Effects of Phonomotor Therapy and Semantic Feature Analysis on Discourse Production

JoAnn P. Silkes,[a,*] Gerasimos Fergadiotis,[b] Kasey Graue,[b] and Diane L. Kendall[c,d]

[a] University of Washington, Seattle

[b] Speech and Hearing Sciences Department, Portland State University, OR

[c] Department of Speech and Hearing Sciences, University of Washington, Seattle

[d] University of Pretoria South Africa

*Correspondence to JoAnn P. Silkes, who is now at San Diego State University: jsilkes@sdsu.edu

Kasey Graue is now at the University of Oregon

Abstract

Background: Anomia treatments typically focus on single word retrieval, although the ultimate goal of treatment is to improve functional communication at the level of discourse in daily situations.

Aims: The focus of this study was to investigate the impact of two effective anomia treatments on discourse production as measured by a story retell task.

Method and Procedure: Fifty-seven people with aphasia were randomized to receive either a phoneme-based treatment, Phonomotor Therapy (PMT; 28 participants), or a lexical–semantic treatment, Semantic Feature Analysis (SFA; 29 participants). Groups were matched for age, aphasia severity, education, and years post onset. All received 56–60 hr of treatment in a massed treatment schedule. Therapy was delivered for a total of 8–10 hr/week over the course of 6–7 weeks. All participants completed testing 1 week prior to treatment (A1), immediately following treatment (A2), and again 3 months later (A3). Discourse was analyzed through the percentage of correct information units at each time point.

Outcomes and Results: Both groups showed nonsignificant improvements from pretreatment to immediately posttreatment. The PMT group showed significant improvement 3 months posttreatment, while the SFA group returned to near-baseline levels.

Conclusion: These results add to our understanding of the effects of both PMT and SFA. Future research should address understanding variability in discourse outcomes across studies and the effects of aphasia severity and individual participant and treatment factors on treatment outcomes for both of these approaches.

The ultimate goal of aphasia treatment is to improve a person's functional communication. Aphasia treatments are known to be effective at improving whatever skills and stimuli are directly treated, but generalization to untrained words, structures, and levels of language function is inconsistent, at best (Wisenburn & Mahoney, 2009), and reliable generalization to the level of discourse has been elusive. This study was undertaken to further investigate the question of whether treatment of word retrieval problems in aphasia, which are ubiquitous and, therefore, a common treatment target, can lead to improvements in discourse-level verbal production. Specifically, we were interested in understanding effects on discourse from treatment of phonology with Phonomotor Therapy (PMT; Kendall et al., 2019, 2015, 2008) and comparing that with effects on discourse from treatment with the current gold standard treatment approach for word retrieval, Semantic Feature Analysis (SFA; Boyle, 2004; Boyle & Coelho, 1995; Coelho et al., 2000).

PMT

PMT focuses on improving phonological processing and phonological sequence knowledge, providing intense practice with phonological manipulation tasks (see details in the Method section). PMT involves training linguistic regularities, which are a finite set of linguistic elements (in this case, phonemes) or operations that are encountered repeatedly across an infinite set of words and phrases. For example, if the phoneme sequence *int* is trained, then words with the same sequence (e.g., mint, lint, intelligence, interrupt, hinted, sprinting) will be affected by any improvements that occur. Improving the function of fundamental operations or elements has widespread effects on the entire system because they are widely connected to all levels of language processing (Nadeau, 2015).

PMT is motivated by a parallel distributed processing (PDP) model of language, which posits that language is represented across multimodal domains that are massively interconnected, with reciprocal activation across domains (Nadeau, 2001; see Kendall et al., 2008, for a discussion of how PDP models motivate PMT). Because of the interactive nature of the linguistic processing system, training phonological sequence knowledge is expected to strengthen related representations and function across lexical contexts. In addition, it should improve function at other levels of processing, including connected discourse. Indeed, prior studies have found generalization from PMT across lexical items (Kendall et al., 2019, 2015, 2008) and from single word production to reading (C. E. Brookshire et al., 2014) and discourse (Kendall et al., 2015; Silkes et al., 2018). Specifically, Silkes et al. (2018) found that individuals who had received 60 hr of PMT showed improvement in discourse immediately following treatment as measured by greater correct information units (CIUs) produced per minute (Nicholas & Brookshire, 1993). In addition, 3 months posttreatment, both CIUs per minute and percent CIUs (%CIUs) produced had improved. The primary objective of the current study was to further explore the effects of PMT on discourse for a new group of people with aphasia (PWA).

SFA

In contrast to PMT, SFA involves focusing on the semantic features of specific target words. In response to a picture stimulus, the client provides semantic features that describe the target item, such as where the target item is found, what it looks like, what category it falls into, what it is used for, who uses it, and any personal associations they may have with it. SFA is motivated by an interactive activation (IA) model of language (Dell, 1986; see Boyle, 2010, for a discussion of how IA models motivate SFA). Similar to the PDP model described above, the IA model is also a network model of linguistic processing in which phonological and semantic information are represented separately within systems that allow bidirectional communication between the two domains. The potential active mechanism of SFA has been described in a few different ways. First, it may train an explicit strategy for word retrieval, with clients learning to use descriptions to assist when word retrieval fails (Massaro & Tompkins, 1992). If this is the case, then generalization to untrained words and to the level of discourse production may occur due to conscious application of the strategy rather than an intrinsic change to the language processing networks. In addition, some have suggested that the process of working in "a structured, methodical way over and over again" (Boyle, 2004, p. 246) can lead to automatization of the process of recruiting semantic features during word retrieval, regardless of the target (Haarbauer-Krupa et al., 1985). If this is the case, then any generalization to untrained words may be the result of improved procedural memory for word retrieval processes, which may have the potential to carry over to more complex levels of discourse. Another possible mechanism that is widely discussed in the literature is that the SFA treatment process reinforces connections within the lexical–semantic system by increasing the spread of activation between related items within these networks (Boyle, 2010; Gravier et al., 2018; Kiran, 2008; Kiran & Johnson, 2008; Quique et al., 2018). As the client describes the semantic features of the target item, the semantic feature representations are strengthened and their connections to their associated lexemes are reinforced; in addition, the more strongly activated representations also spread activation to other related nodes in the lexical–semantic network. Because of this automatic spread of activation within the system, there is potential for improved naming of items that are semantically related to the trained words, as their shared semantic features would be made more readily accessible (Efstratiadou et al., 2018). Given this mechanism, generalization to discourse could occur as the result of the lexical system becoming more efficient and effective overall, which could translate to discourse production. At the same time, however, because trained semantic features are shared only by related items, rather than being shared widely across all words as linguistic regularities, generalization to discourse may not occur.

