Phonomotor versus Semantic Feature Analysis Treatment for Anomia in 58 Persons with Aphasia: A Randomized Controlled Trial

Diane L. Kendall, PhD
VA RR&D Puget Sound DVA Medical Center, Research Service
University of Washington, Speech and Hearing Sciences
Seattle, WA
University of Pretoria, Pretoria, South Africa

Megan Oelke Moldestad, MS
VA RR&D Puget Sound DVA Medical Center, Research Service
University of Washington, Speech and Hearing Sciences
Seattle, WA

Wesley Allen, MS
VA RR&D Puget Sound DVA Medical Center, Research Service
University of Washington, Speech and Hearing Sciences,
Seattle, WA
Portland State University, Speech and Hearing Sciences,
Portland, OR

Janaki Torrence
University of Washington, Speech and Hearing Sciences,
Seattle, WA

Stephen E. Nadeau, M.D.
Malcom Randall VA Medical Center; Research Service
Department of Neurology
University of Florida College of Medicine
Gainesville, FL

Corresponding Author:
Diane L. Kendall, PhD
Department of Speech and Hearing Sciences
University of Washington
1417 N.E. 42nd St. Box 354875
Seattle, WA 98105-6246
Voice: (206) 685-7400
FAX: (206) 543-1093
E-mail: dkendall@uw.edu

Disclosure: The authors have declared that no competing interests existed at the time of publication. The first and second authors were funded by VA RR&D Merit Review Grant #C6572R.

Abstract

Purpose: The ultimate goal of anomia treatment should be to achieve gains in exemplars trained in the therapy session, as well as generalization to untrained exemplars and contexts. The purpose of this study was to test the efficacy of phonomotor treatment, a treatment focusing on enhancement of phonological sequence knowledge, against semantic feature analysis (SFA), a lexical-semantic therapy that focuses on enhancement of semantic knowledge and is well known and commonly used to treat anomia in aphasia.

Methods: In a between group randomized controlled trial, 58 persons with aphasia characterized by anomia and phonological dysfunction were randomized to receive 56-60 hours of intensively delivered treatment over 6 weeks with testing pre-, post- and three-months post treatment termination.

Results: There was no significant between-group difference on the primary outcome measure (untrained nouns phonologically and semantically unrelated to each treatment) at 3 months post-treatment. Significant within group immediately post-treatment acquisition effects for confrontation naming and response latency were observed for both groups. Treatment-specific generalization effects for confrontation naming were observed for both groups immediately and 3 months post-treatment; a significant decrease in response latency was observed at both time points for the SFA group only. Finally, significant within-group differences on the Comprehensive Aphasia Test – Disability Questionnaire were observed both immediately and 3 months post-treatment for the SFA group, and significant within-group differences on the Functional Outcomes Questionnaire were found for both treatment groups 3 months post-treatment.

Discussion: Our results are consistent with those of prior studies that have shown that SFA treatment and PMT generalize to untrained words that share features (semantic or phonological sequence, respectively) with the training set. However, they show that there is no significant generalization to untrained words that do not share semantic features or phonological sequence features.

Introduction

Aphasia, an acquired disorder of language typically following stroke involving the left cerebral hemisphere, negatively impacts daily life, the ability to return to work, and social relationships. One common and particularly disabling symptom of aphasia is anomia, the inability to retrieve words, which results from underlying impairments of semantic, lexical-semantic, and/or phonologic processes. Rehabilitation of aphasia, measured in terms of learning of trained items, is effective (Robey, 1994) and commonly delivered therapies, which typically focus on semantic attributes of words, have been shown to improve naming performance immediately following treatment. Generalization to untrained words and contexts with these therapies, however, is typically limited to words that are semantically related to those in the training corpus (Wiseburn & Mahoney, 2009; Kiran & Thompson, 2003a; Edmonds & Babb, 2011). An alternative to commonly delivered aphasia therapies, called Phonomotor treatment (PMT), focuses on improving knowledge of individual phonemes and phoneme sequences (i.e. phonological sequence knowledge). Through a series of phase I and phase II trials, we have shown that intensively delivered PMT not only improves confrontation naming performance on trained words but, as predicted by the theory motivating it, achieves generalization and maintenance to naming of untrained words, some aspects of discourse production, and indicators of quality of life (Kendall et al., 2008; Kendall, Oelke, Brookshire, & Nadeau, 2015). Further, PMT has been shown to alter the linguistic network, as evidenced by a decrease in omission errors immediately post-treatment and 3 months later (Minkina et al., 2015) and to improve oral reading of real words and nonwords following treatment (Brookshire, Conway, Hunting Pompon, Oelke, & Kendall, 2014). The long-term goal of this line of research is to establish an effective, viable, and evidence-based treatment program for word retrieval deficits in aphasia.

The purpose of the present study is to test the efficacy of PMT, a treatment focusing on enhancement of phonological sequence knowledge, against semantic feature analysis (SFA), a lexical-semantic therapy that focuses on enhancement of semantic knowledge and is well known and commonly used to treat anomia in aphasia. Because the ultimate goal of aphasia treatment should be to improve daily verbal communication, hence performance with untrained exemplars, our primary outcome measure assesses generalization. Both the treatments used in this study have an intrinsic capacity for generalization to untrained exemplars that share features (Nadeau, 2015).

PMT is motivated by a connectionist model of phonology (Nadeau, 2001; Nadeau, 2012; Nadeau, 2015) that has been extensively detailed (Kendall et al., 2008: Kendall et al. 2015; Kendall & Nadeau, 2016) and will be only briefly summarized here. The version of PMT used in the present study and the secondary outcome measure employed (generalization to untrained but phonologically related words) are precisely the same as those used in Kendall et al. (2015). The theoretical foundation for PMT is as follows: through the systematic training of phonemes (sounds) and phoneme sequences, the neural connectivity supporting phoneme sequence knowledge will be enhanced. For example, if one trains the phoneme sequence corresponding to “must” to criterion, the ability to say bust, dust, gust, just, lust, and trust will be enhanced because all these words share the rhyme features of “must.” Because phoneme articulatory sequences correspond to the word forms of concept representations founded on semantic knowledge, through bidirectional spread of activation between cortical substrates for semantic and phonologic sequence knowledge, generalization from treated phonemes can be expected to improve naming of untrained words and discourse production, both immediately after treatment and continued beyond treatment termination. As in normal language development in children, when adults with anomia due to aphasia are trained in the phonemic sequence building blocks of word representations, they should be able to continue to build vocabulary after termination of treatment.

The lexical/semantic-based treatment to which PMT will be compared in this study is SFA. SFA is a treatment approach that aims to improve lexical retrieval through systematic stimulation of semantic features by elicitation of prompts about individual nouns (e.g. group, description, function, context and personal associations). The hypothesis motivating this treatment is that strengthening connectivity in association cortices encoding semantic knowledge will increase the likelihood that trained and semantically related untrained words can be retrieved. Thus, training the inter-feature connectivity in semantic cortices underlying the concept of dog will enhance the ability to form distributed concept representations of wolves, coyotes, and foxes because these creatures share so many attributes with dogs. A recent evidenced-based systematic review of the effectiveness of SFA (Maddy, Capilouto, & McComas, 2014) showed that SFA is an effective intervention for improving confrontation naming of items trained in therapy for individuals with aphasia (medium to large treatment effect), though limited generalization to untrained items and connected speech were reported in many of the included studies. The limited generalization of SFA reflects the challenge of treating a sufficient number of semantic domains (e.g., tools and animals) and treating an adequate number of items in each domain to achieve broad generalization and translation to daily verbal communication. Enhancing neural connectivity supporting one semantic category does not generalize to a category that does not share semantic attributes.

Naming therapy (simply having participants practice naming objects) theoretically should not generalize because, with the exception of onomatopoeic words and derivational forms, there is no relationship between word meaning and word sound. Because semantic knowledge and phonological sequence knowledge share no common features, there is essentially no opportunity for the semantic–phonological sequence knowledge network to acquire implicit knowledge of regularities in the relationship between word meaning and word sound through the course of language acquisition (in dramatic contrast to the domains of semantic knowledge and phonological sequence knowledge). In short, if you have learned to name 40 objects, this knowledge provides no help in naming a heretofore unseen 41st object. Empirical studies bear this out (Wisenburn & Mahoney, 2009).

Intrinsic generalization, the mechanism we sought to engage in this study, is based upon the extent to which untrained exemplars of any type share features with trained exemplars. With PMT, there is a potential for generalization to untrained exemplars that share phonological sequence elements with trained sequences. With SFA, there is the potential for generalization to untrained exemplars that share semantic features with trained entities. To fully understand intrinsic generalization, one needs to understand both the structure of the knowledge domain in question and what qualifies as a shared feature. In the phonological domain, words like “must”, “trust”, and “bust” share the same sequence domain by virtue of their shared rhyme segment. However, it remains to be determined whether some generalization might occur between sequences sharing smaller elements, e.g., the final consonant cluster “st”, both within the “ust” domain and in other sequence domains (e.g., blast, best, fist, most, and first). Within the domain of semantics, the sharing of features between dogs and wolves is extensive and obvious, but is there enough sharing between dogs and more atypical animal exemplars such as platypuses, sharks, or squids, to achieve generalization effects?

