

Application of semantic feature analysis to retrieval of action names in aphasia

Julie L. Wambaugh, PhD;^{1-2*} Morelia Ferguson, MS²

¹Department of Veterans Affairs, Salt Lake City Healthcare System, Salt Lake City, UT; ²University of Utah, Salt Lake City, UT

Abstract—This investigation examined the effects of a semantic feature training procedure on retrieval of action names in a participant with anomic aphasia. Treatment was applied sequentially across two sets of action names in the context of a multiple baseline design across behaviors. Treatment effects were evaluated in terms of naming of trained and untrained actions. Discourse production effects were also examined with respect to verbal productivity, informativeness, and production of nouns and verbs. Increased accuracy of naming was observed for both sets of trained action names, with increases being maintained at 6 weeks posttreatment. However, accuracy of responding did not reach preestablished criterion levels. Repeated exposure to stimulus items without training resulted in unstable and temporary increases in naming accuracy. No changes were observed in accuracy of naming of untrained actions that were measured only at pre- and posttreatment intervals. Increases in verbal productivity and informativeness in discourse production were associated with the treatment.

Key words: anomia, aphasia, language therapy, rehabilitation, semantic feature analysis, single-subject design, treatment, verbs, word-finding, word retrieval.

INTRODUCTION

Anomia, or difficulty with word-finding, is a defining feature of aphasia that crosses aphasia type [1]. Word-retrieval deficits may be observed with all grammatical word forms (e.g., nouns, verbs, adjectives), and persons with aphasia may have more difficulty with retrieval of one grammatical class relative to another [2–4].

Retrieval of object names has received significantly more focus than retrieval of other grammatical form classes in terms of rehabilitation research. However, an increasing number of investigations have addressed treatment of verb retrieval [5–9]. Targeting verb retrieval is potentially an important component of aphasia rehabilitation because verbs not only carry critical meaning but also have important functions in the structural formulation of sentences [10].

Although a limited number of participants have been studied across relatively few investigations, it appears that action-naming may be improved by therapies similar to those used in the treatment of object-name retrieval. For example, McNeil and colleagues applied Lexical-Semantic

Abbreviations: CDC = conservative dual-criterion, CIU = correct information unit, %CIU = percent CIU, L-SAIT = Lexical-Semantic Activation Inhibition Treatment, OANB = Object and Action Naming Battery, PALPA = Psycholinguistic Assessment of Language Processes in Aphasia, PAS = predicate argument structure, PCT = phonological cueing treatment, PICA = Porch Index of Communicative Abilities, SCT = semantic cueing treatment, SD = standard deviation, SEM = standard error of measurement, SFA = Semantic Feature Analysis, TAAWF = Test of Adolescent/Adult Word Finding, VAST = Verb and Sentence Test, wpm = words per minute.

*Address all correspondence to Julie L. Wambaugh, PhD; 500 Foothill Blvd (151A), Salt Lake City, UT 84148; 801-582-1565, ext 1363; fax: 801-584-5621.

Email: julie.wambaugh@health.utah.edu

DOI: 10.1682/JRRD.2006.05.0038

Activation Inhibition Treatment (L-SAIT) to retrieval of verbs, nouns, and adjectives with a participant with anomia [11]. L-SAIT, which entailed production of antonyms and synonyms with semantic, gestural, written, and phonologic cueing, improved retrieval of antonyms for all trained grammatical forms, with no generalization to untrained items within and across grammatical forms.

Cueing treatments have also been studied by Wambaugh and colleagues in the treatment of object names [12–13] and action names [8,14]. Positive treatment effects were found for both a phonological cueing treatment (PCT) and a semantic cueing treatment (SCT), with increases in naming accuracy being observed more consistently in the treatment of object names. More specifically, participants whose verb-retrieval deficits appeared to stem from relatively severe semantic processing deficits did not benefit from either PCT or SCT. The investigators suggested that the relatively limited information provided by the cues may have been insufficient to improve verb retrieval for participants with severe deficits.

The positive effects of treatments for verb retrieval in aphasia have largely been restricted to improvements in naming of trained items. This has been the case for therapies designed specifically to treat verbs [5–6,9,15], and for those treatments derived from or applied to different grammatical form classes [11,14,16]. This lack of generalization to untrained verbs, even to semantically related verbs, concurs with findings from studies of treatment of object name retrieval, in which response generalization has been a relatively rare occurrence [17].

Although improvements in naming have been restricted to trained verbs, changes in the production of sentences and/or discourse have been reported following verb-retrieval training [6–7,9,18]. However, findings regarding such changes have been mixed, both within and across studies examining pre- and posttreatment discourse abilities. For example, Marshall combined semantic tasks, orthographic cueing, gestural tasks, and verbal production tasks in treating a patient with a phonological-level impairment of verb production (unimpaired written verb production) [18]. Following treatment, limited improvements in verb retrieval and production of verb-plus-argument structure in response to questions were noted, but similar changes were not observed in story-retelling tasks.

Marshall et al. reported statistically significant improvements in sentence production following a semantically oriented treatment with a participant with Broca's aphasia

and a selective verb-retrieval deficit [6]. Similarly, Raymer and Ellsworth found improvements in sentence production accuracy in response to a semantic verb-retrieval treatment with a participant with nonfluent aphasia and mild verb-retrieval impairment [7]. They also found phonological treatment and rehearsal to have similar effects on sentence-production accuracy.

In production of lexical items in discourse, Webster et al. found an increased proportion of production of lexical verbs in two-argument structures in narratives [9]. This finding was rather surprising, given that the investigators found no overall improvement in verb retrieval (based on performance on a verb- and noun-retrieval test). In contrast, McNeil et al. [11] reported no changes in verbal productivity and informativeness [19] during discourse production. However, McNeil et al. did not examine production of types of lexical items within discourse [11]. One possible source for the differences in generalization findings in the two preceding investigations, beyond measurement issues, may relate to the types of training used.

