target, and then all 10 items of the remaining target (order of the three sound target groups was randomized). Blocking of sounds continued until at least 70% of the items within each group were produced accurately at Step 1. Then, the remaining trials for the session were alternated between random presentation (all 30 treatment items were presented in a non-predictable order) and blocked presentation. Alternating between blocked and random was continued as long as the 70% criterion was met on the blocked trials. If less than 70% of the items were produced correctly at Step 1, then blocked presentation continued on all trials until this criterion was reached again.	SLPs	Individual, face-to-face	27 per phase	SPT-T =27 sessions over 9 weeks vs. SPT-I = 27 sessions over 3 weeks	50-60 minutes	2 and 8 weeks	Traditional/ non-intense and intense
Zumbansen et al, 2014	Melodic intonation therapy	NR	Participants had to listen and produce 20 phrases, each following a progressive procedure in four steps: two times in unison, two times in unison with therapist fading out at half-way, one time in repetition alone, and finally alone in response to a question. Half of the sentences were New-phrases ranging from two to eight syllables (one phrase of two, three, seven, and eight syllables and two of four, five,	Trained graduate student in SLP	Individual, face-to-face	18	3 one-hour sessions/ week for 6weeks	1 hour	NR	NR

Study	Intervention characteristics									
	Type of intervention	Materials & measurement tools used	Procedure	People involved	Mode of delivery	No. of sessions	Schedule	Duration of one session	Duration of follow up	Intensity or dose
			and six syllables), beginning with the shortest and progressing on to the longest sentences. The other half were Test-phrases repeatedly trained at each session to ultimately assess the direct effect of the treatment. The stimuli were first heard from an iPod connected to speakers and immediately reproduced by the therapist to allow lip-reading. Up to four attempts were allowed in the steps where unison was used. If the participant still failed to produce the phrase successfully, the item was discontinued and the next phrase was presented. When errors occurred at the two last steps, the preceding step was reintroduced before trying again and if this second attempt failed, the item was discontinued.							

Data extraction sheet 3: Conclusion, strengths and limitations and funding source

Study	Conclusion and any other special information	Strengths & limitations of the studies	Funding source
Ballard et al., 2019	All participants showed immediate gain in accuracy of single word production, and a good maintenance for 1 month. 1 person showed less improvements. This approach aids in providing high intensity practice supporting self-administration of speech therapy sessions.	Limitations- Small scale study. Only 124 words were used. Accuracy of ASR may be reduced as dictionary contains more similar words Only single words were tested	National Health and Medical Research Council Grant 630489 and Australian Research Council Future Fellowship awarded to K. J. Ballard (FT120100355) and an Endeavour Postdoctoral Fellowship and Pennsylvania State University faculty funding awarded to N. M. Etter
Bislick, 2020	One showed improvements in trained targets, response generalization and maintenance of treatment outcomes (accuracy of phoneme production) and other one's varied. The modified version of PMT can be used with people with AOS and aphasia.	Limitations- Less generalizability due to small sample size. One participant left the study prior to the end of treatments. A Large number of treatment sessions had been needed to generate restricted treatment effects. Student research assistants were not blinded to the purpose of the study or the study phase.	VPR Advancement of Early Career Researchers Award through the College of Health and Public Affairs at the University of Central Florida
Bislick et al., 2014	Participant showed positive improvements, response generalization and maintenance of treatment effects (accuracy of phoneme production). But results are less generalizable. It is suggested the treatment facilitates speech motor learning in individuals with AOS.	Limitations- Lack of experimental control AB design. Not a multiple baseline design	None
Farias et al., 2014	Participant showed positive effects on trained targets, response generalization and maintenance of treatment outcomes (accuracy in complex consonant blends and simpler consonant blends). Activation of neural areas were observed. Suggest implicit phoneme manipulation approach as an additional method in treating AOS.	Limitations- Lack of a counterbalanced condition. The participant was aware that the treatment tasks were designed to improve his speech.	None
Haley et al., 2021	Participant showed improvements in treated targets (production accuracy of words) was observed. ActionSC treatment can be effective and can be applied in treating people with AOS.	Limitations- This was a preliminary study of the first phase of a program that is still under development. Systematic replication was not conducted. Only one participant was used. Participant's personalities may be the reason for the positive results.	None
Hurkmans et al, 2015	SMTA improved both articulation and communication in daily life. It also reduced the severity of aphasia. SMTA seems an effective treatment to treat AOS.	Limitations- Small sample size The gold standard for treatment research is that the treatment under investigation is compared to a control condition (e.g., no treatment). In the SMTA study, control was included by adding	None