SFA has been shown to reliably improve naming of items that are trained in therapy, with inconsistent evidence for generalization to untrained single words (see reviews by Boyle, 2010; Efstratiadou et al., 2018; Quique et al., 2018). Particularly relevant to this study, there is also inconsistent evidence for generalization from SFA treatment to discourse. Boyle (2004) found generalization on one of three discourse measures (CIUs per minute; described below in Outcome Measure Description and Analysis) for one of two participants, with a drop in this measure noted 1 month after treatment, although performance remained above baseline levels. The second participant showed an improvement in the number of CIUs produced immediately after treatment, but without knowing the total number of words produced in the discourse sample, this result is not interpretable. This participant did not show any changes in CIUs per minute or %CIUs immediately after treatment, and no maintenance data were reported. Coelho et al. (2000) also reported changes in discourse following SFA, but here, too, the evidence is limited. In this single-case study, the participant showed an increase in CIUs per minute and %CIUs during the first phase of treatment. During the second treatment phase, there was a slight continued increase in CIUs per minute and a decrease in %CIUs back to baseline levels. One month after treatment, the increase in CIUs per minute was maintained while %CIUs rose again, but by 2 months posttreatment, both of these measures returned to baseline levels. Finally, DeLong et al. (2015) reported on SFA treatment provided to five PWA. They found that three of the participants showed an increase in %CIUs, though one of these (from 19% to 23%) was arguably not functionally meaningful. No maintenance data were provided.

Collectively, these three studies, with a total of eight participants, have provided inconsistent evidence that discourse-level changes in communication may occur immediately following SFA treatment. There are scant data on whether improved discourse may be maintained over time, with two out of three participants for whom maintenance data were provided showing some continued higher levels of performance relative to baseline up to 2 months posttreatment. Other studies have provided additional data, but these studies have confounding factors that make them less applicable to the current discussion. For instance, Davis and Stanton (2005) reported improvements in some discourse measures both immediately after treatment and inconsistently through 30 weeks posttreatment. While this study provided noun-focused SFA, as did the studies reported above, the participant also received treatment in other domains, such as auditory comprehension, reading comprehension, and writing. In addition, the participant was 4 months postonset of aphasia when the study began, creating the possibility that improvements were influenced by spontaneous recovery mechanisms. Therefore, it is not possible to confidently attribute the observed improvements in discourse to SFA. Wambaugh and Ferguson (2007) also found improvements in discourse, but the participant received SFA focused on verb targets, which presumably influences the lexical processing system in different ways than noun-focused SFA. Generalization to discourse has also been noted after training specific vocabulary to use in discourse contexts (Rider et al., 2008); this improvement is likely largely stimulus-bound, and treatment in a discourse context is not comparable to, and does not inform predictions for, noun-focused SFA. Other studies that have also found generalization from SFA-like treatments to discourse (Antonucci, 2009; Boyle, 2011; Falconer & Antonucci, 2012; Knoph et al., 2015, 2017; Peach & Reuter, 2010) have involved training in a discourse context and/or in conjunction with other treatment methods, rather than using SFA at the single noun level. These studies are not immediately relevant to the current report as there are potentially multiple factors at work to account for generalization that do not apply to the study reported here. Finally, generalization to discourse has been investigated through a recent meta-analysis (Oh et al., 2016), which revealed effect sizes ranging from none to large, across a total of six participants in three studies (two of which provided SFA in a group context; Antonucci, 2009; Falconer & Antonucci, 2012). The authors of that meta-analysis concluded that evidence for generalization to discourse was limited.

With this perspective on SFA in mind, a secondary objective of this study was to assess generalization to discourse in a large sample of PWA receiving SFA. In addition, given that this study was a continuation of an ongoing line of research to develop and establish the efficacy and effectiveness of PMT, the third objective of this study was to compare these outcomes with those obtained for PWA receiving an equally intense program of PMT. To address both primary and secondary objectives, the following three questions were asked:

1. Does PMT lead to changes in discourse production immediately after treatment and 3 months posttreatment? Based on prior research (Silkes et al., 2018) and the theoretical motivation of PMT, we predicted no significant change in %CIUs produced immediately posttreatment but a significant improvement 3 months posttreatment.
2. Does SFA lead to changes in discourse production immediately after treatment and 3 months posttreatment? We were unable to make firm predictions for this question for two reasons. First, as reported above, the available evidence for comparable forms of SFA is inconsistent, making it difficult to predict whether a much larger sample of PWA will show overall significant effects or not. In addition, the amount of treatment provided here is more than double the time provided in any other previous reports, which has the potential to lead to different outcomes.

3. Is there a significant difference in changes to discourse production between PMT and SFA immediately after treatment or 3 months posttreatment? Based on prior research and the models that motivate these treatments and predict their patterns of generalization, we predicted that PMT and SFA would not differ in discourse effects immediately posttreatment (A2) but that PMT would yield significantly greater change in discourse production than SFA 3 months posttreatment (A3).

Method

All procedures were approved by the institutional review boards at the University of Washington, Portland State University, and VA Puget Sound Health Care System. Written informed consent was obtained from all participants.