As part of their treatment, Webster et al. employed a training strategy in which the participant was asked to generate words that filled thematic roles or were associated with the target verb [9]. This strategy appears similar to a treatment approach, termed *Semantic Feature Analysis* (SFA) [20–21], that has been applied to the retrieval of object names. SFA is used to guide the patient in identifying important semantic features of the target word. This approach is believed to help activate the semantic network that surrounds the target word to aid in its retrieval [20]. In the process of identifying features of the target item, nontargeted, and semantically related words may benefit in terms of retrieval because they share features that are being accessed or retrained. Additionally, generalization may be facilitated by application of the compensatory strategy of systematically retrieving features, which may serve a self-cueing function. SFA has been demonstrated to result in improved retrieval of trained nouns and has been shown to promote retrieval of untrained nouns [20–23]. SFA has also been associated with improvements in content production during discourse. For example, Boyle found a 75 percent increase in efficiency of correct information unit (CIU) production in discourse for one participant and a 92 percent increase in mean CIU production for a second participant following treatment with SFA [21]. The effects of SFA

have not been examined with respect to action name retrieval.

Generalization is a current and important challenge in the development of effective word-retrieval treatments. Lack of generalization to untrained items has been observed more frequently than has generalization in the treatment of object-naming (see Nickels [17] for a review). However, advances have been made in generalization *within* semantic category [21,24–26]. That is, generalized improvement in untrained object-naming has been reported for therapies involving strengthening of semantic information/networks [21,25–26], as well as for semantic category rhyme therapy [24]. Such gains have been limited to untrained items that belong to the same semantic category as trained items.

How semantic category membership may pertain to the treatment of verb retrieval in aphasia is not clear. Numerous semantic categories of verbs have been proposed. For example, Levin describes more than 40 “semantically coherent” classes of verbs, with many more subcategories within class [27]. Currently, we are not aware of any agreed-upon methods for categorizing verbs in terms of relatedness. Furthermore, a complex, not fully specified relationship exists between the meaning of verb and its syntactic behavior [27]. Verbs that are closely related in meaning may differ in terms of acceptable alternations [10,27]. That is, even verbs that are clearly similar with respect to core meanings may express their arguments differently. For example, “fill” and “pour” are closely related, but their realization of arguments differs: “She filled the glass with water” is allowable but “She poured the glass with water” is not, and “She poured water into the glass” is acceptable but “She filled water into the glass” is not [27]. Argument structure has been shown to be important in terms of verb retrieval [28], although the interaction of semantic class and diathesis alternation has not been explored relative to treatment in aphasia. Additionally, for verb retrieval treatments involving arguments and thematic roles, generalization may likely be affected by the semantic relatedness of the nouns expressed in the arguments.

Controlling for semantic category factors that may affect generalization was beyond the scope of our investigation. Such control would require extensive normative research. Instead, we chose to focus on potentially facilitating generalization by applying a treatment method that may be useful as a compensatory strategy. As reviewed by Nickels [17], compensatory strategies have been shown to result in generalized responding for some participants.

As noted previously, SFA may serve as a compensatory strategy, which may account for improved responding to items within, as well as beyond, trained semantic categories. In fact, Boyle suggests that generalization to untrained items following application of SFA resulted not only from stimulation of the semantic network but also from “accessing items from a variety of semantic categories in a structured, methodical way over and over again” [21, p. 246].

Therefore, we designed our current investigation to further study the effects of treatment for action name retrieval in aphasia. To promote generalization, we chose SFA for study because of its promising generalization findings with object-name retrieval. Additionally, in light of the array of information carried by verbs (e.g., information concerning lexical semantics, thematic roles, argument structure), SFA appeared amenable, with modification, to systematically providing a variety of cues pertinent to verbs. Although SFA has been applied successfully with object-name retrieval, we could not assume that its effects would be similar when applied to action name retrieval. Consequently, our study was also designed to determine if SFA would result in acquisition of trained action names.

This investigation addresses specific experimental questions: Will SFA applied to action names—

1. Result in increased accuracy of confrontation naming of trained actions?
2. Result in increased accuracy of confrontation naming of untrained action names?
3. Be associated with increased amount and efficiency of production of content during discourse?
4. Be associated with changes in relative production of verbs and nouns during discourse?

METHODS

Participant

The participant was a 74-year-old Caucasian female, who was 4 years, 2 months, postonset of a single, left parietal cerebrovascular accident. She was a right-handed native-English speaker who had 13 years of formal education and was retired from 25 years of secretarial work. She was referred as a potential participant for this investigation by the director of a university speech/language clinic who had been made aware of the study. The participant had a negative history for other neurological conditions, mental illness, and alcohol/substance abuse according

to self-report. She passed a pure tone hearing screening at 500, 1,000, and 2,000 Hz at 45 dB in the left ear unaided.

Specific records of the participant's past speech/language therapy were not available. According to her daughter, the participant received approximately 4 weeks of home health speech/language services after she returned home following hospitalization. At approximately 4 months poststroke, she again received approximately 1 month of speech/language services. According to the participant, her previous speech/language therapy did not include SFA. The participant did not receive any other speech/language services during this investigation.

As shown in **Table 1**, the participant presented with moderate, anomic aphasia, as evidenced by performance on the Western Aphasia Battery [29] and the Porch Index of Communicative Abilities (PICA) [30]. Significant word-retrieval difficulties were present as reflected by a total raw score of 19 on the Test of Adolescent/Adult Word Finding (TAAWF) [31]. No motor speech disorders were evident as defined by Duffy [32], and her single word intelligibility was 92 percent [33].

We conducted various lexical processing assessments to determine the nature of her deficit and to identify potential changes in lexical processing posttreatment (**Table 1**). Responses to subtests from the Psycholinguistic Assessment of Language Processes in Aphasia (PALPA) [34], Verb and Sentence Test (VAST) [35], an Object and Action Naming Battery (OANB) [36], and the TAAWF revealed deficits in semantic association and verb and noun comprehension, suggesting difficulty in semantic processing. A large percentage (60%) of the participant's naming errors on the TAAWF and OANB took the form of no response. However, the majority of her classifiable errors were semantic paraphasias (e.g., hammock → swing, leaf → flower, squeezing → lemon) or descriptions (e.g., dripping → drops are coming down) (**Table 1**). Such errors support the notion that she had deficits at the level of semantic processing. Her sentence production deficits were also in keeping with a theorized deficit in verb semantic processing [37]. Additionally, the participant demonstrated difficulty in judging rhymes and occasionally produced phonemic and mixed paraphasias, indicating possible difficulties in processing at the phonological level.