At the time that this study was initiated, the phonological sequence landscape was poorly understood and the extent to which generalization could occur was unknown. We had only our own published studies and the modest effect sizes we had observed could have been related to many treatment parameters other than generalization constrained by sequence domain and segment length. On the other hand, there exists considerable evidence that capacity for generalization with SFA is limited, as theoretically predicted, to untrained exemplars that share many semantic features (e.g., body parts, food, and clothing), but the role of feature distance, though less well understood, is likely to be important, e.g., the relationship between finger and thumb is likely to be stronger than between finger and lung. Thus, for both phonologic sequence knowledge and semantic knowledge, much remains to be learned about the structure of the knowledge domains and what qualifies as a shared feature for treatment purposes.

All participants in this study had to have both anomia and phonological dysfunction: anomia because improvement of ability to produce major lexical items on confrontation naming was the goal of the study; and phonological impairment because PMT seeks to improve anomia by rehabilitating lost phonological sequence knowledge. The combination of anomia and phonological impairment is by far the most common pattern in post-stroke aphasia because aphasia due to stroke is most often caused by middle cerebral artery (MCA) occlusion or putamenal hemorrhage. With MCA occlusion, ischemia is maximal in the insula and in perisylvian cortex, which is the substrate for phonological sequence knowledge (Nadeau, 2001; Nadeau & Crosson, 1997). With putamenal hemorrhage, perisylvian cortex and its juxtacortical white matter are damaged by compressive and hemato-toxic mechanisms. All participants in this study have lesions in the perisylvian region and are thus representative of the general population of people with stroke-related aphasia.

The primary aim of this study is to determine if PMT or SFA is more effective for improving word retrieval deficits in individuals with aphasia. The following research questions were asked:

Primary Question:

1. Is there a significant *between* group difference in confrontation naming accuracy and response latency, measured 3 months post treatment, for nouns that are *untrained as well as semantically and phonologically unrelated to trained stimuli*?

Secondary Questions:

2. <u>Acquisition</u>: Is there a significant *within* group difference in confrontation naming accuracy and response latency for *trained* nouns measured at treatment completion?
3. <u>Lexical-Semantic Generalization:</u> Is there a significant *within* group difference, immediately post treatment termination and 3 months later, for confrontation naming

of nouns that are both *untrained as well as semantically and phonologically related to trained stimuli*?

4. Ecological validity: Do these treatments achieve a significant *between* or *within* group difference in two measures of ecological validity immediately and 3 months post treatment?

Methods

Study Design

The design of this study was a between group randomized controlled trial with repeated testing. Following randomization, all participants received testing one week prior to treatment (A1), immediately following treatment (A2), and again three months following the termination of treatment (A3). Standardized assessments and outcome measures were administered at all assessment periods (A1, A2 and A3). Individual, impairment-based speech therapy outside of the study was not permitted while participants were enrolled but participation in group communication therapy was allowed.

Participants

Participants were recruited through the Puget Sound Veterans Affairs Healthcare System (Seattle and American Lake) and the University of Washington/Portland State University Northwest Aphasia Registry and Repository, as well as area speech-language pathology clinics. Ninety-six participants were screened and twenty-one failed the screen due to prior neurological events. Fifteen individuals passed the screen and were deemed eligible but declined to participate. A total of 60 participants were recruited and enrolled and 58 completed the treatment protocol. One enrolled individual completed the treatment protocol and A1-A2 testing but did not participate in A3 testing for personal reasons. Two individuals were withdrawn from the study after consent (one following discovery of pre-morbid head trauma and a second following a personal emergency). This study was conducted under the auspices of the

University of Washington Institutional Review Board and all participants personally provided informed consent to take part in the study.

All participants exhibited chronic aphasia (>6 months post-onset) due to predominantly left-hemisphere damage due to stroke. Using computed tomography (CT) or magnetic resonance imaging (MRI) scans, lesion type, locus, and extent of brain damage were characterized and interpreted to determine study eligibility. Neurological damage was characterized using a system like that used in the Locomotor Experience Applied Post-Stroke trial (LEAPS; Nadeau, Dobkin, Wu, Pei, & Duncan, 2016). The LEAPS assessment was completed by a board certified neurologist (last author of this paper). Lacunar infarcts and mild to moderate leukoaraiosis in the nondominant hemisphere were not exclusionary because their impact on cognitive function would be minimal, even as excluding all patients with these findings would have seriously compromised recruitment. There were two patients who were exceptional in terms of their lesions: one had a crossed aphasia and the other had an old infarct in the right calcarine cortex.

Study inclusion required that participants demonstrate anomia, determined via performance on the Comprehensive Aphasia Test (CAT; Swinburn, Porter, & Howard, 2004), and sufficient auditory comprehension to follow basic directions. Inclusion also required the presence of phonological dysfunction, as defined by the Standardized Assessment of Phonology in Aphasia (SAPA; Kendall et al., 2010). Participants with mild to moderate apraxia of speech were included (see details below). Individuals were excluded from study participation if they exhibited untreated depression, degenerative neurological disease, chronic medical illness that would be disrupted by the rigorous study schedule (i.e., chronic kidney disease requiring dialysis), and/or severe, uncorrected impairment of vision or hearing.

Apraxia of speech was determined perceptually using data gathered during the evaluation. Two speech-language pathologists evaluated speech/language behaviors using video-taped data from, but not limited to, picture description, spontaneous conversation, automatic speech, repetition of words of increasing length, and multiple repetition of 3-syllable words. Participants were examined for the following behaviors: slow speech rate (lengthened sounds and/or syllable or word separations), sound distortions (including distorted substitutions), and prosodic abnormalities (Duffy, 2013; Wambaugh, Duffy, McNeil, Robin, & Rogers, 2006). Severe apraxia of speech was defined as a limited repertoire of speech sounds, speech limited to a few meaningful utterances, automatic speech not better than volitional speech, and inability to repeat isolated phonemes.

Fifty-eight total participants completed the treatment protocol. The 30 participants who completed SFA were, on average, 63.4 years of age (*SD*=12.3), had 15.2 years of education (*SD*=2.8), and were 4.1 years post-stroke onset (*SD*=4.7); 18 were males and 12 were females (see Table 1). The 28 participants who completed PMT were, on average, 63.3 years of age (*SD*=10.6), had 14.3 years of education (*SD*=2.0), and were 4.3 years post-stroke onset (*SD*=4.7); 15 were males and 13 were females (see Table 2). Fifty-six individuals were monolingual English speakers, and two were multi-lingual. During screening procedures, multi-lingual speakers were asked about English usage prior to stroke onset, and participants were excluded if English was not spoken fluently prior to that point in time.

Randomization

Participants were randomized in pairs to one of two treatment conditions at the beginning of each new testing cycle, so that one participant received SFA and one received PMT. To control for aphasia severity within each treatment arm of the trial, each participant received a severity rating prior to randomization. Aphasia severity was determined using a binary system of "more" or "less" impaired based on performance on subtests of the CAT (Subtests #7, #9, #11– comprehension of spoken words/sentences/paragraphs; Subtests #12, #13, #16 – repetition of words/complex words/sentences). Individuals whose scores fell predominantly below a standard score of 50 were deemed more impaired, whereas individuals whose scores fell predominantly above a standard score of 50 were deemed less impaired. In

Table 1. Phonomotor treatment participant characteristics.

Participant	Age (years)	Sex	Education level (years)	Duration postonset (years)	Handedness	*Executive Function* Ravens (out of 36)	*Lexical Retrieval* BNT (out of 60)	*Semantic Processing* CAT Memory *t* score	*Auditory Comprehension* CAT Comprehension of Spoken Language *t* score	*Reading Comprehension* CAT Comprehension of Written Language *t* score	*Writing* CAT Production of Written Language *t* score	*Phonological Processing* SAPA (out of 144)	*Verbal Short-Term Memory* TALSA Digit Span (out of 7)	*Verbal Short-Term Memory* TALSA Word Span (out of 7)
1	71	F	15	1.42	R	35	49	62	59	66	69	116	3.2	3
2	46	F	16	1.25	R	28	23	54	58	62	69	79	4.05	4.05
3	59	F	16	5.25	R	33	42	62	50	54	56	110	4	3.1
4	67	M	12	2.75	R	29	46	62	60	59	59	94	4.15	3.1
5	70	F	12	6.75	R	23	6	38	39	48	44	61	1.05	1.05
6	40	F	16	1.83	R	36	5	62	51	48	52	73	3	2
7	59	M	18	2.25	R	32	6	54	54	50	55	58	2.2	2.05
8	71	F	14	0.83	R	31	36	62	58	60	52	111	2.15	2
9	65	M	16	8.42	R	34	4	62	44	49	53	62	2.1	3.05
10	73	M	16	2.33	R	28	6	54	49	44	43	49	1.1	2
11	73	F	13	2.5	R	34	12	50	50	54	53	100	3.05	2
12	67	M	16	4.17	R	32	20	54	53	56	57	84	3.15	2.2
13	46	M	13	7.08	L	*	14	50	45	44	41	48	1.1	1.15
14	59	F	12	3.67	R	26	0	16	35	40	48	30	*	*
15	71	F	16	3.67	R	34	41	54	52	66	62	109	3.15	3.1
16	90	F	12	6.42	L	29	26	54	58	59	60	79	4.2	3.05
17	63	M	13	4	R	35	32	62	45	56	50	77	3	2
18	60	M	14	2.75	L	19	2	62	46	50	46	50	2.05	2
19	46	M	11	24.83	R	35	3	62	41	37	51	41	*	*
20	74	F	14	0.92	R	25	3	62	44	48	44	53	*	1
21	63	M	16	2	L	32	2	62	50	42	45	35	*	1.15
22	62	M	16	0.83	R	35	46	54	58	55	59	105	3.15	3.1
23	50	M	12	2.08	R	29	1	62	52	51	48	79	*	*
24	67	F	16	1.67	R	32	39	62	45	54	51	93	1.2	2.05
25	70	M	12	4.42	R	21	32	54	56	49	50	55	2.15	2
26	66	F	12	9.42	R	22	40	62	58	54	*	94	3.1	2.2
27	65	M	14	1.42	L	22	52	62	55	54	49	88	3.1	2.15
28	59	M	18	4.17	L/R	33	16	50	47	54	48	60	3.05	1.2
AVE	63.3		14.3	4.3		29.8	21.6	55.9	50.4	52.3	52.4	74.8	2.7	2.2
SD	10.6		2.0	4.7		5.0	17.8	9.8	6.6	7.2	7.3	24.9	1.0	0.8

Note. Asterisk (*) indicates missing data. BNT = Boston Naming Test; CAT = Comprehensive Aphasia Test; SAPA = Standardized Assessment of Phonology in Aphasia; TALSA = Temple Assessment of Language and Short-Term Memory in Aphasia; AVE = average; F = female; M = male; R = right; L = left.