Study Design

The analysis presented here was based on data collected from 57 participants, all of whom were included in the data reported by Kendall et al. (2019). Because the participants, protocol, and stimuli have been detailed in this previous publication, they are only briefly summarized here.

Participants

Participants for this project 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. Computed tomography and magnetic resonance imaging scans and/or reports were used to document the presence of the stroke. Fifty-seven individuals completed the entirety of treatment and provided discourse samples at all time points during posttreatment (A2) and maintenance testing (A3). These 57 PWA were included in this discourse analysis (see Table 1 for a summary of participant characteristics and Appendix for detailed information on each participant in each group).

Table 1. Averages (and standard deviations) for the demographic and clinical characteristics of the participant sample.

Group	Sex	Education	Age	Years post onset
PMT (*n* = 28)	13F, 15M	14.2 (2.0)	63.3 (10.6)	4.3 (4.7)
SFA (*n* = 29)	11F, 18M	15.2 (2.8)	64 (12)	4.1 (4.3)

Note. PMT = Phonomotor Therapy; F = female; M = male; SFA = Semantic Feature Analysis.

Study inclusion required that participants demonstrate chronic aphasia (> 6 months post onset), with anomia determined via performance on the Comprehensive Aphasia Test (Swinburn et al., 2004), and sufficient auditory comprehension to follow basic directions. All participants demonstrated impaired phonological processing, as seen by performance on the Standardized Assessment of Phonology in Aphasia (Kendall et al., 2010). All participants passed hearing screening at 500, 1000, and 2000 Hz at 25 dB HL in at least one ear.

Participants with mild to moderate apraxia of speech were included. Presence of apraxia of speech was determined perceptually by two speech-language pathologists (SLPs) who observed a video-recorded speech sample showing picture description, spontaneous conversation, automatic speech, repetitions of words of increasing length, and multiple repetitions of three-syllable words. These samples were evaluated for slow speech rate, sound distortions, and prosodic abnormalities. Individuals were excluded from study participation if they exhibited untreated depression, degenerative neurological disease; chronic medical illness that would be disruptive to the rigorous study schedule (e.g., chronic kidney disease requiring dialysis), and/or severe, uncorrected impairment of vision, all determined by participant or caregiver report. None of the participants received any other speech-language therapy during the course of their participation in this study, including during the 3-month maintenance interval.

The 28 participants who completed PMT were, on average, 63.3 years of age (*SD* = 10.6), had 14.2 years of education (*SD* = 2.0), and were 4.3 years post stroke onset (*SD* = 4.7); 15 were men, and 13 were women. All were monolingual English speakers. The 29 participants who completed SFA were, on average, 64 years of age (*SD* = 12), had 15.2 years of education (*SD* = 2.8), and were 4.1 years post stroke onset (*SD* = 4.3); 18 were men, and 11 were women. One participant in the SFA group (SF01PDX) was bilingual but used primarily English prior to and after their stroke. Given that the participant was not identified as a univariate or multivariate outlier, as described below, they were included in the analysis.

Treatment

Participants were randomized to receive either SFA or PMT, with the two groups matched for age, aphasia severity, education, and years post onset. All participants received 56–60 hr of treatment in a massed treatment schedule. Therapy was administered by a licensed and certified research SLP and was delivered for a total of 8–10 hr/week over the course of 6–7 weeks. Treatment fidelity was monitored by graduate students who evaluated 10-min, randomly selected audio samples that were recorded in 20% of the therapy sessions. The average treatment fidelity across weeks and participants was 96.75% for PMT and 99.51% for SFA. Details of the treatment protocols are fully described in Kendall et al. (2019), but an overview of the major points is provided here.

PMT Treatment and Stimuli

Phonomotor treatment focuses on manipulating phonology through multimodality tasks, beginning at the level of single phonemes and proceeding to increasingly complex nonwords and, periodically throughout treatment, real words. Treatment tasks focus exclusively on phonological awareness and manipulation, including activities such as identifying, recognizing, and repeating single phonemes; recognizing phonological contrasts; matching phonemes to mouth pictures that represent them; parsing and blending syllables; identifying the number of syllables in stimuli; matching phonemes to graphemes; and chaining from one syllable to another through single phoneme changes (e.g., “if this says ‘peef,' make it say ‘seef’”). Because phonological representations are widely distributed in the brain, the treatment tasks engage the phonological system across multiple domains, including acoustic, tactile–kinesthetic, articulatory–motor, visual, and orthographic processing. Throughout all tasks, the clinician engages the client in evaluation of stimuli and responses through Socratic questioning (see Kendall et al., 2019, for a full summary of treatment procedures, or a detailed Guide to Phonomotor Therapy is available upon request to the first author).

Stimuli used in PMT are highly varied. They primarily involve a wide range of nonword stimuli that are created spontaneously in the course of a variety of phonological manipulation tasks. In addition, there are 39 specific real words and 69 specific phonotactically legal nonwords that are intentionally incorporated across tasks by interweaving them with other nonword stimuli throughout all tasks. All of these specific, prescribed items are one- and two-syllable items of low phonotactic probability and high phonological neighborhood density, as determined using methods similar to those outlined by Vitevitch and Luce (1999). Phonotactic probability refers to the frequency with which a phonologic sequence occurs in the language, while phonological neighborhood density is the number of words in the language that differ from a target word by a single phoneme. These phonotactic/phonological criteria were established based on the word-learning literature (Storkel et al., 2006) to maximize the likelihood of generalizing the effects of treatment through engaging widespread phonological networks. The real words were included to facilitate the spread of activation between phonological processing and semantic processing. Importantly, however, when real words were incorporated into tasks, they occurred within the stream of nonwords that comprised the majority of treatment stimuli, and the focus remained solely on phonological aspects of processing; the semantic aspects of the word were never invoked or addressed.