Experimental Stimuli

Confrontation Naming Stimuli

The participant was asked to name 100 black and white line drawings depicting actions from OANB on

two separate occasions. We then selected 40 of the 100 items to be experimental stimuli. The words were divided into four lists of 10 items each (**Appendix 1**, available online only at <http://www.rehab.research.va.gov/>). Each list was matched as closely as possible for factors that could influence retrieval or production: frequency, age of acquisition, number of syllables, familiarity, imageability, and visual complexity [36,38]. Argument structure was balanced across lists in terms of number of one- and two-place argument verbs represented in each list (**Appendix 1**).

Discourse Content Stimuli

We obtained samples of narrative and procedural discourse using procedures developed by Nicholas and Brookshire [19]. The procedures involve requesting six picture descriptions (single and sequenced pictures), two descriptions of procedures (e.g., "How do you wash the dishes?") and two descriptions of personal information (e.g., "What do you do on a Sunday?").

Pre- and posttreatment discourse samples were audio recorded and transcribed orthographically. Transcriptions were verified by a master's student not involved in the investigation, who was experienced in transcription of aphasic discourse.

We scored each sample for production of CIUs using procedures detailed by Nicholas and Brookshire [19]. As described by Nicholas and Brookshire, CIUs are "words that are intelligible in context, accurate in relation to the picture(s) or topic, and relevant to and informative about the content of the picture(s) or the topic. Words do not have to be used in a grammatically correct manner to be included in the correct information count" [19, p. 348]. CIUs were coded from the transcripts independently by two examiners (second author and master's student) with audio tape supplementation as necessary. Any disagreements in point-to-point CIU scoring were resolved by agreement of a third examiner (Dr. Wambaugh).

The test-retest stability of Brookshire and Nicholas' discourse elicitation method has been demonstrated to be good for the scoring of words per minute (wpm) and "percent CIU" (%CIU) [39]. Specifically, Brookshire and Nicholas obtained discourse samples from a group of 20 adults with aphasia on three separate sampling occasions; two samples were obtained on the same day and a third sample was obtained 7 to 10 days later [39]. For the group, the mean number of wpm ranged from 81 to 85, with a standard error of measurement (SEM) from time

Table 1.

Pre- and posttreatment lexical processing assessment results for 74-year-old female participant.

Measure	Pretreatment	Posttreatment
Test of Adolescent/Adult Word Finding (TAAWF)		
Raw Score	19/107	40/107
Comprehension (%)	75	96*
Object and Action Naming Battery (OANB)		
Nouns Only (Number Correct)	50/81	64/81
Types of Errors on TAAWF and OANB: Total Errors (%), Excluding		
No Responses		
Semantic Paraphasias	57.5	—
Phonemic Paraphasias	2.0	—
Mixed Paraphasias	6.5	—
Unrelated Responses	6.5	—
Descriptions	19.0	—
Perseverations	8.5	—
Verb and Sentence Test		
Verb Comprehension	35/40	35/40
Sentence Comprehension	29/40	32/40
Grammaticality Judgment	38/40	38/40
Action Naming	8/40	18/40
Filling in Finite Verbs	4/10	2/10
Filling in Infinitives	2/10	4/10
Sentence Construction	9/20	14/20
Porch Index of Communicative Ability		
Overall	11.74	11.94
Percentile	60th	64th
Psycholinguistic Assessment of Language Processing in Aphasia		
Word Rhyme Judgment	48/60	47/60
Word Semantic Association		
High Imageability	8/15	9/15
Low Imageability	8/15	6/15
Auditory Comprehension, Verbs/Adjectives	35/41	34/41
Western Aphasia Battery		
Aphasia Quotient	67.7	—
Diagnostic Category	Anomic	—
Assessment of Intelligibility of Dysarthric Speech		
Word—Transcription (%)	92	—
Coloured Progressive Matrices		
Overall	18	—
Percentile	10th	—

*Considered within normal limits according to guidelines of test published.

1 to time 3 of 5.5 wpm. Mean %CIU was 63 (at all sampling times), with a SEM of 2.1 from time 1 to time 3.

All words that were included in the CIU count were then coded in three grammatical form classes: nouns (including pronouns), verbs, and “other” (e.g., adjectives, adverbs, articles, conjunctions).

Experimental Design

A single-subject, multiple baseline design across behaviors was used. This design allowed for the examination of both acquisition and response generalization effects of treatment. The behavior of interest (i.e., confrontation naming of actions) was measured repeatedly

prior to the initiation of treatment during the baseline phase for three lists of action pictures. Then, treatment was applied to one of the lists of items while repeated probing of all lists of items continued. When the criterion for terminating treatment was reached (i.e., 80% correct in three consecutive probes, or completion of 12 treatment sessions), treatment was withdrawn from the first list of items and was applied to a second list. Probing was conducted continuously until treatment was terminated with the second list and was again conducted at follow-up intervals.

The third list of items and an additional fourth list of actions were used for assessing the effects of generalization to items with repeated and limited exposure. Specifically, List 3 was probed continually along with Lists 1 and 2. List 4 was probed only once during the baseline phase and at the completion of treatment. We included List 4 to permit comparison of potential generalization effects with List 3. Repeated exposure of items to be named has been demonstrated to improve naming performance, even when items have not received treatment [40]. Such improvement is a confound to the determination of generalization. Therefore, we considered including a group of generalization items that had limited exposure important. We expected that if repeated exposure, but not generalization, contributed to improved performance on untrained lists, then performance on List 3 would improve, but performance on List 4 would not change. However, if generalization contributed to improved performance on untrained lists, then performance on both Lists 3 and 4 would improve, with differences in magnitude being attributed to exposure (if performance on List 3 was higher).

Baseline Phase

Responses to the 30 action pictures (Lists 1–3) were repeatedly measured during the baseline phase. Each of the 30 items was presented one at a time in random order and the participant was asked to “use an action word to describe the picture.” We prompted the participant again with the same phrase if she used a noun instead of a verb (we provided this prompt a maximum of three times per probe session). We allowed a response interval of 30 seconds, and provided no feedback other than general encouragement (e.g., you’re working hard, good job, etc). A 30-second response window was provided because accuracy of retrieval was of greater interest than latency of responding. Furthermore, because SFA may

function as a compensatory strategy for retrieval, it was desired that sufficient time be provided for the feature strategy to be employed. A score of 7 or higher, using a multidimensional scoring system [13] (**Appendix 2**, available online only at <http://www.rehab.research.va.gov>) was considered a correct response.