Table 2. Semantic feature analysis participant characteristics.

Participant	Age (years)	Sex	Education level (years)	Duration postonset (years)	Handedness	*Executive Function* Ravens (out of 36)	*Lexical Retrieval* BNT (out of 60)	*Semantic Processing* CAT Memory *t* score	*Auditory Comprehension* CAT Comprehension of Spoken Language *t* score	*Reading Comprehension* CAT Comprehension of Written Language *t* score	*Writing* CAT Production of Written Language *t* score	*Phonological Processing* SAPA (out of 144)	*Verbal Short-Term Memory* TALSA Digit Span (out of 7)	*Verbal Short-Term Memory* TALSA Word Span (out of 7)
29	72	M	20	2	R	27	11	62	45	46	47	59	1.05	2
30	69	M	19	8.5	L	35	28	62	50	54	51	61	2.2	1.2
31	38	F	12	3.42	R	36	38	62	48	49	57	100	3	3.05
32	72	F	16	2.25	R	35	46	54	56	62	57	106	4.15	3.1
33	57	M	10	0.92	R	25	4	54	47	44	46	49	2	1.15
34	44	F	16	0.75	R	36	50	62	57	60	55	112	4.1	3.1
35	45	M	14	1	R	34	7	41	52	56	52	73	3.15	2.05
36	91	F	18	0.67	R	*	3	62	39	49	44	32	1.05	1.15
37	69	M	13	10.33	R	32	30	54	56	52	50	74	4	3.05
38	63	M	12	3.17	L	28	17	54	58	56	53	70	3	2.2
39	70	F	13	3.58	R	34	2	39	43	45	47	44	*	1.2
40	59	F	16	9.25	R	35	14	39	45	49	50	90	2.15	2.1
41	65	M	13	0.83	R	25	50	54	58	54	59	92	5.1	3.05
42	56	F	12	2.08	R	27	46	50	50	52	48	80	2	2
43	77	F	16	14.75	R	35	11	62	50	57	59	69	2.05	1.2
44	74	M	16	2.75	R	24	1	41	38	38	44	36	2	1.05
45	64	M	19	2.17	R	28	38	50	52	58	50	84	2.1	1.2
46	55	M	15	1.58	R	35	54	50	63	63	60	118	6.05	4.2
47	75	F	12	14.08	R	22	1	50	52	56	52	44	3	2
48	77	M	18	11	R	34	1	54	47	51	49	53	*	2.1
49	55	M	16	1.5	R	32	47	62	57	60	62	111	*	*
50	66	M	16	1	R	33	52	62	56	57	55	99	5.05	4
51	68	F	18	1.83	R	22	11	16	41	45	48	58	1.1	1.05
52	62	M	18	6.17	R	25	17	16	39	50	49	100	1.05	1
53	58	M	12	1.83	R	*	28	54	56	50	58	87	*	*
54	44	F	16	1.58	L	31	0	47	39	45	47	44	*	1.05
55	79	F	14	9.25	R	19	47	35	55	59	52	96	4.1	3.15
56	75	M	20	0.5	R	31	36	54	46	52	62	102	3.2	2.15
57	47	M	11	1.83	R	36	7	62	48	54	54	69	4.1	2.15
58	55	M	16	1.17	R	34	52	62	65	65	69	117	6.2	5
AVE	63.4		15.2	4.1		30.4	25.0	50.9	50.3	52.9	52.9	77.6	3.1	2.2
SD	12.3		2.8	4.2		5.1	19.6	12.5	7.3	6.3	6.0	25.6	1.5	1.1

Note. Asterisk (*) indicates missing data. BNT = Boston Naming Test; CAT = Comprehensive Aphasia Test; SAPA = Standardized Assessment of Phonology in Aphasia; TALSA = Temple Assessment of Language and Short-Term Memory in Aphasia; AVE = average; M = male; F = female; R = right; L = left.

rare cases in which severity was ambiguous (i.e., half of subtests scores above 50, half below 50), the administering speech-language pathologist shared clinical judgment and scores from other standardized measures with the research team to collectively arrive at a severity rating. Once a severity rating was determined for each participant, the participant ID number and severity rating were entered into a program that used a modified version of Pocock's (1983, 1975) minimization method for randomization. In the current study, a paired, rather than an unpaired, version was used. Once individual participant information was entered, an imbalance score was computed for each of the two possible randomization assignments (SFA/PMT or PMT/SFA). Each treatment group was closely balanced in its constitution of "more" or "less" severe participants. It should be noted that while the CAT subtests used for randomization typically produce continuous variables, these were converted to categorical variables in order to use the paired version of Pocock's minimization randomization method. However, as can be seen in Tables 1 and 2, the SFA and PMT groups were also balanced when CAT subtest scores were reported in their standardized, continuous variable form.

Treatment administration

All participants received 56-60 hours of treatment in a massed treatment schedule. Therapy was administered by a licensed and certified research speech-language pathologist (second and third authors). Therapy was delivered for a total of 8-10 hours/week over the course of six to seven weeks. Each participant was seen for approximately two hours of therapy per day (two 45-50 minute sessions with a 10-minute break between sessions). Participants were seen at the site most convenient for them, which was either their home, the university, or another quiet location. The two research speech-language pathologists received training on the treatment protocol by the first author. Training involved direct observation, supervised treatment administration, and daily review of therapy tasks and procedures for the first 60 hours of treatment. Ongoing training was conducted through discussions during weekly lab meetings with the two research speech-language pathologists and first author.

Treatment fidelity was monitored by graduate students who evaluated ten-minute, randomly selected audio samples that were recorded in 20% of the therapy sessions (i.e., one day each week). The evaluator analyzed the recordings using a pre-determined checklist of essential therapy components. There were six key therapy components identified for PMT and nine key therapy components for SFA. When the speech-language pathologist who was delivering treatment delivered the key component, the evaluator gave it a score of 1. If the element was not present in the ten-minute sample, it was given a score of 0. Fidelity was calculated by averaging the percent of treatment elements that were present in each sample. The average treatment fidelity across weeks and participants was 96.75% for PMT and 99.51% for SFA.

Treatment procedures

This study investigated the effects of two intensively delivered treatments: SFA (Boyle, 2004; Boyle & Coelho, 1995), a lexical-semantic based treatment, and PMT (Kendall et al., 2008; Kendall et al., 2015), a phoneme-based treatment. Both treatments are described in detail in the Appendices (Appendix 1 and Appendix 2, respectively).

Treatment stimuli.

Phonomotor treatment stimuli.

Treatment stimuli were created using the same methods as those described in previous studies (Brookshire et al., 2014; Kendall et al., 2015; see Table 3). Thirty-nine real words and 72 nonwords were trained. These were phonotactically-legal one and two-syllable words of low phonotactic probability (PP) and high neighborhood density (ND), as determined by methods similar to those outlined in Vitevitch and Luce (1999). The rationale for using words of low PP and high ND derives from work done by Storkel, Armbruster, and Hogan (2006), in which greater learning occurred with words characterized by these phonological sequence properties.

Semantic feature analysis treatment stimuli.

Table 3. Phonomotor treatment stimuli.