Study	Conclusion and any other special information	Strengths & limitations of the studies	Funding source
		multiple baseline measurements and including related and unrelated control tests.	
Johnson, 2018	Participant showed improvements in retention measures for treated words/ phrases practiced low dose (therapy only) condition and high dose (therapy+ self-controlled practice) but no generalization. Fewer no. of targets resulted in improvements for treated, generalization and maintenance. The study supports the evidence to prove the effectiveness of MLG treatment on AOS.	Limitations- Use of only three baseline measures in both participants	None
Johnson, Lasker, et al., 2018	Participant showed improvements for all trained phrases only with MLG treatment, maintenance 6 mo. post-treatment. Fewer sessions were required to meet the mastery criterion for high dose (therapy+ self-controlled practice) condition than low dose (therapy only) condition and untrained (after training began). The study contributes to the evidence supporting MLG treatment.	NR	None
Johnson, Lott, et al., 2018	Accuracy of word production in treated phrases and maintenance effects were identified through approach. Both qualitative and quantitative measures are recommended to use when measuring improvements from the intervention.	Limitations- Small sample size	None
Jungblut et al., 2014, Germany	While imaging data yielded either no significant activation or right hemisphere activation before therapy, significant activation in perilesional regions was identified for all chronic patients after therapy. Participants with global aphasia indicated brain activation in homologous areas in the right hemisphere too. Speech and language improvements were observed in all cases. The intervention indicates reorganization through brain activation in individuals with AOS and aphasia.	NR	Gebr̈uder Werner Stiftung
Marangolo et al., 2013	Upon bihemispheric tDCS, participants exhibited a significant recovery not only in terms of better accuracy and speed in articulating the treated stimuli but also in other language tasks (picture description, noun and verb naming, word repetition, word reading) which persisted in the follow-up session. Bihemispheric stimulation is a useful tool in treating individuals with AOS.	NR	None
Mauszycki, Nessler, et al., 2016	Both participants had modest gains in articulatory accuracy for treated and generalization stimuli. Further research is recommended to investigate the effect of MIT on speech and language. Further studies are required to examine the efficacy of MIT for individuals with AOS and aphasia.	NR	Department of Veterans Affairs, Rehabilitation Research and Development
Mauszycki & Wambaugh, 2020	Both approaches improve the production accuracy in treated behaviors (accuracy of phonemes), response generalization and maintenance, but the gains are greater for SPT. Further research is needed with lager sample.	Limitations- Small sample size, participant 1 was unable to fully participate in both treatment protocols	None
Mauszycki, Wright, et al., 2016	Positive acquisition of treated items (accuracy of phonemes) by all, 50% showed response generalization and superior maintenance.	NR	None

Study	Conclusion and any other special information	Strengths & limitations of the studies	Funding source
	Results suggest people with AOS may benefit from VBFB+ SPT. Further studies are warranted.		
Mozeiko et al., 2019	Participant showed positive results on acquisition, generalization and maintenance. Findings support the use of massed practice in treatment sessions.	Limitations- Generalization of findings to the larger population of individuals with AOS. The appearance of loss of experimental control given the rising probes prior to treatment of lists 2–4.	None
Preston & Leaman, 2014	Acquisition was observed for both rhotics. Generalization and retention were identified for postvocalic rhotics. Results support the feasibility of using ultrasound VBFB on treating individuals with AOS.	Limitations- Because this is a single case study, determining the specific elements of the treatment program that facilitated improvement is challenging. It is possible that adapting the treatment protocol to transition from a focus on acquisition to a focus on motor learning could help to aid learning earlier in the process. For example, the feedback type and feedback frequency were designed to facilitate acquisition initially (high-frequency feedback with KP when working at the syllable level) with a transition to motor learning at more complex levels (low frequency with KR at the phrase level); however, it is possible that this procedure may have hindered learning and generalization, as fewer practice trials were achieved when high-frequency feedback was provided. The present study used a consultative model in a real clinical situation to implement the visual feedback. Therefore, the treating clinician had only minimal training in ultrasound. The participant in the study had AOS with concomitant residual aphasia. Because of her strong language comprehension, she was able to follow instructions well and to monitor her errors. Thus, ultrasound visual feedback is not necessarily appropriate for all patients with speech problems following a CVA.	Southern Connecticut State University/AAUP research grant.
Varley et al., 2016	With speech program- Improvements and large maintenance in naming and repetition. This computer based self-administered therapy is beneficial in providing intense treatments.	NR	Bupa UK Foundation specialist grant program
Wambaugh et al., 2016	Both SPT- blocked and SPT- random cause in improvements in treated items (accuracy of phonemes or clusters in multisyllabic words), some may benefit more for maintenance with random schedule. Large scale studies are warranted.	Limitations- control of various aspects of the stimuli Small sample size	None
Wambaugh, Nessler, et al., 2014	Both SPT-blocked and SPT-random cause in improvements in treated items and generalization (accuracy in phonemes). Some may benefit through SPT-random.	NR	None
Wambaugh et al., 2013	Results suggest that the similar outcomes are obtained in terms of acquisition, generalization and maintenance of treatment targets (accuracy of words) regardless of the variation of treatment intensity or practice schedules.	NR	None
Wambaugh et al., 2017	Improvements in treated items (production accuracy in phonemes), generalization and maintenance were present for both practice schedules. Gains in treated items and maintenance was greater for SPT random.	NR	None