The remaining 10 items, which we used to control for exposure effects, were presented separately from the other three sets, in random order, on one occasion prior to treatment. All probes were audio and video recorded throughout the study.

A relatively stable baseline was established for the lists prior to treatment. The a priori criterion for stability was no more than 20 percent variability in responding across three consecutive probes immediately preceding the application of treatment for that list.

Treatment Phase

Probes of naming performance, identical to those conducted during baseline, were completed at the beginning of each session prior to treatment.

Maintenance and Follow-Up Phase

Maintenance of naming of the initially trained word list was measured during subsequent training of the second list. Follow-up probes were conducted for the sets of action pictures at 2 and 6 weeks posttreatment.

Posttreatment testing was conducted immediately following the completion of treatment: the VAST, TAAWF, PICA, OANB (nouns only), and subtests of the PALPA were readministered. Additionally, discourse samples were obtained immediately following treatment.

Treatment

We used a modified version of SFA to accommodate the use of verbs instead of nouns as the pictured target stimuli. Specifically, we employed the basic training procedure used by Boyle and Coelho [20], but the features we targeted were appropriate for verb retrieval. We used a semantic feature diagram (**Appendix 3**, available online only at <http://www.rehab.research.va.gov>) to elicit the following features from the participant:

1. Subject (“Who usually does this?”).
2. Purpose of action (“Why does this happen?”).
3. Part of body or tool used to carry out action (“What part of the body or what tool is used to make this happen?”).
4. Description of physical properties (“Tell me what it looks like.”).

5. Usual location (“Where does this action usually take place?”).
6. Related objects or actions that reminded the participant of target verb (“What does it make you think of?”).

We wanted the features in this investigation to target not only lexical semantic information but also information concerning thematic roles. We also considered adhering as closely as possible to the original SFA features important. The SFA features of “properties,” “location,” and “association” were retained because these features appeared useful for eliciting lexical semantic information. To address thematic role information, we included “subject” to attempt to target the role of agent/experiencer and “purpose” and item “used to carry out the action” to attempt to elicit features related to theme/patient (of course, these features also are pertinent to eliciting lexical semantic information).

A picture of the targeted action was placed in the center of the diagram. We guided the participant in the production of semantically related words for each of the six features corresponding to the targeted illustration by asking the six questions just described. Each feature was addressed, one at a time, and the same order of features was used for each item. After providing the question/statement pertaining to the feature, the clinician wrote the participant’s responses on the diagram. Following elicitation of all six features, the clinician asked the participant to name the action being carried out in the picture. If the participant was unable to respond appropriately, the clinician provided a correct verbal response, requested a verbal repetition, and then reviewed all the features again. If the patient responded correctly, the clinician provided feedback and presented the next picture. (See **Appendix 3** for examples of actual responses at <http://www.rehab.research.va.gov/>.)

One presentation of the 10 treatment items constituted one trial. The participant typically completed one trial during one treatment session. Treatment sessions were conducted three times a week for 45 to 60 minutes a session (including probing prior to treatment). We should note that participant illness (respiratory complaint) and scheduling conflicts resulted in a changed treatment/probing schedule on three occasions. The second author, a master’s student, served as the clinician and was supervised by the first author, an American Speech-Language-Hearing Association-certified speech-language pathologist. All treatment sessions were conducted in the participant’s home and were audio and video recorded.

Treatment was applied sequentially to two sets of pictures. Before the treatment began, we determined that treatment would continue with each picture set until the participant reached at least 80 percent accuracy in naming the trained actions in three consecutive probe sessions or until 12 treatment sessions were completed.

Reliability

Dependent Variable

We selected 20 percent of all baseline and probe sessions randomly to determine interjudge reliability. An individual other than the original examiners used audiotapes and/or videotapes to rescore the selected probe sessions. That individual received training in scoring using probe sessions not selected for reliability calculation. Point-to-point agreement for determination of correct or incorrect responses within the allowed response time was verified. The average agreement was 97 percent, with a range of 86 to 100 percent across all rescored sessions.

Independent Variable

We randomly selected recordings of four treatment sessions (17% of sessions) to determine reliability of application of the following aspects of treatment: (1) elicitation of the six features in correct order, (2) request for naming response following feature elicitation, and (3) determination of correct or incorrect naming response, with appropriate clinician response option applied. Accuracy of application was 100 percent for the three preceding aspects of treatment in the reviewed treatment sessions.

RESULTS

The percentage of actions named correctly in probe sessions is depicted in the **Figure**. Graph (a) shows responses to List 1, which was the first group of items submitted to SFA. Graph (b) illustrates responses to List 2 items, which received treatment following completion of treatment with List 1 items. Graph (c) shows responses to List 3 items, which we included to examine the effects of repeated exposure without training. That is, these items were probed prior to each treatment session but were not submitted to treatment. Graph (d) displays responses to items that were limited to pre- and posttreatment probing (List 4) to provide a comparison to the exposed, but untrained list (List 3).