Sounds in isolation		Real words				Nonwords			
Trained		Trained		Untrained		Trained		Untrained	
IPA	Graphemes	1-Syllable	2-Syllable	1-Syllable	2-Syllable	1-Syllable	2-Syllable	1-Syllable	2-Syllable
p	p	ape	clover	beef	baby	doi (dɔɪ)	chootee (tʃuti)	ain (eɪn)	wurkee (wɝki)
b	b	bird	diver	boot	iron	af (æf)	zhuree (ʒɝi)	poom (pum)	koetoe (koutou)
f	f	bride	father	bow	jury	toos (tus)	foekoe (foukou)	gee (gi)	wayzer (weɪzɚ)
v	v	bruise	genie	eel	ladder	sheev (ʃiv)	leber (lɛbɚ)	haje (heɪdʒ)	rooit (ruɪt)
t	t	cave	gravy	fig	lasso	ek (ɛk)	doem (doʊʌm)	loy (lɔɪ)	sayvay (seɪveɪ)
d	d	ditch	halo	hay	leather	dach (dætʃ)	mefoe (mɛfou)	heeg (hig)	fooer (fuɚ)
k	k	fire	heater	thigh	razor	peenz (pinz)	shever (ʃɛvɚ)	jong (dʒɑŋ)	laybee (leɪbi)
g	g	fur	ivy	tire	shadow	poa (poʊə)	feether (fiðɚ)	poy (pɔɪ)	grayzee (greɪzi)
θ	th	jail	jockey	toy	turkey	meeth (miθ)	toiler (tɔɪlɚ)	awb (ɑb)	ekee (ɛki)
ð	th	jeans	level	whip	valet	ri (rɪ)	izel (ɑɪzl)	jeef (dʒif)	badow (bædou)
s	s	knee	meadow	wire		ish (ɪʃ)	shaybee (ʃeɪbi)	tay (teɪ)	nider (nɑɪdɚ)
z	z	knob	movie			whup (wʌp)	veeder (vidɚ)	mirth (mɝθ)	eepoe (ipou)
ʃ	sh	knot	polo			breek (brik)	zower (zaʊɚ)	vank (vænk)	vaylow (veɪlou)
ʒ	zh	maze	ranger			voo (vu)	tawthee (tɑθi)	bap (bæp)	sheefur (ʃifɚ)
tʃ	ch	mop	shoulder			eep (ip)	jiver (dʒɪvɚ)	ka (kæ)	hoower (huwɚ)
dʒ	j	owl	shower			reesh (riʃ)	wooter (wutɚ)	ool (ul)	eeshur (iʃɚ)
l	l	pie	speaker			nie (nɑɪ)	dungee (dʌŋi)	wog (wɑg)	rayger (reɪgɚ)
r	r	plane	teacher			iej (ɑɪdʒ)	turmee (tɝmi)	glane (gleɪn)	zopper (zɑpɚ)
h	h	wheel	tiger			zine (zɑɪn)	lekzher (lɛkʒɚ)	ieg (ɑɪg)	joah (dʒoʊə)
w	w	witch				broiz (brɔɪz)	lekee (lɛki)	dite (dɑɪt)	tawkee (tɑki)
wh	wh					thag (θæg)	juroe (dʒɝo)	grabe (greɪb)	zire (zɑɪər)
m	m					oit (ɔɪt)	shashoe (ʃæsou)	jie (dʒɑɪ)	thiver (θɪvɚ)
n	n					kur (kɝ)	hoyter (hoɪtɚ)	wawj (wɑdʒ)	wiver (wɑɪvɚ)
ŋ	ng					froos (frus)	neenee (nini)	fie (fɑɪ)	uzher (ʌʒɚ)
i	ee					grake (greɪk)	rayzel (reɪzl)	oozh (uʒ)	chafter (tʃæftɚ)
ɪ	i					choy (tʃɔɪ)	highger (hɑɪgɚ)	whike (wɑɪk)	osay (oseɪ)
ɛ	e					oos (us)	woewuh (woʊwə)	gride (grɑɪd)	doojee (dudʒi)
eɪ	ae					wap (wæp)	unger (ʌngɚ)	loich (bɪtʃ)	fayshur (feɪʃɚ)
æ	a					faps (fæps)	miver (mɑɪvɚ)	moy (mɔɪ)	shiloe (ʃɪlo)
ʌ, ə	u					woy (wɔɪ)	jawvee (dʒɑvi)	jurl (dʒɝl)	voker (voukɚ)
ɑ, ɔ	o, aw					awch (ɑtʃ)	prezhur (prɛʒɚ)	thed (θɛd)	haybee (heɪbi)
o, oʊ	oe					plown (plaʊn)	foover (fʊvɚ)	eem (im)	rieger (rɑɪgɚ)
ʊ	oo					zae (zeɪ)	pire (pɑɪɚ)	riz (rɪz)	layfee (leɪfi)
u	oo					hob (hɑb)	dryper (drɑɪpɚ)		meevee (mivi)
ɑɪ	ie					veed (vid)	gower (gaʊɚ)		tycher (tɑɪtʃɚ)
ju	ue						teever (tivɚ)		kloper (kloupɚ)
ɔɪ	oi, oy						ibee (ɑɪbi)		nyer (nɑɪɚ)
aʊ	ow, ou								langee (leɪŋi)
ɝ, ɚ	er, ir, ur								gainjer (geɪndʒɚ)
ɔr	or								skonner (skɑnɚ)
ɑr	ar								

Note. IPA = International Phonetic Alphabet.

Table 4. Semantic feature analysis stimuli.

Lo								
Category	Body Parts	Clothing and Accessories	Food and Beverages	Household	Hobbies, Recreation/ Sports	Nature	Occupations	Transportation
Targets	trachea	bowtie	avocado	armoire	rafting	spiderweb	umpire	tugboat
	bellybutton	cardigan	pineapple	bookshelf	origami	geyser	veterinarian	rickshaw
	toenail	kilt	cinnamon	quilt	archery	sunflower	miner	boxcar
	pinkie	beret	macaroni	mixer	croquet	beehive	pianist	blimp
	calves	mittens	lime	silverware	fencing	petal	ballerina	rowboat
	heel	scarf	oatmeal	mattress	skiing	volcano	mechanic	canoe
	elbow	slippers	gravy	stove	photography	avalanche	musician	skates
	cheek	vest	lemonade	candle	yoga	autumn	rabbi	tractor
	forehead	sleeve	tomato	cabinet	camping	seeds	nun	ferry
	ankle	jeans	garlic	ceiling	chess	pond	magician	submarine

Hi								
Category	Body Parts	Clothing and Accessories	Food and Beverages	Household	Hobbies, Recreation/ Sports	Nature	Occupations	Transportation
Targets	lungs	buttons	pudding	refrigerator	wrestling	rainbow	photographer	carriage
	chin	necklace	cereal	oven	bowling	hurricane	farmer	limo
	thumb	skirt	potatoes	bench	swimming	clouds	dentist	motorcycle
	toes	pockets	corn	rug	hockey	lawn	policeman	jeep
	palm	leather	rice	pillow	soccer	waves	chef	subway
	bone	tie	apple	plate	fishing	snow	pilot	elevator
	tongue	belt	soup	closet	painting	desert	priest	tank
	lips	shoe	juice	gate	baseball	coast	actor	bike
	ear	glasses	bread	roof	golf	trees	artist	taxi
	fingers	jacket	pizza	bedroom	football	mountain	nurse	traffic

Body Parts	Clothing and Accessories	Food and Beverages	Household	Hobbies, Recreation/ Sports	Nature	Occupations	Transportation
Trained stimuli (high frequency)							
lungs	buttons	pudding	refrigerator	wrestling	rainbow	photographer	carriage
chin	necklace	cereal	oven	bowling	hurricane	farmer	limo
thumb	skirt	potatoes	bench	swimming	clouds	dentist	motorcycle
toes	pockets	corn	rug	hockey	lawn	policeman	jeep
palm	leather	rice	pillow	soccer	waves	chef	subway
bone	tie	apple	plate	fishing	snow	pilot	elevator
tongue	belt	soup	closet	painting	desert	priest	tank
lips	shoe	juice	gate	baseball	coast	actor	bike
ear	glasses	bread	roof	golf	trees	artist	taxi
fingers	jacket	pizza	bedroom	football	mountain	nurse	traffic
Trained stimuli (low frequency)							
trachea	bowtie	avocado	armoire	rafting	spiderweb	umpire	tugboat
bellybutton	cardigan	pineapple	bookshelf	origami	geyser	veterinarian	rickshaw
toenail	kilt	cinnamon	quilt	archery	sunflower	miner	boxcar
pinkie	beret	macaroni	mixer	croquet	beehive	pianist	blimp
calves	mittens	lime	silverware	fencing	petal	ballerina	rowboat
heel	scarf	oatmeal	mattress	skiing	volcano	mechanic	canoe
elbow	slippers	gravy	stove	photography	avalanche	musician	skates
cheek	vest	lemonade	candle	yoga	autumn	rabbi	tractor
forehead	sleeve	tomato	cabinet	camping	seeds	nun	ferry
ankle	jeans	garlic	ceiling	chess	pond	magician	submarine
Untrained stimuli (high frequency)							
muscle	sweater	beans	blanket	tennis	cave	dancer	airplane
tooth	gloves	dough	lamp	shopping	mud	surgeon	rocket
knee	boots	steak	basket	hunting	grass	cowboy	jet
skull	socks	salad	furniture	reading	trunk	waiter	helicopter
shoulder	purse	cookies	newspaper	writing	rocks	maid	ambulance
Untrained stimuli (low frequency)							
pelvis	swimsuit	lettuce	dresser	knitting	pollen	fisherman	escalator
pupil	apron	brownie	vase	gymnastics	seaweed	plumber	skis
vein	zipper	biscuit	fireplace	crossword	waterfall	mailman	trolley
waist	tuxedo	grapes	crib	pottery	puddle	astronaut	skateboard
spine	sneakers	onion	bathtub	sewing	orchid	runner	scooter

Treatment stimuli consisted of 80 highly imageable nouns from eight semantic categories (see Table 4). Participants were given either a high- or low-frequency word set based on naming performance on standardized tests and the POM screen(see below under primary outcome measure). When the appropriate word set was not easily determined, a larger corpus of 240 nouns, taken from both frequency sets, was administered to determine whether participants would receive training on the high- or low-frequency treatment set of words. Of the nouns in each semantic category, ten from each frequency set were chosen for training, and an additional five items were used to assess generalization to untrained responses.

Outcome measure description

Primary outcome measure

The primary outcome measure addresses Research question #1 testing generalization and maintenance: Is there a significant between group difference in confrontation naming accuracy and response latency of untrained nouns that are linguistically unrelated to either treatment measured 3-months post treatment termination?