SFA Treatment and Stimuli

SFA involves exploring the semantic features of specific target words. Each treatment trial began with the client attempting to name the target picture and then, regardless of whether they named it accurately or not, the client provided descriptions of the target such as where it is found, what it looks like, what category it falls into, what it is used for, who uses it, and any personal associations they may have with it. Clinician assistance was provided if the client was unable to generate semantic features. After the features had been generated, the client was asked to name the picture 3 times in a row. If they did not do so, the clinician reviewed the features again and repeated the target word along with one feature descriptor (e.g., "*dog* belongs to the group *animals*"). The participant was then asked to try to repeat the target name 3 times before moving on to the next training item.

There were two different sets of treatment stimuli used in SFA treatment. Each comprised 80 nouns, with 10 nouns in each of eight different semantic categories. One set included high-frequency words and was administered to participants whose initial testing demonstrated that their naming skills were relatively more impaired (see Kendall et al., 2019, for details of this process). The other set included low-frequency words and was administered to participants whose initial testing demonstrated that their naming skills were relatively less impaired. All participants received training on all items for their assigned frequency group. During the treatment session, stimuli were color photographs of items in isolation (i.e., no other context was provided in the pictures). Selected photographs were pretested by five neurologically healthy control participants, who achieved 100% interrater agreement on the name of each item.

Outcome Measure Description and Analysis

Discourse language samples were collected and audio-recorded for all participants in a single session pretreatment (A1), immediately posttreatment (A2; 7–8 weeks after A1), and 3 months posttreatment (A3; i.e., maintenance) through a structured, face-to-face interaction between the participant and the research SLP who had conducted treatment with that participant. Discourse samples were elicited using the stimuli from the story retell procedure

(SRP; Doyle et al., 1998; McNeil et al., 2001). We used this approach because the SRP is associated with less variability during administration and elicitation of language samples, relative to other discourse elicitation procedures, and the elicited language samples have known and constrained content, allowing more reliable transcription, coding, and analysis. The stimuli for the SRP are based on the 12 stories from the Discourse Comprehension Test (R. H. Brookshire & Nicholas, 1993), each of which has an associated six-plate black-and-white illustration. Stories were presented using the digital audio files developed by Doyle et al. (1998), which were read and recorded at a rate of 170 words per minute. Doyle et al. (2000) developed four parallel forms, each consisting of three stories. These four sets of stories do not differ significantly across four connected language production domains including verbal productivity, information content, grammatical structure and word errors, and verbal disruptions. Furthermore, Doyle et al. reported strong, positive, and significant correlations among test forms with rho values ranging from .61 to .99 ($p < .01$). For this study, two of the four parallel forms were merged to create a single form with six stories and the other two of the four parallel forms were merged to create a second form with six stories; this allowed a longer discourse sample from each participant at each time point. The first form was used to elicit discourse at A1 and A3, while the other form was used to elicit discourse at A2.

For the SRP, participants were seated in front of a computer and listened to each of the prerecorded stories while viewing the associated pictures on the screen. Prior to the presentation of each story, participants heard the following instructions: “You are about to hear a short story. As you listen, pictures that go with the story will appear on the screen. Listen to the story and watch the pictures carefully. When the story is completed you will be asked to retell the story.” Following the presentation of the story, the complete six-plate sequence (in two rows of three pictures each) was displayed on the computer screen, and participants were instructed to retell the story in their own words. A unidirectional microphone connected to the computer was used to record the participants' retellings. The task was administered using PowerPoint.

The effects of PMT and SFA treatment on discourse abilities was measured by the %CIUs produced by participants in language samples collected at Time Points A1, A2, and A3. CIUs are a rule-based method of scoring the informativeness of connected speech (Nicholas & Brookshire, 1993). CIUs are defined as words used in connected speech that are “intelligible in context, accurate in relation to the picture(s) or topic, and relevant to and informative about the content of picture(s) or the topic” (Nicholas & Brookshire, 1993, p. 348). Evidence suggests that %CIUs reliably reflect listeners' perception of a speaker's overall informativeness (Cameron et al., 2010; Carlomagno et al., 2011; Doyle et al., 1996).

Each language sample was independently transcribed orthographically in CHAT format (MacWhinney, 2000) by two graduate assistants. A third graduate assistant resolved 100% of the disagreements and generated a final orthographic transcription. Then, each finalized transcript was independently analyzed by two graduate assistants to generate the number of words and CIUs with which %CIUs were calculated. To determine the disagreement across raters, we estimated the root-mean-square deviations of %CIUs between raters at each time point, which ranged between 8 and 6.2 %CIUs. To maximize the precision of estimating %CIUs, 100% of the disagreements were resolved by a third graduate student. Raters were unaware of treatment assignment (i.e., SFA or PM) or time (e.g., A1, A2). Alphanumeric codes were used to identify the samples throughout the process.

Data Processing

Data were screened for missing values and univariate and multivariate outliers separately within each group. For identifying univariate outliers, we used *z* scores that had %CIU values greater than 3.3 *SD*s from the mean and that were disconnected from their respective distributions. For identifying multivariate outliers, we used Mahalanobis distance at an α level of .005. Cases that were flagged as multivariate outliers were further explored by estimated *z* scores of the differences in %CIUs across time points.

Data Analysis

The data were analyzed in SPSS 26 using mixed repeated-measures analysis of variance with two factors (group and time). Group was included as a between-subjects factor with two level (SFA, PMT), and time was included as a within-subject factor with three levels (A1, A2, and A3). In addition, the model included the interaction (Group × Time).

Results

Preliminary analysis was conducted to identify outliers. No univariate outliers were identified using *z* scores, which ranged from 2.05 to −1.56 across groups and time points. However, one participant in the phonomotor group (PM02SEA) was identified as a multivariate outlier using Mahalanobis distance at the .005 level. The participant's %CIUs at A1 was approximately 55% and improved to approximately 88% at A2, where it remained at A3. To further explore the performance of this participant, we *z* transformed the changes in %CIUs for each participant in the phonomotor group across time points and found that the change exhibited by this participant between A1 and A2 was 3.3 *SDs* from the mean change exhibited by the rest of the participants. Given that this participant was on the threshold of being an outlier, we opted to exclude this person's data from the main analysis to avoid an inflated Type I error rate. We do, however, return to a discussion of her performance below. Overall, data from 27 participants who received PMT and 29 participants who received SFA were included in the primary analysis.