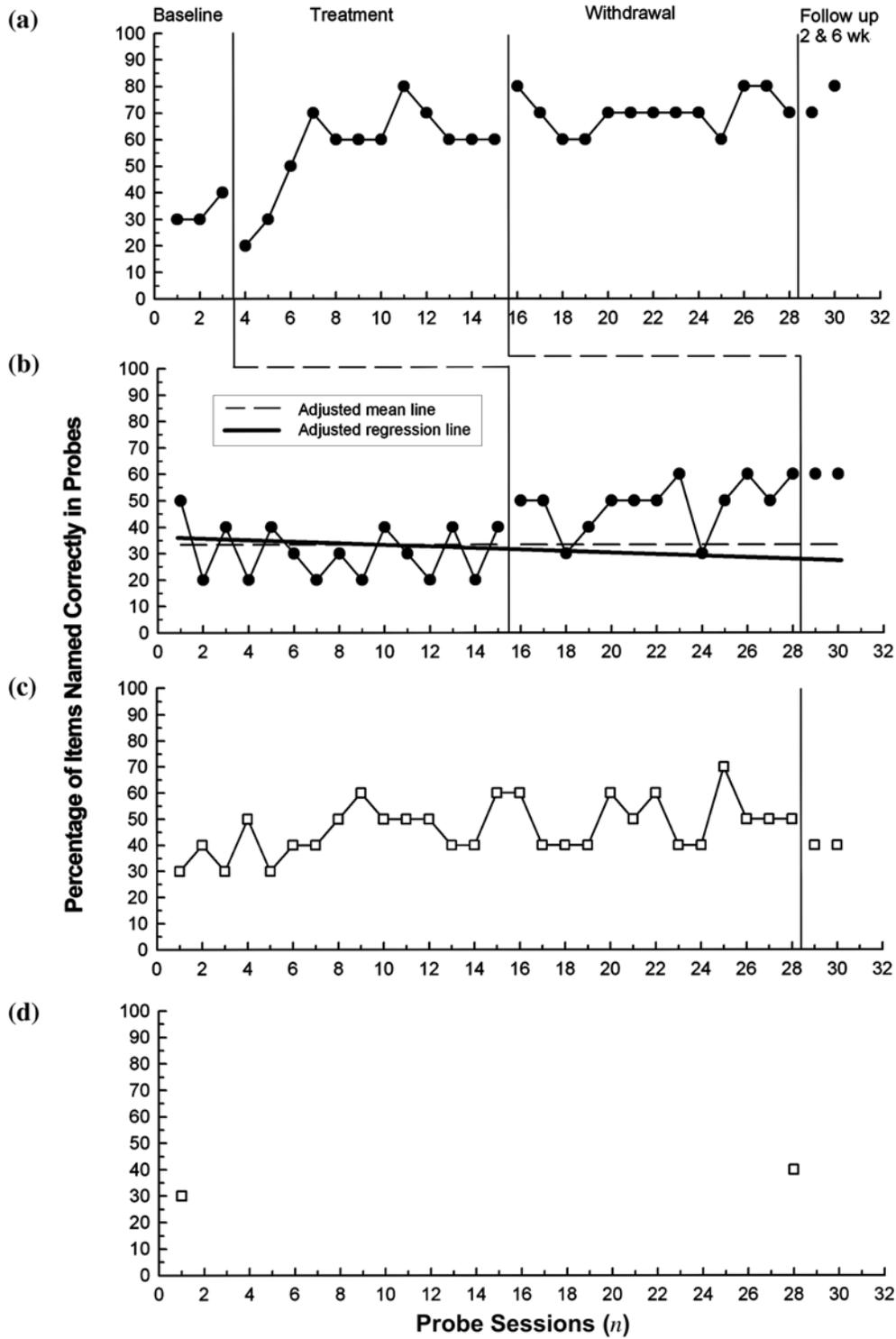


Figure. Accuracy of verb retrieval during probes: (a) treatment List 1, (b) treatment List 2, (c) exposed/untreated List 3, and (d) pre- and posttreatment List 4.

Acquisition Effects: Naming of Trained Items

Prior to treatment, responses for List 1 ranged between 30 and 40 percent accuracy, with an average of 34 percent correct (**Figure (a)**). Application of SFA resulted in improved accuracy of naming of List 1 items. Specifically, production of trained items during probe sessions reached a high of 80 percent accuracy and stabilized at 60 percent in the final three probes of the first training phase. To compare the amount of change from the baseline phase with the treatment phase, we calculated ES using the *d*-index. The *d*-index is a variation of the Δ -index and has been recommended for use when the data involve a small number of baseline observations [41]. The calculated *d*-index was 1.5 (43.3% increase) [42], indicating a moderate effect.

As seen in **Figure (b)**, variability in responding was apparent during the true and extended baseline phases for List 2 items, with accuracy averaging 31 percent across the 15 baseline points. To assist in the visual inspection of these data, we used the conservative dual-criterion (CDC) method described by Fisher, Kelley, and Lomas [43]. The CDC has been recommended for use in helping interpret single-case designs because it has been found to control type I error better than the split-middle method, the general linear model, and interrupted time series analysis when data are autocorrelated [43]. Additionally, the CDC has been found to have higher power than those methods [43]. In keeping with the CDC method, we used the 15 baseline data points (List 2) to plot adjusted mean and least squares trend lines. (The lines were adjusted upward by 0.25 standard deviation [SD].) As shown in **Figure (b)**, the lines were extended through the treatment phase. With 13 points in the treatment phase, the CDC method required that at least 10 data points in that phase be above both criterion lines to conclude that a reliable treatment effect existed. After SFA treatment was applied to List 2, accuracy of naming of List 2 items increased to a high of 60 percent correct, with an average in the final three probes of 57 percent correct. Eleven of the thirteen data points during the treatment phase fell above both criterion lines, indicating positive treatment effects for List 2 items. (The CDC method requires a minimum of five baseline points; consequently, List 1 responses could not be submitted to this analysis.)

We estimated the magnitude of the difference in baseline probe data compared with treatment phase probe data for List 2 items using the Δ -index calculation, which

yielded an ES index of 1.72 [42], representing a medium effect size (*d*-index = 1.76).

Response Generalization Effects: Naming of Untrained Items

During treatment of List 1, no changes were evident in responses to List 2. A trend of improving, but unstable, accuracy was noted for List 3 during training of List 1, which continued during training of List 2. During the true baseline phase, accuracy levels for List 3 ranged from 30 to 40 percent correct, with an average of 33 percent correct. In the final three probes of the second treatment phase, List 3 accuracy levels averaged 50 percent correct. Negligible changes were noted in responses to List 4, which was probed only pre- and posttreatment.

Maintenance: Withdrawal and Follow-Up Effects

Maintenance of treatment effects for List 1 was measured during a withdrawal phase (i.e., treatment was withdrawn from List 1 but was being applied to List 2). During the withdrawal phase, accuracy levels for List 1 remained at levels consistent with those achieved during the treatment phase (i.e., ranged from 60%–80% accuracy).

In addition, maintenance of treatment effects was measured at 2 and 6 weeks following treatment for all lists of items. Follow-up probing revealed that lasting improvements were evident for only Lists 1 and 2. Specifically, accuracy levels at 6 weeks were 80 and 60 percent, respectively for Lists 1 and 2. At follow-up, accuracy levels for Lists 3 and 4 were similar to the levels obtained during the initial baseline phase.

Effects on Production of Content During Discourse

Table 2 illustrates the possible generalization effects of SFA on production of discourse. Pre- and posttreatment discourse samples that represent the entire sample are shown in **Appendix 4**, available online only at <http://www.rehab.research.va.gov/>.