Stimuli for POM were comprised of untrained nouns taken from the Philadelphia Naming Test (PNT; Roach, Schwartz, Martin, Grewal, & Brecher, 1996) and the Object and Action Naming Battery (O&A; Druks & Masterson, 2000). A corpus of 50 nouns was chosen from the two tests. To mitigate effects of generalization within semantic category (for SFA) or within phonological sequence domain (for PMT), the nouns used in POM did <u>not</u> belong to one of the eight trained SFA categories used in treatment and did <u>not</u> include the phonological sequence domains trained in PMT.

Although the PNT and O&A stimuli were balanced according to the Francis and Kučera (1982) written frequencies, for this study the smaller set of 50 nouns was categorized and balanced according to SUBTLEX-US verbal word frequencies (Brysbaert, New, & Keuleer, 2012). High- versus low-frequencies were categorized based on a median split (Storkel, et al., 2006). Once the word lists were finalized, corresponding non-contextualized, color photographic images were collected. In the first testing session, a POM screen was administered which included all 50 noun images presented randomly on a computer screen to elicit confrontation naming responses. Based on performance in this initial testing session, participants were assigned to either the high- or low-frequency word list in subsequent sessions.

Within group outcome measures

These secondary outcome measures address Research Questions #2 and #3: Is there a significant within group difference in confrontation naming accuracy and response latency for trained and untrained nouns measured at treatment completion and 3-months post treatment termination?

For participants receiving SFA, 240 total nouns were chosen from the English noun imageability dataset (N=2877 English nouns; Reilly & Kean, 2007) and supplemented by the SUBTLEX-US database (Brysbaert et al., 2012). These 240 nouns were sorted into eight different semantic categories: Body Parts, Clothing and Accessories, Food and Beverages, Household, Sports/Hobbies/Recreation, Nature, Occupations, and Transportation. Each category contained 15 high- and 15 low-frequency nouns. Noun verbal frequency information was found using the SUBTLEX-US word list. High- versus low-frequencies were categorized based on a median split (Storkel et al., 2006). All semantic categories had a median frequency of 308-475. Within semantic categories, high- and low-frequency word groups were statistically different from one another ($p < 0.01$). Finally, across semantic categories, averages were balanced (low-frequency range 111-229, SD = 43; high-frequency range 845-995, SD = 48) and were not statistically different. Once the word lists were finalized, corresponding non-contextualized, color photographic images were collected. Prior to administration to participants, these stimulus pictures were given to five neurologically healthy controls to ensure stimulus fidelity. Stimulus pictures were only used if inter-rater agreement was 100%.

For participants receiving PMT, both trained (N=39) and untrained (N=21) nouns were further assessed through confrontation naming. Phonotactic probabilities and neighborhood densities for the 60 total nouns were calculated using The Irvine Phonotactic Online Dictionary (IPHOD) calculator Version 2.0 (Vaden, Halpin, & Hickok, 2009). A number of linguistic properties were controlled, including frequency, imageability, age of acquisition, syllable number and complexity, and semantic category. Once the word lists were finalized, corresponding non-contextualized, color photographic images were collected. Prior to administration to participants, these stimulus pictures were given to five neurologically healthy controls to ensure stimulus fidelity. As with SFA stimuli, stimulus pictures were only used if inter-rater agreement was 100%.

Ecological validity

These secondary outcome measures address Research Question #4: Is there a between group difference in measures of ecological validity?

Ecological validity was assessed using two paper-based communication questionnaires: the Functional Outcome Questionnaire-Caregiver (FOQ-A; Glueckauf et al., 2003) and the Comprehensive Aphasia Test Disability Questionnaire (CAT-DQ; Swinburn et al., 2004). The FOQ is a measure given to a proxy rater, an unpaid family member or friend intimately familiar with the participant's communication, to assess everyday communicative function. The CAT-DQ is a participant-rated questionnaire administered by a speech-language pathologist to determine the impact of aphasia.

Outcome measure administration

Primary and secondary outcome measures involving confrontation naming (RQs 1-3) were administered two to three times over the course of one week at pre-treatment (A1), immediately post-treatment (A2), and three-months post-treatment (A3) time points. Measures of ecological validity (RQ 4) were administered one time during each testing period.

All confrontation naming measures were administered using Microsoft Office 2013 Power Point using a Dell Latitude 7370 laptop and HP EliteDisplay S140u (14-inch) display, or a Dell Optiplex 9020 Desktop and a Dell 24-inch monitor. Participants were seated approximately 30 inches from the screen and wore an Audio-Technica Power Module AT8531 head-mounted microphone connected to a Tascam US-125M USB mixing audio interface to audio-record responses. Participant verbal responses were recorded using Adobe Audition CS6 Version 5.0 software. For naming probes, participants received a brief practice session to train to task before stimuli were presented. They were asked to name the stimulus on the screen, which was presented for ten seconds, followed by a blank white screen for a brief period (2-10 seconds, depending on individual participant needs) before the next stimulus was presented. No cues were given to participants during confrontation naming tasks. Each stimulus and blank screen was paired with a click to mark the onset of stimulus presentation for later analyses of response latency.

Outcome measure analysis

The recorded verbal responses were scored for accuracy and response latency by trained research assistants (undergraduate or post-baccalaureate students in the University of Washington department of Speech and Hearing Sciences). The first verbal response was scored as either correct or incorrect. Verbal fillers or stereotypic verbalizations were not counted as a response. Phonologic errors (substitutions, additions, omissions, transposition, or mixed), semantic errors, neologisms, unrelated real words, nonresponses, and nonresponses with circumlocutory descriptions were all deemed incorrect. Motor speech errors (distortions that did not cross a phonemic boundary with perceptible place, manner, voice, or timing deviation; inserted schwa) were counted as correct (Bislick, McNeil, Spencer, Yorkston, & Kendall, 2017). Response latency was measured from the offset of the click to the onset of the participant response.

Reliability

Point-to-point reliability was performed on 10% of all confrontation naming outcome measures. Inter-rater reliability scoring was performed by a member of the data analysis team using audio recordings of the session. Ten percent of items were randomly selected by the rater for scoring (i.e. the scorer would choose a place in the audio recording of the probe session and score the first production after a click at that time in the recording). For participants with predominantly anomic, neologistic or unintelligible responses, where it was not possible for the rater to determine the target word when listening to random points in the recording, the rater scored the first 10% of items starting at the beginning of the recording. The time period between initial scoring and intra-rater reliability scoring was not controlled. Inter-class correlations demonstrated intra-rater reliability of 1.0 and inter-rater reliability of 0.98.

Effect Size

Effect size for each group was calculated for each of the outcome measures. Effect size was calculated by taking the mean of the change in scores for each participant in the group, and dividing by the standard deviation of the baseline scores. Criteria for judgment of effect size magnitude are those of Cohen (1998): 0.2 = small; 0.5 = medium; and 0.8 = large.

Statistical analysis

The primary analysis tested whether the improvement in POM differed between the two treatment groups. Analysis of variance was used to test whether change in POM from A1 to A2 differed between the groups after controlling for POM at A1. This analysis was repeated for change in POM from A1 to A3. Additional analyses tested whether there was evidence of improvement in POM within each treatment group. Within each treatment group, paired t-tests were used to test whether the mean change in POM differed significantly from zero. Secondary outcome measures were analyzed using these same methods. No multiple comparison adjustments were applied to p-values. This is appropriate since the primary analysis was specified in advance. However, results for secondary outcomes need to be interpreted cautiously in light of the multiple comparisons.

Results

Compliance with therapy completed by participants was 100%. An overview of the results is provided here and details are outlined in the Primary and Secondary Outcome Results sections. In brief, there was no significant between-group difference on POM (untrained nouns that were semantically and phonologically unrelated to each treatment, at 3 months post-treatment). Significant within group immediately post-treatment acquisition effects for confrontation naming and response latency were observed for both groups. Treatment-specific generalization effects for confrontation naming were observed for both groups immediately and 3 months post-treatment; a significant decrease in response latency was observed at both time points for the SFA group only. Finally, significant within-group differences on the CAT-DQ were observed both immediately and 3 months post-treatment for the SFA group, and significant within-group differences on the FOQ were found for both treatment groups 3 months post-treatment.

Primary outcome results

Research Question 1 examined between-group difference in generalization to untrained nouns at 3 months post-treatment. The primary outcome measures for this study were confrontation naming and response latency for untrained nouns that were unrelated, semantically or in terms of phonological sequence, to the stimuli used in treatment for the two groups. Results are summarized in Table 5.

Pretreatment naming accuracy was 40.33% (SD = 24.0) for the PMT group and 44.09% (SD = 28.7) for the SFA group, with no significant between-group difference ($p = .592$). Absolute change scores at 3 months post-treatment were 3.67% (SD = 8.9) (ES=0.15) for the PMT group and 4.26% (SD = 8.6) (ES = 0.15) for the SFA group, with no significant between-group difference in change score ($p = .797$).

Pretreatment response latency was 1.49s (SD = 0.54) for the PMT group and 1.40s (SD = 0.67) for the SFA group, with no significant between-group difference (p = .621). Change scores at 3 months post-treatment were -0.03s (SD = 0.36) for the PMT group and -0.06s (SD = 0.63) for the SFA group, with no significant between-group difference in change score (p = .632).

Secondary outcome results

Acquisition

Research Question 2 examined within group differences for trained nouns immediately following treatment completion. Results are summarized in Table 6.

PMT Group.

Pre-treatment naming accuracy was 34.97% (SD = 24.58) and immediately post-treatment was 53.01% (SD = 29.71), indicating an 18.04% absolute increase in performance (p = .000; ES = 0.73). Pre-treatment response latency (in seconds) was 1.66s (SD = 0.63) and immediately post-treatment was 1.33s (SD = 0.50), indicating a 0.33s decrease in latency (p = .012; ES = 0.52).