To evaluate the assumption of sphericity, we used Mauchly's test. Based on Mauchly's test (Mauchly's $W = .962$), approximate $\chi^2(2) = 2.028$, $p = .363$, no violations of sphericity were noted. To evaluate the assumption of compound symmetry, we inspected the variance–covariance matrix of the dependent variable (i.e., %CIUs; see Table 2), and no gross differences in terms of variances and covariances were noted across time points.

Table 2. The variance/covariance matrix for each group along with means and standard deviations.

Time point	Phonomotor Therapy group		
1. A1	0.052		
2. A2	0.054 (.96*)	0.061	
3. A3	0.051 (.94*)	0.061 (.95*)	0.055
M	0.334	0.356	0.381
SD	0.228	0.248	0.235
	Semantic Feature Analysis group		
1. A1	0.059		
2. A2	0.056 (.96*)	0.062	
3. A3	0.051 (.94*)	0.054 (.96*)	0.054
M	0.343	0.373	0.341
SD	0.242	0.250	0.232

Note. Values in parentheses indicate Pearson product–moment correlations. A1 = pretreatment; A2 = posttreatment; A3 = 3 months posttreatment.

*Correlation is significant at the .01 level (two-tailed).

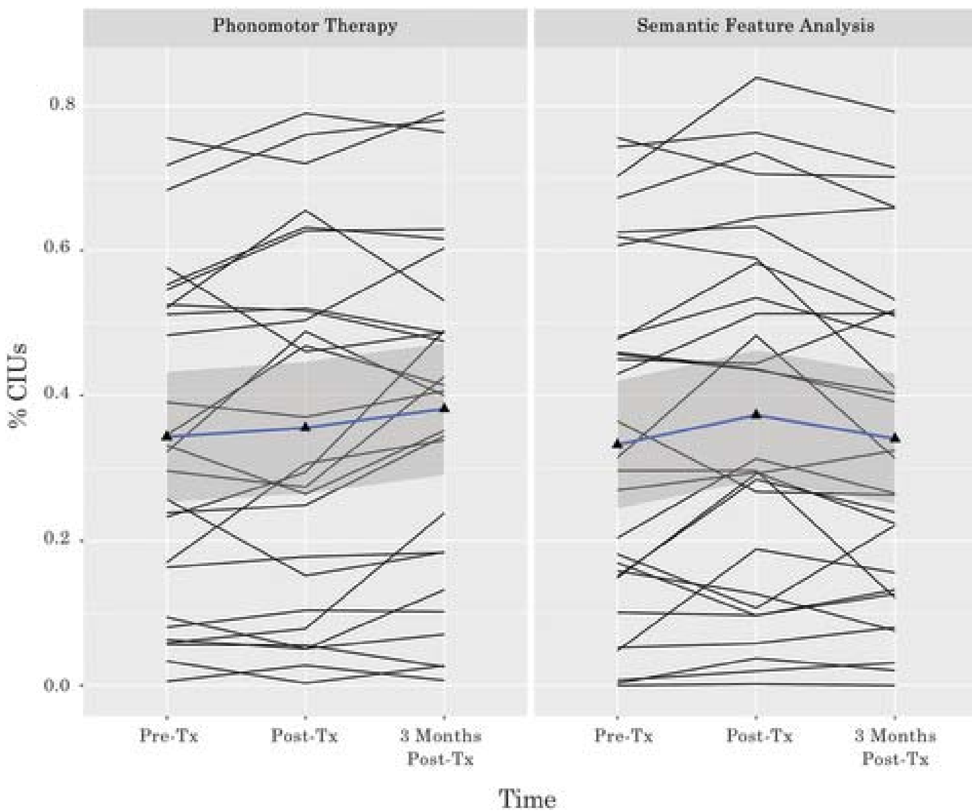

Figure 1. Individual participant data as a function of time and treatment group. Triangles indicate mean performance at each time point. The gray-shaded area corresponds to 95% confidence intervals. %CIUs = percentage of correct information units; Tx = treatment.

Primary Analysis

Significant results were found for the time main effect, Wilk's $\lambda = .875$, $F(2, 53) = 3.769$, $p = .029$, partial $\eta^2 = .125$. The interaction was also statistically significant, Wilk's $\lambda = .845$, $F(2, 53) = 4.877$, $p = .011$, partial $\eta^2 = .155$. The group main effect was not statistically significant, $F(1, 53) = .006$, $p = .94$ (see Figure 1 for a visual representation of the interaction).

Table 3. Pairwise comparisons of group performance at each time point.

Comparison	Mean %Δ[a]	95% CI	*df*	p^b
		Phonomotor Therapy group		
A1 vs. A2	2.17	–0.64, 4.97	26	.749
A1 vs. A3	4.78	1.61, 7.94	26	.028
A2 vs. A3	2.61	–3.33, 5.55	26	.479
		Semantic Feature Analysis group		
A1 vs. A2	3.00	0.26, 5.75	28	.198
A1 vs. A3	–0.25	–3.43, 2.94	28	> .99
A2 vs. A3	–3.25	0.67, 5.82	28	.091

Note. CI = confidence interval; A1 = pretreatment; A2 = posttreatment; A3 = 3 months posttreatment.

[a]Mean %Δ = difference in percentage of correct information units.
[b]Adjusted for multiple comparisons.

To better understand the significant interaction and answer the substantive questions of the study, group and time simple effects were explored by conducting a series of pairwise comparisons (see Table 3). To control for Type I error, familywise error rate across the pairwise comparisons was controlled using the Bonferroni procedure. For the PMT group, there was a statistically significant difference only between pretreatment and maintenance testing (A3−A1), with an average change of 4.8% (adjusted $p = .028$). None of the pairwise comparisons were statistically significant for the SFA group after controlling for Type I error due to multiple comparisons within the simple main effects analysis.