Study	Conclusion and any other special information	Strengths & limitations of the studies	Funding source
Wambaugh et al., 2021	Changes in the accuracy of articulation needed only few sessions, mastery needed 12-14 sessions.	NR	None
Wambaugh, Wright, Boss, et al., 2018	Improvements in treated items (production accuracy of phonemes) and response generalization were present for all regardless of the intensity. Maintenance effects were greater for traditional. Patients with AOS may benefit more from traditional than intense.	NR	None
Wambaugh, Wright, Mauszycki, et al., 2018	Improvements in speech production for trained items has increased compared with the previous study on CAAST. Further development in CAAST is recommended.	Limitations- Fatigue associated with higher dose frequency may have impacted the findings	None
Wambaugh, Wright, et al., 2014	Gain in speech production varied across participants. Therefore, suggested further development of the CAAST approach.	NR	None
Wambaugh et al., 2020	Improvements in treated items (production accuracy of phonemes) and generalization were present irrespective of the intensity; and majority showed large maintenance effects for both traditional and intense practice. People with AOS may benefit from both intense and traditional practice	Limitations- The finding that only very few investigations were designed to address activity/participation and contextual factors is still pertinent at this time Restriction of the number of initial baseline probe sessions to three probes per participant.	None
Zumbansen et al, 2014	The three interventions (MIT, rhythmic only, and normally spoken treatments) all improved speech accuracy in trained sentences, but the MIT (combination of rhythm and pitch) elicited the strongest generalization effect both to untreated stimuli and connected speech. No significant change was measured in motor-speech agility with either treatment Combination of rhythm and pitch can bring beneficial effect of MIT.	NR	Centre for Research on Brain, Language and Music (CRBLM) and by scholarships to Anna Zumbansen from the Collaborative Research and Training Experience (CREATE) Program in Auditory Cognitive Neuroscience from the Natural Sciences and Engineering Research Council of Canada (NSERC), the Quebec Bio-Imaging Network (QBIN), and the Faculty of Graduate Studies of Université de Montréal

SCED = Single case experimental design; AOS = Apraxia of speech; M = Male; F = Female; L/CVA = Left cerebral vascular accident; R/CVA = Right cerebral vascular accident; P = Participant; DIAS = Diagnostic Instrument of Apraxia of Speech; NR = Not reported; GSW = Gunshot wound; RCT = Randomized controlled trial; ASRS = Apraxia of Speech Rating Scale; ASR = Automatic speech recognition; SLP = Speech-language pathologist; C = Consonant; V=Vowel; ActionSC treatment = Action for Speech and Communication treatment; SMTA = Speech–Music Therapy for Aphasia; MLG treatment=Motor learning guided treatment; EPG = Electropalatography; VBFB = Visual biofeedback; CAAST = Combined Aphasia and Apraxia of Speech Treatment; MIT = Melodic intonation therapy; tDCS = Transcranial direct current stimulation; SML = Speech motor learning; SPT = Sound Production Treatments; B = Blocked; R = Random; I = Intense; T = Traditional; PMT = Phonomotor Treatment