SFA Group.

Pre-treatment naming accuracy was 34.81% (SD = 21.28) and immediately post-treatment was 50.37% (SD = 29.79), indicating a 15.56% absolute increase in performance (p = .000; ES = 0.73). Pre-treatment response latency was 1.92s (SD = 0.57) and immediately post-treatment was 1.51s (SD = 0.58), indicating a -0.42s change in latency (p = .000; ES = 0.48).

Treatment-specific Generalization

Research Question 3 examined within group generalization to untrained nouns related to treatment stimuli, either in phonological sequence (for PMT) or semantically (for SFA), immediately and 3 months post-treatment. Results are summarized in Table 7.

Table 5. Research Question 1 results: between-groups differences in generalization to untrained nouns 3 months posttreatment.

	A1 score		A3 change score	
Variable	PMT	SFA	PMT	SFA
% Accuracy (*SD*)	40.33 (24.0)	44.09 (28.7)	3.67 (8.9) ES = 0.15	4.26 (8.6) ES = 0.15
p	*p* = .592		*p* = .797	
Median latency, s (*SD*)	1.49 (0.54)	1.40 (0.67)	−0.03 (0.36)	0.06 (0.63)
p	*p* = .621		*p* = .632	

Note. A1 = pretreatment; A3 = 3 months posttreatment; PMT = phonomotor treatment; SFA = semantic feature analysis; ES = effect size.

Table 6. Research Question 2 results: within-group differences for trained nouns immediately following treatment.

	PMT (*n* = 39 trained items)		SFA (*n* = 80 trained items)	
Variable	A1 score	A2 change score	A1 score	A2 change score
% Accuracy (*SD*)	34.97 (24.5)	18.04 (10.7) **ES = 0.73**	34.81 (21.3)	15.56 (10.7) **ES = 0.73**
p	***p* = .000**		***p* = .000**	
Median latency, s (*SD*)	1.66 (0.63)	−0.33 (0.58) **ES = 0.52**	1.92 (0.57)	−0.42 (0.42) **ES = 0.48**
p	***p* = .012**		***p* = .000**	

Note. Significant *p* values (< .05) and effect sizes (0.2 = small, 0.5 = medium, 0.8 = large) are in bold. PMT = phonomotor treatment; SFA = semantic feature analysis; A1 = pretreatment; A2 = immediately posttreatment; ES = effect size.

PMT Group.

Pre-treatment naming accuracy was 36.25% (SD = 25.70). Immediately post-treatment accuracy was 40.39% (SD = 27.36), indicating a 4.14% absolute increase in performance (p = .000; ES = 0.17). Three months post-treatment accuracy was 40.47% (SD = 26.54), indicating a 4.23% absolute increase in performance (p = .001; ES = 0.17). Pre-treatment response latency was 1.61s (SD = 0.67). Immediately post-treatment response latency was 1.58s (SD = 0.66), indicating a -0.03s change in latency (p = .687; ES = -0.13). Three months post-treatment response latency was 1.49s (SD = 0.58), indicating a -0.06s change in latency (p = .563; ES = 0.12).

SFA Group.

Pre-treatment naming accuracy was 35.42% (SD = 22.18). Immediately post-treatment accuracy was 42.24% (SD = 25.57), indicating a 6.82% absolute increase in performance (p = .000; ES = 0.31). Three months post-treatment accuracy was 41.70% (SD = 25.45), indicating a 5.06% absolute increase in performance (p = .002; ES = 0.31). Pre-treatment response latency was 2.00s (SD = 0.67). Immediately post-treatment response latency was 1.54s (SD = 0.57), indicating a -0.46s change in latency (p = .003; ES = 0.65). Three months post-treatment response latency was 1.50s (SD = 0.51), indicating a -0.48s change in latency (p = .001; ES = 0.69).

Ecologic Validity

Research Question 4 examined between and within group differences on two measures of ecologic validity immediately and 3 months post-treatment: the FOQ (caregiver report) and the CAT-DQ (patient report). Results are summarized in Table 8.

FOQ.

For the PMT group, scores on the FOQ were 3.49 (SD = 0.84) pre-treatment, 3.68 (SD = 0.74) immediately post treatment (change score = 0.192; p = .057; ES = 0.0), and 3.76 (SD =

Table 7. Research Question 3 results: within-group generalization to untrained nouns related to treatment stimuli.

	PMT (*n* = 21 untrained items)			SFA (*n* = 40 untrained items)		
Variable	**A1 score**	**A2 change score**	**A3 change score**	**A1 score**	**A2 change score**	**A3 change score**
% Accuracy (*SD*)	36.25 (25.7)	4.14 (5.5) ES = 0.17	4.23 (6.3) ES = 0.17	35.42 (22.2)	6.82 (7.9) **ES = 0.31**	5.06 (8.1) **ES = 0.31**
p	N/A	***p* = .000**	***p* = .001**	N/A	***p* = .000**	***p* = .002**
Median latency, s (*SD*)	1.61 (0.67)	−0.04 (0.44) ES = −0.13	−0.06 (0.51) ES = 0.12	2.00 (0.67)	−0.46 (0.64) **ES = 0.65**	−0.48 (0.61) **ES = 0.69**
p	N/A	*p* = .687	*p* = .563	N/A	***p* = .003**	***p* = .001**

Note. Significant *p* values (< .05) and effect sizes (0.2 = small, 0.5 = medium, 0.8 = large) are in bold. N/A means it was not appropriate to calculate a *p* value for the A1 score. PMT = phonomotor treatment; SFA = semantic feature analysis; A1 = pretreatment; A2 = immediately posttreatment; A3 = 3 months posttreatment; ES = effect size.

Table 8. Research Question 4 results—between-groups and within-group differences on measures of ecologic validity.

Within-group	PMT A1 score	PMT A2 change score	PMT A3 change score	SFA A1 score	SFA A2 change score	SFA A3 change score
CAT-DQ patient *t* score (*SD*)	53.79 (5.2)	1.07 (3.4) **ES = 0.21**	1.86 (5.0) **ES = 0.36**	54.59 (4.2)	1.48 (3.6) **ES = 0.34**	2.15 (3.6) ES = −0.36
p	N/A	*p* = .110	*p* = .060	N/A	***p* = .036**	***p* = .005**
FOQ caregiver score (*SD*)	3.49 (0.84)	0.192 (0.45) ES = 0.0	0.274 (0.48) **ES = 0.27**	3.69 (0.80)	0.136 (0.49) ES = 0.15	0.309 (0.59) ES = 0.03
p	N/A	*p* = .057	***p* = .013**	N/A	*p* = .161	***p* = .016**

Between-groups	A1 score PMT	A1 score SFA	A2 change score PMT	A2 change score SFA	A3 change score PMT	A3 change score SFA
CAT-DQ patient *t* score (*SD*)	53.79 (5.2)	54.59 (4.2)	1.07 (3.4)	1.48 (3.6)	1.86 (5.0)	2.15 (3.6)
p	*p* = .521		*p* = .574		*p* = .795	
FOQ caregiver score (*SD*)	3.49 (0.84)	3.69 (0.80)	0.192 (0.45)	0.136 (0.49)	0.274 (0.48)	0.309 (0.59)
p	*p* = .382		*p* = .983		*p* = .661	

Note. Significant *p* values (< .05) and effect sizes (0.2 = small, 0.5 = medium, 0.8 = large) are in bold. N/A means it was not appropriate to calculate a *p* value for the A1 score. PMT = phonomotor treatment; SFA = semantic feature analysis; A1 = pretreatment; A2 = immediately posttreatment; A3 = 3 months posttreatment; ES = effect size.

0.82) 3 months post-treatment (change score = 0.274; p = .013; ES = 0.27). For the SFA group, scores on the FOQ were 3.69 (SD = 0.80) pre-treatment, 3.83 (SD = .75) immediately post-treatment (change score = 0.136; p = .161; ES = 0.15); and 3.96 (SD = 0.85) 3 months post-treatment (change score = .309; p = .016; ES = 0.03). There was no significant difference in change scores between the PMT and SFA groups immediately or 3 months post-treatment.

CAT-DQ.

For the PMT group, scores on the CAT-DQ were 53.79 (SD = 5.2) pre-treatment, 52.72 (SD = 3.4) immediately post treatment (change score = 1.07; p = .110; ES 0.21) and 55.65 (SD 5.0) 3 months post treatment (change scores = 1.07; p = .060; ES 0.36). For the SFA group, scores on the CAT-DQ were 54.59 (SD = 4.17) pre-treatment, 56.07 (SD = 5.28) immediately post-treatment (change score = 1.48; p=.036; ES = 0.34), and 56.96 (SD = 5.02) 3 months post-treatment (change score = 2.15; p = .005; ES = 0.36). There was no significant difference in change scores between the PMT and SFA groups immediately or 3 months post-treatment.

Discussion

The objective of this study (Research Question 1) was to compare, using a randomized, parallel group trial design, the efficacy of PMT relative to SFA in improving production of untrained nouns on an object naming task 3-months after conclusion of treatment. There was no significant between-groups difference on the primary outcome measure (untrained nouns phonologically and semantically unrelated to each treatment) at 3 months posttreatment (p = .797).