When comparing the levels of performance between groups at each time point using pairwise comparisons, there was no significant difference between the two groups at the pretreatment (A1), posttreatment (A2), or 3 months posttreatment (A3) time points (see Table 4).

Table 4. Mean percent correct information unit scores and standard deviations for Phonomotor Therapy (PMT) and Semantic Feature Analysis (SFA) groups and group differences for each time point.

Time point	PMT		SFA		Group difference	
	M	*SD*	*M*	*SD*	*t (df)*	*p*
A1	33.39%	22.84	34.31%	24.24	−0.145 (54)	.885
A2	35.56%	24.77	37.31%	24.98	−0.263 (54)	.981
A3	38.17%	23.53	34.06%	23.21	0.657 (54)	.514

Note. Unlike Table 3 where *p* values were adjusted, they are presented here unadjusted to avoid masking their differences. A1 = pretreatment; A2 = posttreatment; A3 = 3 months posttreatment.

Discussion

This study investigated the effects of intensive PMT and SFA on discourse production in individuals with aphasia. We compared discourse performance within and between treatment groups with equivalent amounts of treatment provided. The results revealed different outcome patterns for the two treatment approaches.

The within-group PMT results, comparing pretreatment (A1) to the posttreatment (A2) and maintenance (A3) periods, are consistent with previous data on PMT's effects on discourse level production (Silkes et al., 2018) in that participants did not show significant improvement from A1 to A2 but showed continued improvement after treatment ended such that there was a significant difference between A1 and A3. We have hypothesized (Silkes et al., 2018) that this delay in changes in communicative effectiveness reflects the time it takes to make enough improvement in the lexical networks that there is a measurable change in the number of word substitutions and repetitions, which influence the calculation of %CIUs. Unfortunately, we could not verify this through analysis of error types in the current data set due to resource limitations, so this remains a speculative explanation that should be further explored in the future. As predicted by PDP models of language processing, this improvement after treatment ended is possible because of the self-reinforcing nature of the lexical processing system; once appropriate connections are established, they continue to be strengthened through ongoing exposure to language, as a small set of phonemes is engaged in all language processing. This redundant recruitment of each phoneme across countless words and word positions facilitates massive reinforcement, allowing further growth once direct treatment ends.

The within-group SFA analysis showed a different pattern, with nonsignificant improvements occurring in discourse from A1 to A2 but a return to near-baseline levels at A3. If SFA is effective because it trains the use and proceduralization of a retrieval process (Haarbauer-Krupa et al., 1985), then this strategy should lead to widespread improvements at the discourse level and should be maintained, or even strengthened, with practice over time. Finding limited generalization and a return to baseline at maintenance, as reported here, suggests that any treatment effects obtained for discourse following SFA are more likely related to intrinsic changes within the lexical–semantic networks, as described earlier. According to this mechanism, using SFA to strengthen connections between semantic and lexical representations may serve to make those particular lexical–semantic elements more readily retrieved, along with others that are semantically related; however, the elements that are reinforced in SFA are not shared universally, so this improvement is not widespread and does not self-reinforce in a robust manner through ongoing language use after treatment ends. The result is that treatment effects for untrained items are not likely to be maintained over time, as was seen in these data.

While our findings for PMT are consistent with prior literature, those for SFA are not; previous reports on SFA have shown significant improvements in discourse measures immediately after treatment, whereas this study found a nonsignificant improvement at that time point. It is possible that this is due to differences in sample size, as there are a large number of participants involved in this study as compared with small single cases reported in the past. Alternatively, it might seem that these differences could be the product of different discourse elicitation procedures used across studies; the SRP involves the participant hearing a story and recalling it for retell, while other procedures often involve self-generated stories based on a common referent. We do not believe that this is the case, however, since McNeil et al. (2007) did not find any significant differences for %CIUs across language sampling procedures (Figure 2, p. 785).

The two treatments tested here presented with different patterns from A2 to A3, with PMT showing increased growth while SFA returned to near-baseline levels. Despite these differences, however, they did not significantly differ from each other in mean accuracy at A3. This finding is likely related to the large variance noted in both groups. We considered whether aphasia severity may have contributed to this large variance, but visual inspection of the individual data does not reveal any striking differences in the degree or direction of change over time between participants in either group who are more or less severe, as reflected by their measure of %CIUs at A1. Future analyses that consider aphasia severity, individual response to treatment at the level of single word retrieval, and other individual factors in conjunction with treatment type may shed further light on differences between these two treatments.

It is worth specifically considering the one PMT participant who responded better than all of the other PMT participants, leading to her removal from the primary analysis. We were curious about factors that may have influenced her treatment response. It is unclear what made her a particularly good responder to this treatment. While she was relatively young and well educated and had a relatively mild aphasia, none of these factors have been shown to be good predictors of treatment response for PMT (Hunting Pompon et al., 2017). We can report that she is a person who is very engaged in her world. She lives independently, with an active social life, no mental health concerns, high levels of involvement in her community, and high efforts at ongoing language use (e.g., writing poetry). She had an excellent ability to engage in the treatment process (particularly Socratic questioning) and keen interest and curiosity about language and its recovery. It is unclear, however, whether any of these characteristics played a role in her treatment response. We can only speculate that perhaps there was something intrinsic to her approach to treatment or her ability to engage with her environment that led her to make changes more quickly than other participants, leading to a high degree of change by A2 that was not apparent in the rest of the group until later posttreatment.