As seen in **Table 2**, wpm increased from 46 prior to treatment to 82.8 following treatment, with total talking time being similar (532 s and 558 s). Compared with the group data provided by Brookshire and Nicholas [39], the participant's wpm score prior to treatment fell below the mean of 81 but was within 1 SD of that group's average performance (SD = 39). On the basis of doubling Brookshire and Nicholas' [39] SEM for time 1 to time 3 (SEM = 5.5), we would have expected that the participant's repeated sample would have yielded a wpm score falling

between 35 to 57 with no influence of treatment or other external forces.

Number and percentage of words that were CIUs also increased following treatment. Prior to treatment, the participant's %CIU was 51, which was lower than Brookshire and Nicholas' group mean %CIU [39], but was within 1 SD (15%). Following treatment, the participant's %CIU increased to 65. Brookshire and Nicholas reported a SEM of 2.1 %CIU for time 1 compared with time 3 for the group with aphasia [39]. Consequently, a doubling of that SEM would predict that the %CIU at posttreatment would have fallen between 46 and 56 with no extraneous influences on CIU production in discourse.

Counts of verb and noun production were made for the verbs or nouns included in the total CIU count. The relative distribution of nouns and verbs remained unchanged following treatment.

Posttreatment Test Results

Results of posttreatment assessment are shown in **Table 1**. Increases in naming accuracy were observed for the TAAWF; the raw score increased from 19 to 40, with increases seen in all areas except verb-naming. Noun-naming accuracy also increased on the OANB from a raw score of 50 to 64. Performance accuracy on the VAST increased from 9/20 to 14/20 on the Sentence Construction subtest and from 8/40 to 18/40 on the Action-Naming subtest. Because a wide array of errors occurred during pretreatment administration of the Sentence Construction subtest (e.g., missing article, implausible subject/object relationship, incomplete sentences, incorrect pronoun usage, missing subjects, missing verbs), no discernible pattern of improvement was noted at posttreatment testing beyond improved completeness of sentences. Of the actions named correctly at posttreatment testing, only one

(“diving”) had been present in the experimental word lists (and this word had received treatment).

Normative data concerning test-retest reliability are unavailable for the tests just discussed; consequently, we could not determine whether changes exceeded normal variation. All other tests showed negligible changes following treatment.

DISCUSSION AND CONCLUSIONS

The results of our study revealed a modest increase in retrieval of trained action names as measured during production in probes. Although we obtained ES suggestive of positive treatment effects, our participant did not achieve levels desired for clinically significant change (i.e., a prior behavioral criterion for terminating treatment; 80 percent correct performance on probes in consecutive sessions). As with previous investigations of verb-retrieval training in aphasia, we noted improvements in naming only for trained items. However, treatment effects did apparently extend beyond trained items in that the participant's accuracy in retrieving object names during formal testing improved and verbal productivity and informativeness during discourse increased.

Acquisition Effects

Although we observed clear acquisition effects for List 1 items, positive changes were less robust for List 2 items. The use of the CDC method indicated a treatment effect for this list (and we obtained a medium ES), but improvements were less than we expected and desired. During extended baseline probing we found considerable variability for List 2, which may indicate greater difficulty

Table 2.

Production of content during discourse for 74-year-old female participant.

Measure	Pretreatment	Posttreatment
Total Words	408	770
Total Sample Time, s (min)	532 (8.9)	558 (9.3)
Words Per Minute	46	82.8
Total CIUs	208	497
CIUs Per Minute	23.4	53.4
%CIU	51	65
Verbs, <i>n</i> (% of total CIU)	68 (32)	144 (29)
Nouns, <i>n</i> (% of total CIU)	70 (34)	167 (34)
Other Grammatical Forms, <i>n</i> (% of total CIU)	70 (34)	186 (37)

CIU = correct information unit

with List 2 items for this participant. Additionally, our extended probing of these items prior to treatment, without providing feedback concerning incorrect responses, may have strengthened those incorrect responses, which conceivably could have had a detrimental effect on treatment.

We were surprised that the participant did not achieve higher levels of naming accuracy on probes following treatment. Her pretreatment testing suggested that she should have been a good candidate for treatment. That is, she demonstrated significant naming difficulty, but her deficit was not severe enough to totally prohibit accurate naming of some verbs and nouns. Similarly, semantic processing deficits were evident but did not appear to be devastating. She seemed to be motivated to receive treatment, was an attentive and active participant during the sessions, and gave the impression of enjoying the therapy.

One possible poor prognostic indicator identified in the pretreatment assessment was the participant's score on the Coloured Progressive Matrices [44]. Relatively low scores on this measure have been associated with poorer response to word-retrieval treatment [8]. However, throughout the investigation, we had no concerns regarding the participant's intellectual functioning.

The precise nature of the participant's verb-retrieval deficits was not well specified through assessment. Her semantic processing abilities may have been minimally impacted by treatment, with significant deficits remaining such that larger-scale improvements were undermined. And the case may be that in addition to semantic deficits, her naming errors stemmed from difficulties in accessing the phonological form. The treatment may not have provided sufficient opportunities for stimulation with the phonological form of the target action or for practice in accessing the phonological form.

Response Generalization Effects

Performance on List 3 items, which we used for exposure control, showed inconsistent increases and decreases during the study but returned to baseline levels of 40 percent accuracy during follow-up probes. The lack of change in accuracy of production of List 4 items suggests that repeated exposure during probing may have played a role in improved naming of List 3 items and possibly of Lists 1 and 2. Such improvement has been reported occasionally in the aphasia-naming literature [32]. However, as illustrated by the follow-up data, the changes

in naming accuracy for List 3 were not lasting, whereas the changes in Lists 1 and 2 remained. Exposure is a necessary component of treatment that likely factors into the effects of most treatments. The use of an exposed but untreated list (List 3), as well as a limited exposure list (List 4), can help estimate the effects of repeated exposure. Given that improvements were maintained only for the treated lists, exposure alone did not likely significantly affect treatment.