Both treatments were efficacious in improving naming of trained words (ES for both 0.73)(Research Question 2) and in improving naming of untrained words that shared features (semantic or phonologic sequence) with the training set (Research Question 3), though the effect sizes were small (0.31 for SFA, 0.17 for PMT). Thus, there is evidence of learning of trained material with both treatments and there is evidence (small ESs) of generalization to untrained exemplars with shared features. However, no between group differences were noted at 3 months on words that did not share features with the training sets. Both treatments yielded improvements in measures of ecological validity (Research Question 4) but the effect sizes were small (0.03 for SFA, 0.27 for PMT on the FOQ; 0.36 for both on the CAT-DQ). Group differences in these gains were not statistically significant.

Prior studies using SFA have shown mixed generalization results likely due to the type of SFA treatment delivered or the effect of repeated exposure to generalization probes throughout the study (Boyle, 2010). The lack of generalization to untrained semantic domains also could be explained on the basis of lack of semantic features shared between domains. More recent work, employing an innovative semantic therapy, Verb Network Strengthening Treatment (VNest; Edmonds, Mammino, & Ojeda, 2014), suggests that there may be ways of getting around this limitation. VNest centers on the use of verbs. Because verbs prime nouns (agents or patients; Ferretti, McRae, & Hatherell, 2001) and nouns prime verbs (McRae, Hare, Elman, & Ferretti, 2005) and many verbs have synonyms (e.g., weigh, measure, quantify, assess, gauge, and plumb), VNest is theoretically capable of training the relationships between semantic features across a wide extent of semantic knowledge and transcending individual semantic domains. It is also a treatment that could easily be translated into a "game" that would be played by the participant with family and friends. PMT, in its present iteration, was designed with the concept that trained phonological sequence knowledge would generalize to untrained sequence knowledge. Ultimately, enhancement of phonological sequence knowledge could benefit the production of all words in the language. Unfortunately, the results of the present trial suggest that such generalization did not occur.

In neural network terms, failure of clinically significant generalization is likely to have occurred because phonological sequences are generally too sparsely coded, providing little basis for shared features between sequences. This stands in contrast to semantic domains, e.g., animals, which tend to be densely coded, hence a high prevalence of shared features (even as coding between semantic domains, such as between animals and tools, is sparse). These considerations raise two questions. First, do there exist subdomains of phonological sequence knowledge that share features and that are sufficiently common in spoken language that achievement of generalization in these subdomains would likely benefit performance with untrained exemplars and in daily verbal communication? Second, can PMT be modified to incorporate a set of training stimuli that are sufficiently commonly used in daily communication that this knowledge would be useful, even if it did not generalize to all phonological sequences? In a learning study, Storkel et al. (2006) showed a moderate advantage for low probability, high neighborhood density phonologic sequences — a finding that motivated the design of the PMT training set. However, their study did not speak to generalization effects.

Low frequency phonological sequences were employed in PMT on the assumption that they represented atypical exemplars. In a parallel distributed processing simulation involving “rehabilitation” of a trained semantic-phonological (lexical) network that had been damaged, Plaut (1996) found that rehabilitation of atypical exemplars benefitted network performance with both typical and atypical exemplars. This was because rehabilitating atypical exemplars also strengthened connectivity supporting typical exemplars but training typical exemplars could not strengthen connectivity incorporating atypical features. The generalization value of training atypical exemplars was subsequently borne out in a clinical trial employing a semantic feature treatment that involved naming, picture sorting, identifying semantic attributes, and answering yes/no questions about the semantic features of an item (Kiran & Thompson, 2003b). The current study, however, suggests that in the domain of phonological sequence knowledge, it is a mistake to regard low frequency exemplars as atypical in this same sense because they do not share enough features with higher frequency exemplars to generalize.

Finally, the iteration of PMT employed in this trial also incorporated training stimuli with large phonological neighborhoods. This was to maximize the number of semantic representations that, over time and with further acquisition of semantic-phonological (lexical knowledge) through Hebbian learning in day to day life, could be linked to the trained phonological sequences. The results of this trial do not enable us to judge the wisdom of this approach.

Future directions

In a preliminary analysis of the words and nonwords used in the PMT training set and the words used in the POM, we have applied the methodology of Vitevitch, Luce and Castro (Vitevitch & Castro, 2015; Vitevitch & Luce, 2016). This has revealed that there is very little phonological sequence similarity between the training set and POM exemplars. However, there do exist large realms of phonological sequence knowledge in which the sequences share many similarities (in analogy to semantic domains). Thus, it appears possible to redesign PMT such that it will have a much higher probability of yielding generalization to a redesigned POM. However, this redesign would involve explicitly training phonological sequences held in common with words that might be tested in follow-up. The treated words would need to generalize to untreated words in the same phonological sequence domains (a secondary outcome measure). However, the value of such domain-limited treatment for generalization to every day verbal communication, which spans all phonological sequence domains, would also have to be tested. Therefore, the primary outcome measure would necessarily become a discourse measure. These issues and our analysis of our PMT stimuli and the POM will be the subject of a follow-up paper.

Very strict scoring criteria were used in this study to assess performance on the POM. However, the ultimate goal of therapy for aphasia is to enhance daily verbal communication. In such communication, errors might be perfectly acceptable so long as they do not obscure meaning. Thus, many, if not most, phonemic substitutions (so long as they do not yield abstruse neologisms) and semantic near-miss errors might be acceptable.

Conclusion

The results of this randomized controlled trial of intensively delivered PMT and SFA in 58 participants with chronic post-stroke aphasia with anomia provides evidence that neither PMT nor SFA yielded clinically superior generalization to words that were untrained and neither semantically nor phonologically related to either training set (ES's were small). Each treatment produced predicted improvement in trained and untrained words that were semantically (SFA) and phonologically (PMT) related to the training sets, consistent with prior studies.

The lack of generalization is problematic in all anomia treatments and was evidenced in this trial. As such, we do not advocate translating either PMT or SFA to clinical practice, both because the limited generalization does not justify the magnitude of the treatment effort and because what we are reporting is the results of a phase II trial. A larger, phase III trial would be necessary to assure that the results reported are statistically representative of the general population of patients with aphasia characterized by anomia and phonological impairment. We are hopeful that further refinements of PMT and SFA will achieve sufficiently large effect sizes on measures of generalization to warrant their translation to the clinic, either individually or, in many patients, in sequence, since many patients have both phonological and semantic impairment. Other treatments tested in large numbers of aphasic patients, e.g. intensive comprehensive aphasia programs (Babbitt, Worrall, Cherney, 2015; Breitenstein, Grewe, Floel, Ziegler, Springer and Martus, 2017; Dignam, Copland, McKinnon, Burfein, O'Brien, Farrell, 2015; WinansMitrik, Hula, Dickey, Shumacher, Swoyer, Doyle, 2014) and constraint induced language therapy (Meinzer, Djundja, Barthel, Elbert, Rockstroh, 2005; Meinzer, Streitfau, Rockstroch, 2007), treatments that may achieve their effect through improving intentionality to speak, have also achieved small effect sizes in the 0.18 to 0.32 range. There are ways in which these treatments could potentially be improved to increase effect size. There are still other mechanisms of generalization that have scarcely been tested (Nadeau, 2015). These small effect sizes also point to the need to increase the neural substrate for language function through neurobiological interventions in the days, weeks and months following stroke (Cramer, 2015). Because, on average, the treatment control that is achieved in clinical trials is probably not achieved in the clinic, we believe that research is needed on mechanisms by which speech language therapists benefit patients and how these mechanisms can be better leveraged. Clinical observation suggests that therapists may have a profound influence on the psychological wellbeing and outlook of patients dealing with one of the worst turns of their lives. Finally, generalization is not a sine qua non of all aphasia therapy: Some patients, particularly those with rudimentary language function, may benefit from learning a very limited number of words of particularly high communicative significance.

Acknowledgements: The authors would like to acknowledge the participants and their families for their time and efforts. We are very grateful to Kevin Cain for his statistical input, Kerry Lam for her assistance in recruiting and scheduling participants, and Jenny Erpenbeck for managing the data analysis team. We acknowledge the students at the University of Washington who were involved in data analysis and the Aphasia Network (Oregon) for participant recruitment. The contents of this publication do not represent the views of the U.S. Department of Veterans Affairs or the United States government.

References

Babbitt EM, Worrall L, Cherney LR. Structure, processes, and retrospective outcomes from an intensive comprehensive aphasia program. Journal of Speech Language and Hearing Research. 2015;24:S854-S63.

Bislick, L., McNeil, M., Spencer, K.A., Yorkston, K., & Kendall, D.L. (2017). The nature of error consistency in individuals with acquired apraxia of speech and aphasia. *American Journal of Speech-Language Pathology, 26*(2S), 611-630. https://doi.org/10.1044/2017_AJSLP-16-0080

Boyle, M. (2004). Semantic Feature Analysis treatment for anomia in two fluent aphasia syndromes. *American Journal of Speech-Language Pathology, 13*(3), 236-249. https://doi.org/10.1044/1058-0360(2004/025)

Boyle, M., & Coelho, C. (1995). Application of Semantic Feature Analysis as a treatment for aphasic dysnomia. *American Journal of Speech-Language Pathology, 4*(4), 94-98. https://doi.org/10.1044/1058-0360.0404.94

Boyle, M., (2010) Semantic Feature Analysis Treatment for Aphasic Word Retrieval Impairments: What's in a Name? *Topics in Stroke Rehabilitation*, 17(6):411-22. https://doi.org/10.1310/tsr1706-411

Breitenstein C, Grewe T, Flöel A, Ziegler W, Springer L, Martus P, et al. Intensive speech and language therapy in patients with chronic aphasia after stroke: a randomised, open-label, blinded-endpoint, controlled trial in a health-care setting. Lancet. 2017;389:1528-38.