Limitations

The data presented here provide interesting insights into the potential of PMT and SFA to have an impact on discourse-level communication, reinforced by the PMT data being consistent with data from a prior study. There are, however, a few limitations. The first to consider regards our ability to understand the mechanism(s) responsible for acquisition and generalization in response to SFA. As discussed earlier, SFA may work through strengthening lexical networks but may also work through the application of systematic description to assist when word retrieval fails or through the automatization of a systematic lexical search method. Although we interpret the data presented here as indicative that these were not the active mechanisms of change in this study, these method-based explanations for change were not directly investigated. In addition, participants in the SFA group were never encouraged to use the semantic descriptions as a compensatory strategy during the course of treatment. It is possible that including an aspect of explicit compensatory training in the SFA protocol might lead to a different outcome than that reported here. A second limitation is that, while the %CIU measure reported here replicated what was found in Silkes et al. (2018), this is not a complete replication of that earlier study. Due to resource limitations, we were unable to evaluate changes in CIUs per minute for this study, which were previously found to change significantly from pretreatment to immediately posttreatment and then to maintain to 3 months posttreatment. Future investigations should further explore whether this change in communication efficiency, as measured by CIUs per minute, can be replicated in another large sample.

Conclusion

The data presented here suggest that both SFA and PMT can lead to small positive impacts on discourse-level communication in PWA immediately after treatment; in this data set, however, they differed in their effects 3 months posttreatment relative to pretreatment performance. In light of the high individual variability noted in responses to treatment for both groups, further research is warranted to determine which treatment is likely to be most effective for whom and how to maximize treatment effects and carryover to discourse. Regardless, the finding that it is possible to influence discourse by treating at a lower level of processing, such as phonology or semantic features, highlights the importance of developing and using model-driven aphasia treatments and provides evidence that it is possible for aphasia treatments focused at the lexical and/or sublexical level to generalize to discourse.

Acknowledgments

This research was partially supported by U.S. Department of Veterans Affairs Office of Rehabilitation Research and Development Merit Review Grant C6572R. We are especially grateful to the study participants. We would like to thank Megan Moldestad and Wesley Allen for delivering the treatment protocol, Janaki Torrence for providing training and ongoing support for student transcriptions, and Kerry Lam for administrative support of this project.

References

Antonucci, S. M. (2009). Use of semantic feature analysis in group aphasia treatment. *Aphasiology*, 23(7–8), 854–866. https://doi. org/10.1080/02687030802634405

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. (2010). Semantic feature analysis treatment for aphasic word retrieval impairments: What's in a name. *Topics in Stroke Rehabilitation*, 17(6), 411–422. https://doi.org/10.1310/tsr1706-411

Boyle, M. (2011). Discourse treatment for word retrieval impairment in aphasia: The story so far. *Aphasiology*, 25(11), 1308–1326. https://doi.org/10.1080/02687038.2011.596185

Boyle, M., & Coelho, C. A. (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

Brookshire, C. E., Conway, T., Pompon, R. H., Oelke, M., & Kendall, D. L. (2014). Effects of intensive phonomotor treatment on reading in eight individuals with aphasia and phonological alexia. *American Journal of Speech-Language Pathology*, 23(2), S300–S311. https://doi.org/10.1044/2014_AJSLP-13- 0083

Brookshire, R. H., & Nicholas, L. E. (1993). *Discourse Comprehension Test*. Communication Skill Builders.

Cameron, R. M., Wambaugh, J. L., & Mauszycki, S. C. (2010). Individual variability on discourse measures over repeated sampling times in persons with aphasia. *Aphasiology,* 24(6–8), 671–684. https://doi.org/10.1080/02687030903443813

Carlomagno, S., Giannotti, S., Vorano, L., & Marini, A. (2011). Discourse information content in non-aphasic adults with brain injury: A pilot study. *Brain Injury*, 25(10), 1010–1018. https://doi.org/10.3109/02699052.2011.605097

Coelho, C. A., McHugh, R. E., & Boyle, M. (2000). Semantic feature analysis as a treatment for aphasic dysnomia: A replication. *Aphasiology*, 14(2), 133–142. https://doi.org/10.1080/ 026870300401513

Davis, L. A., & Stanton, S. T. (2005). Semantic feature analysis as a functional therapy tool. *Contemporary Issues in Communication Sciences and Disorders*, 32, 85–92. https://doi.org/10.1044/ cicsd_32_F_85

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review,* 93(3), 283–321. https://doi.org/10.1037/0033-295X.93.3.283

DeLong, C., Nessler, C., Wright, S., & Wambaugh, J. (2015). Semantic feature analysis: Further examination of outcomes. *American Journal of Speech-Language Pathology*, 24(4), S864–S879. https://doi.org/10.1044/2015_AJSLP-14-0155

Doyle, P. J., McNeil, M. R., Park, G. H., Goda, A. J., Rubenstein, E., Spencer, K., Carroll, B., Lustig, A., & Szwarc, L. (2000). Linguistic validation of four parallel forms of a story retelling procedure. *Aphasiology*, 14(5–6), 537–549. https://doi.org/10.1080/ 026870300401306

Doyle, P. J., McNeil, M. R., Spencer, K. A., Goda, A. J., Cottrell, K., & Lustig, A. P. (1998). The effects of concurrent picture presentations on retelling of orally presented stories by adults with aphasia. *Aphasiology,* 12(7–8), 561–574. https://doi.org/ 10.1080/02687039808249558

Doyle, P. J., Tsironas, D., Goda, A. J., & Kalinyak, M. (1996). The relationship between objective measures and listeners' judgements of the communicative informativeness of the connected discourse of adults with aphasia. *American Journal of Speech-Language Pathology*, 5(3), 53–60. https://doi.org/10.1044/ 1058-0360.0503.53

Efstratiadou, E. A., Papathanasiou, I., Holland, R., Archonti, A., & Hilari, K. (2018). A systematic review of semantic feature analysis therapy studies for aphasia. *Journal of Speech, Language, and Hearing Research*, 61(5), 1261–1278. https://doi.org/ 10.1044/2018_JSLHR-L-16-0330

Falconer, C., & Antonucci, S. M. (2012). Use of semantic feature analysis in group discourse treatment for aphasia: Extension and expansion. *Aphasiology*, 26(1), 64–82. https://doi.org/ 10.1080/02687038.2011.602390

Gravier, M. L., Dickey, M. W., Hula, W. D., Evans, W. S., Owens, R. L., Winans-Mitrik, R. L., & Doyle, P. J. (2018). What matters in semantic feature analysis: Practice-related predictors of treatment response in aphasia. *American Journal of Speech Language Pathology*, 27(1S), 438–453. https://doi.org/10.1044/ 2017_AJSLP-16-0196

Haarbauer-Krupa, J., Moser, L., Smith, G., Sullivan, D. M., & Szekeres, S. F. (1985). Cognitive rehabilitation therapy: Middle stages of recovery. In M. Ylvisaker (Ed.), *Head injury rehabilitation: Children and adolescents* (pp. 287–310). College-Hill Press.