The results of this study partially concur with previous findings of SFA investigations [20–23], showing that treatment resulted in positive changes for trained items. However, we found no generalization to naming of untrained actions in this investigation, whereas improvements in untrained object-naming have been reported for all previous SFA participants who demonstrated improvements with trained items. (However, the generalization findings have varied across participants and have not always been robust.) As discussed earlier, one reason that response generalization may occur with SFA is that untrained items that belong to the same semantic category as trained items may benefit from stimulation of the shared semantic network. Previous studies with SFA have not controlled categorical membership closely, which may explain why the degree of generalization has varied across participants. In our investigation, various factors associated with verb retrieval were controlled, but categorical membership was not. Post hoc examination of the lists of actions indicates that some overlapping representation of semantic categories was present (e.g., “manner of movement” verbs are represented across lists). Marshall et al. did control semantic category membership and, as in our investigation, did not find significant changes in control items that belonged to the same categories as trained items [6]. The categorical organization of actions is not well understood and numerous categories have been proposed. A better understanding of categorical processing of actions/verbs may permit organization of treatment stimuli that facilitates generalization.

Also, some of the purported generalization effects observed with SFA may have actually represented repeated exposure effects. Previous studies of SFA have not controlled for such effects.

Stimulus Generalization

The finding of increases in accuracy of object-naming was not necessarily expected but is not incongruous with the treatment provided. The nature of the treatment required

retrieval of a variety of nouns associated with target actions. Consequently, a substantial portion of each treatment session was devoted to noun retrieval. Furthermore, as part of their semantic representation, verbs are considered to carry information regarding predicate argument structure (PAS) [10,27,37]. Improved retrieval of specific verbs is likely to improve access to the nouns associated with the PAS of those verbs.

The substantial increase in verbal productivity and efficiency of production of content in discourse is encouraging but somewhat difficult to explain. As stated earlier, Webster et al. also reported limited improvement in general verb-retrieval abilities accompanied by increased production of nouns in sentences and increased variety of argument structures in connected speech [9]. A component of the treatment applied by Webster et al. was similar to SFA in that the participant was asked to generate words to describe features (i.e., where, what, who, and with) of targeted verbs [9]. Interestingly, Webster et al. indicated that therapy had probably not strengthened access to specific PAS information but had likely provided the participant with a "general strategy that enabled him to specify arguments around verbs he could produce" [9, p. 761]. Such may have been the case in this investigation. The lack of generalization to untrained verbs and limited improvement in trained verbs would indicate that treatment did not provide sufficient strengthening of underlying semantic networks. Furthermore, the participant either was unable to or did not apply the general retrieval strategy of reviewing features to these verbs that were difficult to retrieve. However, the treatment strategy may have enhanced retrieval of PAS information in more easily retrieved verbs, which could account for the improvements seen in sentence production and in discourse.

The postulated explanations for the increases in noun retrieval may also apply to the increases in discourse. In addition, for some speakers, a discourse task, which is less constrained than a confrontation naming task and divergent in nature may allow the realization of improved semantic processing. Conversely, the convergent nature of naming tasks may mask generalized improvement in lexical processing. With our treatment, improved descriptive abilities might be expected, even if access to specific, untrained lexical items remains problematic. Descriptions were not quantified in this investigation, but the pre- and posttreatment discourse samples suggest increased use of descriptions (**Appendix 4**, available online only at <http://www.rehab.research.va.gov/>).

Perhaps a simpler explanation for the changes in discourse production may be that treatment had a general effect of facilitating willingness to attempt verbal communication. Our participant's daughter, in a follow-up conversation with the first author, reported that her perception of the effects of treatment was that her mother was more willing to join in conversations (this comment was unsolicited). To control for this type of potentially confounding effect, we may have found an extended baseline/pretesting phase that included repeated administrations of the discourse task useful.

Future Directions

Future research efforts with SFA applied to verbs may be strengthened in several ways. A few design modifications would enhance the interpretation of obtained data: (1) extended baselines for discourse purposes, as well as for clearer measurement of the primary dependent variable; (2) implementation of a multiple-probe procedure to reduce exposure to items; and (3) repeated pre-treatment measurement of the limited exposure list to establish stability of responding and ensure that it is equated in difficulty to the treatment lists (i.e., repeated baseline probing followed only by a posttreatment probe).

One of the probable benefits of the SFA procedure is the repeated, systematic use of the semantic feature review [21]. The treatment may be enhanced by overtly training participants to use the feature review as a retrieval strategy. That is, SFA currently relies on participants to intuit that they should apply feature review as a compensatory strategy. Rather than hoping that participants apply the strategy, therapists may find explicitly training its use useful. As in the case of object name retrieval [25], treatment effects may be improved by selecting or arranging treatment stimuli according to additional factors related to verb processing.

The findings of this investigation are certainly preliminary and require replication. However, SFA shows promise for application to the training of action names in aphasia.

ACKNOWLEDGMENTS

Thanks are extended to Kelly Morgan for her assistance with this project.

This material was based on work supported in part by the Department of Veterans Affairs, Rehabilitation Research and Development Service, grant C3826R.

The authors have declared that no competing interests exist.