Brookshire, C.E., Conway, T., Hunting Pompon, R., Oelke, M., & Kendall, D. (2014). Effects of intensive phonomotor treatment on reading in eight individuals with aphasia and phonological alexia. *American Journal of Speech-Language Pathology, 23*(10): S300-S311. https://doi.org/10.1044/2014_AJSLP-13-0083

Brysbaert, M., New, B., & Keuleers, E. (2012). Adding Part of Speech information to the SUBTLEX-US word frequencies. *Behavior Research Methods, 44*(4), 991-997. https://doi.org/10.3758/s13428-012-0190-4

Cohen, J. (1998). *Statistical power analysis for the behavioral sciences* (2nd ed). Mahwah, NJ: Erlbaum.

Cramer SC. Treatments to promote neural repair after stroke. Journal of Stroke. 2018;20:57-70.

Dignam J, Copland D, McKinnon E, Burfein P, O'Brien K, Farrell A, et al. Intensive versus distributed aphasia therapy. A nonrandomized, parallel group, dosage-controlled study. Stroke. 2015;46:2206-11.

Druks, J., & Masterson, J. (2000). *An object and action naming battery.* Hove, East Sussex, UK: Psychology Press.

Duffy, J. R. (2013). *Motor speech disorders: Substrates, differential diagnosis and management* (3rd ed.). St. Louis, MO: Elsevier Mosby.

Edmonds, L. A., & Babb, M. (2011). Effect of verb network strengthening treatment in moderate-to-severe aphasia. *American Journal of Speech-Language Pathology, 20*(2), 131-145. https://doi.org/10.1044/1058-0360(2011/10-0036

Edmonds, L.A., Mammino, K., & Ojeda, J. (2014). Effect of verb network strengthening treatment (VNeST) in persons with aphasia: extension and replication of previous findings. *American Journal of Speech-Language Pathology, 23*(2)*,* S312-S29. https://doi.org/10.1044/2014_AJSLP-13-0098

Ferretti, T. R., McRae, K., & Hatherell, A. (2001). Integrating verbs, situation schemas, and thematic role concepts. *Journal of Memory and Language, 44*(4), 516-547. https://doi.org/10.1006/jmla.2000.2728

Francis, W., & Kučera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar.* Boston: Houghton Mifflin.

Glueckauf, R. L., Blonder, L. X., Ecklund-Johnson, E., Maher, L., Crosson, B., & Gonzalez-Rothi, L. (2003). Functional outcome questionnaire for aphasia: Overview and preliminary psychometric evaluation. *NeuroRehabilitation, 18*(4), 281-290.

Kendall, D. L., del Toro, C., Nadeau, S. E., Johnson, J., Rosenbek, J., & Velozo, C. (2010, June). *The development of a standardized assessment of phonology in aphasia*. Paper session presented at the Clinical Aphasiology Conference, Isle of Palm, SC.

Kendall, D. L., Rosenbek, J. C., Heilman, K. M., Conway, T., Klenberg, K., Gonzalez-Rothi, L., & Nadeau, S. E. (2008). Phoneme-based rehabilitation of anomia in aphasia. *Brain and Language, 105*(1), 1-17. https://doi.org/10.1016/j.bandl.2007.11.007

Kendall, D., Oelke, M., Brookshire, C.E., & Nadeau, S.E. (2015). The influence of phonomotor treatment on word retrieval abilities in 26 individuals with chronic aphasia: An open trial**.** *Journal of Speech, Language and Hearing Research, 58*(3): 798-812. https://doi.org/10.1044/2015_JSLHR-L-14-0131

Kendall, D., & Nadeau, S.E. (2016). The phonomotor approach to treating phonological-based language deficits in people with aphasia. *Topics in Language Disorders*, *36*(2), 109-122. https://doi.org/10.1097/TLD.0000000000000085

Kiran, S., & Thompson, C.K. (2003a). Effect of typicality on online category verification of animate category exemplars in aphasia. *Brain and Language, 85*(3), 441-450. https://doi.org/10.1016/S0093-934X(03)00064-6

Kiran, S., & Thompson, C.K. (2003b). The role of semantic complexity in treatment of naming deficits: training semantic categories in fluent aphasia by controlling exemplar typicality. *Journal of Speech, Language, and Hearing Research, 46*(3), 773-87. https://doi.org/10.1044/1092-4388(2003/048)

McRae, K., Hare, M., Elman, J., & Ferretti, T. (2005). A basis for generating expectancies for verbs from nouns. *Memory and Cognition 33*(7), 1174-1184. https://doi.org/10.3758/BF03193221

Maddy, K.M., Capilouto, G.J., & McComas, K.L. (2014). The effectiveness of semantic feature analysis: An evidence-based systematic review. *Annals of Physical and Rehabilitation Medicine*, *57*(4), 254-267. https://doi.org/10.1016/j.rehab.2014.03.002

Meinzer M, Djundja D, Barthel G, Elbert TR, Rockstroh B. Long-term stability of improved language functions in chronic aphasia after constraint-induced aphasia therapy. Stroke. 2005;36:1462-6.

Meinzer M, Streiftau S, Rockstroh B. Intensive language training in the rehabilitation of chronic aphasia: efficient training by laypersons. Journal of the International Neuropsychological Society. 2007;13:846-53.

Minkina, I., Oelke, M., Bislick, L., Brookshire, C.E., Hunting Pompon, R., Silkes, J., & Kendall, D. (2015). An investigation of aphasic naming error evolution following phonomotor treatment. *Aphasiology, 30*(8), 962-980. https://doi.org/10.1080/02687038.2015.1081139

Nadeau, S. E. (2001). Phonology: A review and proposals from a connectionist perspective. *Brain and Language, 79*(3), 511-579. https://doi.org/10.1006/brln.2001.2566

Nadeau, S.E. (2012). *The neural architecture of grammar*. Cambridge, MA: MIT Press.

Nadeau, S.E. (2015). Neuroplastic mechanisms of language recovery after stroke. In J.E. Tracy, B.M. Hampstead, & K. Sathian (Eds.), *Cognitive Plasticity in Neurologic Disorders* (61-84). Oxford: Oxford University Press.

Nadeau, S.E., Dobkin, B., Wu, S.S., Pei, Q., & Duncan, P.W. (2016). The Effects of Stroke Type, Locus, and Extent on Long-Term Outcome of Gait Rehabilitation: The LEAPS Experience. *Neurorehabilitation and Neural Repair, 30*(7), 615–625. https://doi.org/10.1177/1545968315613851

Nadeau, S.E., & Crosson, B. (1997). Subcortical aphasia. *Brain and Language, 58*(3), 355-402. https://doi.org/10.1006/brln.1997.1707

Plaut, D. C. (1996). Relearning after damage in connectionist networks: Toward a theory of rehabilitation. *Brain and Language, 52*(1), 25-82. https://doi.org/10.1006/brln.1996.0004

Pocock, S.J. (1983). *Clinical trials: a practical approach*. New York: John Wiley and Sons.

Pocock, S.J., & Simon, R. (1975). Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. *Biometrics, 31*(1), 103-115.

Reilly, J., & Kean, J. (2007). Formal distinctiveness of high- and low-imageability nouns: Analyses and theoretical implications. *Cognitive Science, 31*(1)*,* 157-168. https://doi.org/10.1080/03640210709336988

Roach, A., Schwartz, M.F., Martin, N., Grewal, R.S., & Brecher, A. (1996). The Philadelphia Naming Test: Scoring and rationale. *Clinical Aphasiology, 24*, 121–133.

Robey, R. (1994). The efficacy of treatment for aphasic persons: a meta-analysis. *Brain and Language*, *47*(4), 582-608. https://doi.org/10.1006/brln.1994.1060

Storkel, H. L., Armbruster, J., & Hogan, T. P. (2006). Differentiating phonotactic probability and neighborhood density in adult word learning. *Journal of Speech, Language and Hearing Research, 49*(6), 1175-1192. https://doi.org/10.1044/1092-4388(2006/085)

Swinburn, K., Porter, G., & Howard, D. (2004). *Comprehensive Aphasia Test*. New York, NY: Psychology Press.

Vaden, K. I., Halpin, H. R., & Hickok, G. S. (2009). Irvine Phonotactic Online Dictionary, Version 2.0 [Data file]. Retrieved from www.iphod.com.

Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language, 40*(3), 374-408. https://doi.org/10.1006/jmla.1998.2618

Vitevitch, M.S., & Castro, N. (2015). Using network science in the language sciences and clinic. *International Journal of Speech-Language Pathology, 17*(1),13-25. https://doi.org/10.3109/17549507.2014.987819

Vitevitch, M.S., & Luce, P.A. (2016). Phonological neighborhood effects in spoken word perception and production. *Annual Review of Linguistics, 2*, 75-94. https://doi.org/10.1146/annurev-linguistics-030514-124832

Wambaugh, J. L., Duffy, J. R., McNeil, M. R., Robin, D. A., & Rogers, M. A. (2006). Treatment guidelines for acquired apraxia of speech: A synthesis and evaluation of the evidence. *Journal of Medical Speech-Language Pathology*, *14*(2), xv–xxxiii.

Winans-Mitrik RL, Hula WD, Dickey MW, chumacher JG, Swoyer B, Doyle PJ. Description of an intensive residential aphasia treatment program: rationale, clinical processes, and outcomes. American Journal of Speech-Language Pathology. 2014;23:S330-S42.

Wisenburn, B., & Mahoney, K. (2009). A meta-analysis of word-finding treatments for aphasia. *Aphasiology, 23*(11), 1338-1352. https://doi.org/10.1080/02687030902732745