Hunting Pompon, R., Bislick, L., Elliott, K., Madden, E. B., Minkina, I., Oelke, M., & Kendall, D. (2017). Influence of linguistic and nonlinguistic variables on generalization and maintenance following phonomotor treatment for aphasia. *American Journal of Speech-Language Pathology*, 26(4), 1092–1104. https://doi.org/ 10.1044/2017_Ajslp-16-0175

Kendall, D., Moldestad, M. O., Allen, W., Torrence, J., & Nadeau, S. (2019). Phonomotor versus semantic feature analysis treatment for anomia in 58 persons with aphasia: A randomized controlled trial. *Journal of Speech, Language, and Hearing Research*, 62(12), 4464–4482. https://doi.org/10.1044/2019_JSLHR-L18-0257

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

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., Rosenbek, J. C., Heilman, K. M., Conway, T., Klenberg, K., Gonzalez Rothi, L. J., & 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

Kiran, S. (2008). Typicality of inanimate category exemplars in aphasia treatment: Further evidence for semantic complexity. *Journal of Speech, Language, and Hearing Research,* 51(6), 1550–1568. https://doi.org/10.1044/1092-4388(2008/07-0038)

Kiran, S., & Johnson, L. (2008). Semantic complexity in treatment of naming deficits in aphasia: Evidence from well-defined categories. *American Journal of Speech-Language Pathology*, 17(4), 389–400. https://doi.org/10.1044/1058-0360(2008/06-0085)

Knoph, M. I. N., Lind, M., & Simonsen, H. G. (2015). Semantic feature analysis targeting verbs in a quadrilingual speaker with aphasia. *Aphasiology*, 29(12), 1473–1496. https://doi.org/ 10.1080/02687038.2015.1049583

Knoph, M. I. N., Simonsen, H. G., & Lind, M. (2017). Crosslinguistic transfer effects of verb-production therapy in two cases of multilingual aphasia. *Aphasiology*, 31(12), 1482–1509. https://doi.org/10.1080/02687038.2017.1358447

MacWhinney, B. (2000). *The CHILDES Project: Tools for analyzing talk* (3rd ed.). Erlbaum.

Massaro, M., & Tompkins, C. A. (1992). Feature analysis for treatment of communication disorders in traumatically brain injured patients: An efficacy study. *Clinical Aphasiology*, 22, 245–256.

McNeil, M. R., Doyle, P. J., Fossett, T. R. D., Park, G. H., & Goda, A. J. (2001). Reliability and concurrent validity of the information unit scoring metric for the story retell procedure. *Aphasiology,* 15(10–11), 991–1006. https://doi.org/10.1080/ 02687040143000348

McNeil, M. R., Sung, J. E., Yang, D., Pratt, S. R., Fossett, T. R. D., Doyle, P. J., & Pavelko, S. (2007). Comparing connected language elicitation procedures in persons with aphasia: Concurrent validation of the story retell procedure. *Aphasiology*, 21(6–8), 775–790. https://doi.org/10.1080/02687030701189980

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. (2015). Neuroplastic mechanisms of language recovery after stroke. In J. Tracy, B. Hampstead, & K. Sathan (Eds.), *Cognitive plasticity in neurologic disorders* (pp. 61–84). Oxford University Press.

Nicholas, L. E., & Brookshire, R. E. (1993). A system for quantifying the informativeness and efficiency of the connected speech of adults with aphasia. *Journal of Speech and Hearing Research,* 36(2), 338–350. https://doi.org/10.1044/jshr.3602. 338

Oh, S. J., Eom, B., Park, C., & Sung, J. E. (2016). Treatment efficacy of semantic feature analyses for persons with aphasia: Evidence from meta-analyses. *Communication Sciences and Disorders*, 21(2), 310–323. https://doi.org/10.12963/csd. 16312

Peach, R. K., & Reuter, K. A. (2010). A discourse-based approach to semantic feature analysis for the treatment of aphasic word retrieval failures. *Aphasiology*, 24(9), 971–990. https://doi.org/ 10.1080/02687030903058629

Quique, Y. M., Evans, W. S., & Dickey, M. W. (2018). Acquisition and generalization responses in aphasia naming treatment: A meta-analysis of semantic feature analysis outcomes. *American Journal of Speech-Language Pathology*, 28(1S), 230–246. https:// doi.org/10.1044/2018_AJSLP-17-0155

Rider, J. D., Wright, H. H., Marshall, R. C., & Page, J. L. (2008). Using semantic feature analysis to improve contextual discourse in adults with aphasia. American *Journal of Speech-Language Pathology*, 17(2), 161–172. https://doi.org/10.1044/1058-0360 (2008/016)

Silkes, J. P., Fergadiotis, G., Hunting Pompon, R., Torrence, J., & Kendall, D. L. (2018). Effects of phonomotor treatment on discourse production. *Aphasiology*, 33(2), 125–139. https:// doi.org/10.1080/02687038.2018.1512080

Storkel, H. L., Armbrüster, 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*. Routledge.

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

Wambaugh, J. L., & Ferguson, M. (2007). Application of semantic feature analysis to retrieval of action names in aphasia. *Journal of Rehabilitation Research & Development*, 44(3), 381–394. https://doi.org/10.1682/Jrrd.2006.05.0038

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