REFERENCES

1. Goodglass H, Wingfield A. *Anomia: Neuroanatomical and cognitive correlates*. San Diego (CA): Academic Press; 1997.
2. Berndt RS, Mitchum CC, Haendiges AN, Sandson J. Verb retrieval in aphasia. 1. Characterizing single word impairments. *Brain Lang*. 1997;56(1):68–106. [PMID: 8994699]
3. Laiacona M, Caramazza A. The noun/verb dissociation in language production: Varieties of causes. *Cogn Neuropsychol*. 2004;21(2/4):103–24.
4. Zingeser LB, Berndt RS. Retrieval of nouns and verbs in agrammatism and anomia. *Brain Lang*. 1990;39(1):14–32. [PMID: 2207618]
5. Fink RB, Martin N, Schwartz MF, Saffron EM, Myers JL. Facilitation of verb retrieval skills in aphasia: A comparison of two approaches. *Clin Aphasiol*. 1991;21:263–75.
6. Marshall J, Pring T, Chiat S. Verb retrieval and sentence production in aphasia. *Brain Lang*. 1998;63(2):159–83. [PMID: 9654430]
7. Raymer AM, Ellsworth TA. Response to contrasting verb retrieval treatments: A case study. *Aphasiology*. 2002;16(10/11):1031–46.
8. Wambaugh JL, Cameron R, Kalinyak-Fliszar M, Nessler C, Wright S. Retrieval of action names in aphasia: Effects of two cueing treatments. *Aphasiology*. 2004;18(11):979–1004.
9. Webster J, Morris J, Franklin S. Effects of therapy targeted at verb retrieval and the realisation of the predicate argument structure: A case study. *Aphasiology*. 2005;19(8):748–64.
10. Druks J. Verbs and nouns—A review of the literature. *J Neurolinguistics*. 2002;15(3):289–315.
11. McNeil MR, Doyle PJ, Spencer KA, Goda AJ, Flores D, Small SL. A double-blind, placebo-controlled study of pharmacological and behavioural treatment of lexical semantic deficits in aphasia. *Aphasiology*. 1997;11(4/5):385–400.
12. Wambaugh JL, Linebaugh CW, Doyle PJ, Martinez AL, Kalinyak-Fliszar M, Spencer KA. Effects of two cueing treatments on lexical retrieval in aphasic speakers with different levels of deficit. *Aphasiology*. 2001;15(10/11):933–50.
13. Wambaugh JL. A comparison of the relative effects of phonologic and semantic cueing treatments. *Aphasiology*. 2003;17(5):433–42.
14. Wambaugh JL, Doyle PJ, Martinez AL, Kalinyak-Fliszar M. Effects of two lexical retrieval cueing treatments on action naming in aphasia. *J Rehabil Res Dev*. 2002;39(4):455–66.
15. Fink RB, Schwartz MF, Myers JL. Effects of multilevel training on verb retrieval: Is more always better? *Brain Lang*. 1997;60(1):41–43.
16. Linebaugh CW, Baron CR, Corcoran KJ. Assessing treatment efficacy in acute aphasia: Paradoxes, presumptions, problems, and principles. *Aphasiology*. 1998;12(7/8):519–36.
17. Nickels L. Therapy for naming disorders: Revisiting, revising, and reviewing. *Aphasiology*. 2002;16(10/11):935–80.
18. Marshall J. Doing something about a verb impairment: Two therapy approaches. In: Byng S, Swinburn K, Pound C, editors. *The aphasia therapy file*. Hove (England): Psychology Press; 1999. p. 111–30.
19. Nicholas LE, Brookshire RH. A system for quantifying the informativeness and efficiency of the connected speech of adults with aphasia. *J Speech Hear Res*. 1993;36(2):338–50. [PMID: 8487525]
20. Boyle M, Coelho CA. Application of semantic feature analysis as a treatment for aphasic dysnomia. *Am J Speech Lang Pathol*. 1995;4(4):94–98.
21. Boyle M. Semantic feature analysis treatment for anomia in two fluent aphasia syndromes. *Am J Speech Lang Pathol*. 2004;13(3):236–49. [PMID: 15339233]
22. Coelho CA, McHugh R, Boyle M. Semantic feature analysis as a treatment for aphasic dysnomia: A replication. *Aphasiology*. 2000;14(2):133–42.
23. Lowell S, Beeson PM, Holland AL. The efficacy of a semantic cueing procedure on naming performance of adults with aphasia. *Am J Speech Lang Pathol*. 1995;4(4):109–14.
24. Spencer KA, Doyle PJ, McNeil MR, Wambaugh JL, Park G, Carroll B. Examining the facilitative effects of rhyme in a patient with output lexicon damage. *Aphasiology*. 2000;14(5/6):567–84.
25. Kiran S, Thompson CK. Effect of typicality on online category verification of animate category exemplars in aphasia. *Brain Lang*. 2003;85(3):441–50. [PMID: 12744956]
26. Edmonds LA, Kiran S. Effect of semantic naming treatment on crosslinguistic generalization in bilingual aphasia. *J Speech Lang Hear Res*. 2006;49(4):729–48. [PMID: 16908872]
27. Levin B. *English verb classes and alternations: A preliminary investigation*. Chicago (IL): University of Chicago Press; 1993.
28. Thompson CK, Lange KL, Schneider SL, Shapiro LP. Agrammatic and non-brain-damaged subjects' verb and verb argument structure production. *Aphasiology*. 1997;11(4/5):473–90.
29. Kertesz A. *Western Aphasia Battery test manual*. New York (NY): Grune & Stratton; 1982.
30. Porch BE. *Porch index of communicative ability*. 3rd ed. Palo Alto (CA): Consulting Psychologists Press; 1981.
31. German DJ. *Test of Adolescent/Adult Word Finding*. Allen (TX): PRO-ED; 1990.

32. Duffy JR. Motor speech disorders: Substrates, differential diagnosis, and management. St. Louis (MO): Mosby, Inc; 1995.
 33. Yorkston KM, Beukelman DR. Assessment of intelligibility of dysarthric speech. Tigard (OR): C.C. Publications; 1981.
 34. Kay J, Lesser R, Coltheart M. Psycholinguistic assessment of language processes in aphasia. Hove (England): Lawrence Erlbaum Associates, Ltd; 1992.
 35. Bastiaanse R, Edwards S, Rispens J. Verb and sentence test. Windsor (England): Thames Valley Test Company; 2002.
 36. Druks J, Masterson J. An Object and Action Naming Battery. Hove (England): Psychology Press; 2000.
 37. Levelt WJM. Speaking: From intention to articulation. Cambridge (MA): MIT Press; 1989.
 38. Kemmerer D, Tranel D. Verb retrieval in brain-damaged subjects: 1. Analysis of stimulus, lexical, and conceptual features. *Brain Lang.* 2000;73(3):347–92. [\[PMID: 10860561\]](#)
 39. Brookshire RH, Nicholas LE. Speech sample size and test-retest stability of connected speech measures for adults with aphasia. *J Speech Hear Res.* 1994;37(2):399–407. [\[PMID: 8028321\]](#)
 40. Nickels L. Improving word finding: Practice makes (closer to) perfect? *Aphasiology.* 2002;16(10/11):1047–60.
 41. Bloom M, Fischer J, Orme JG, editors. Evaluating practice: Guidelines for the accountable professional. 4th ed. Boston (MA): Allyn & Bacon; 2003.
 42. Auerbach C, Schnall D, Heft LaPorte H, SingWin. In: Bloom M, Fischer J, Orme JG, editors. Evaluating practice: Guidelines for the accountable professional. 2nd ed. Boston (MA): Allyn & Bacon; 1995.
 43. Fisher WW, Kelley ME, Lomas JE. Visual aids and structured criteria for improving visual inspection and interpretation of single-case designs. *J Appl Behav Anal.* 2003; 36(3):387–406. [\[PMID: 14596583\]](#)
 44. Court JH, Raven J. Coloured progressive matrices. Los Angeles (CA): Western Psychological Services; 1995.
- Submitted for publication May 3, 2006. Accepted in revised form November 29, 2006.