Published in final edited form as:

Aphasiology. 2011 November 18; 25(11): 1327–1346. doi:10.1080/02687038.2011.599364.

A comparison of drill- and communication-based treatment for aphasia

Daniel Kempler and

Communication Sciences and Disorders, Emerson College, 120 Boylston St, Boston, MA 02116

Mira Goral

Speech-Language-Hearing Sciences, Lehman College, The City University Of New York, 250 Bedford Park Blvd., Bronx, NY 10468

INTRODUCTION

Some methods of treating clients with impaired speech and language production associated with aphasia emphasize predominantly drill-based exercises; others emphasize more social, discourse-based client-clinician interactions (Brookshire 2003; Chapey, 2008; Helm-Estabrooks & Holland, 1998; Hengst, Duff & Dettmer, 2010; Howard and Hatfield, 1987; Kempler, 2005; Whitworth, Webster & Howard, 2005).

The “didactic” and “stimulation” methods (Howard & Hatfield, 1987) rely on drill repetition (imitation), phrase and sentence-completion, picture naming and, in general, eliciting multiple productions of the target items during each session. These client-clinician interactions do not typically involve exchange of novel information as generally occurs in communication-based verbal interactions in discourse. Hengst, Duff and Dettmer (2010) wrote that the use of drill and repetition achieve learning through reinforcement of behaviors and/or “strengthening neuro-pathways that underlie target skills” (p. 888).

In contrast, language can be elicited in treatment in a communicative context, methods sometimes referred to as a “cognitive-contextual,” “functional,” “socio-linguistic,” “communicative” or “pragmatic” (Chapey, 2008; Kempler, 2005). These reflect the fact that language is largely “dialogic” (Hengst et al., 2010) and “transactive” (Wertz, 1998). That is, language is typically used to communicate in a variety of contexts to transfer information between individuals. The verbal interactions are generative or novel insofar as they use words and sentences (as well as gestures, etc.) creatively to communicate information. These aphasia therapy methods generally elicit language production that is relevant, unambiguous and sufficiently informative for that context. The underlying features of these methods are largely adapted from the pragmatically based cooperative principle that guides conversation (Grice, 1975). In order to preserve the pragmatics of conversation, these methods may incorporate exchange of novel information between the clients and clinicians in individual dialogues, open-ended group discussions as well as structured barrier-tasks.

It should be noted that drill- and communication-based methods are not mutually exclusive and both can be integrated into a single treatment protocol. For instance, script training (e.g., Cherney, Halper, Holland & Cole, 2008) uses a small number of conversation-based scripts, mastered through repetition. Melodic intonation therapy (Albert, Sparks & Helm, 1973; Belin et al., 1996) uses a small set of (drill-learned) intonation patterns superimposed on

conversational exchanges. The relative preference of any one method might depend on severity of aphasia (more severe symptoms might warrant eliciting speech production through simple sound repetition while milder aphasia might allow for a concentration on discourse structure through novel narrative generation), the goals of the client, the perspective of the clinician and whether the therapy is provided in an individual or group setting.

Clinical efficacy can be evaluated on the basis of a direct treatment effect such as improved performance on the treated items in the treatment context. However, one of the crucial questions about aphasia therapy is the extent of generalization achieved by the treatment. Generalization refers to demonstration of a behavioral response in a different linguistic, interpersonal or other context not explicitly practiced in the treatment. Ideally, treatment will result in improvement of untrained lexical items, linguistic structures, communication tasks or settings and interaction partners.

It is our hypothesis that the type of interaction used in therapy will determine the extent and type of improvement and generalization found following therapy. We predict that drill-based exercises, because they typically work with a restricted set of items in a non-communicative context, are more likely to result in direct treatment effects, such as improvement on the particular items and tasks practiced in therapy. In contrast, we hypothesize that communicative interactions using a less restricted set of practice items are more likely to result in generalization to a wider array of items and contexts. Greater generalization in this case might be attributed to the similarity between the therapeutic interactions and other discourse contexts insofar as both incorporate the core discourse principle of information exchange.

While most aphasia treatment research does not explicitly compare drill-based and communication-based treatment methods, recent studies of Constraint Induced Aphasia Therapy (CIAT) have addressed this to some extent. CIAT is built on three principles borrowed from work in Constraint-Induced Movement Therapy with hemiparetic patients (Taub, Uswatte & Pidikiti, 1999). Based on the notion that brain injury can result in adaptations that substitute easier but non-optimal compensations for lost function (called “learned nonuse”), constraint induced therapy restrains the use of compensatory adaptations (e.g., use of left-hand for patients with right hemiplegia), is intense by emphasizing “massed practice” (i.e., movement is often practiced all day long), and is functional (i.e., clients engage in real-life activities). These principles have been adapted to aphasia therapy by constraining interaction to verbal channels, implementing relatively intense therapy schedules (massed practice), and using communicative card games such as “Go Fish” to simulate functional communication. Results available to date for the efficacy of CIAT are promising, with participants showing improvement in standardized test measures, direct treatment effects on baseline measures, and generalization to other contexts by measuring independent ratings of participant communication (Goral & Kempler, 2009; Maher et al., 2006; Pulvermüller et al., 2001). It is not clear from this body of research which aspect(s) of CIAT are most crucial to improved performance. Several studies have demonstrated the importance of constraining therapy to verbal output only, in finding that this type of therapy was more effective than a comparison therapy in which participants could use other modalities as well (Maher et al., 2006; Pulvermüller et al., 2001). However, Barthel and colleagues (Barthel, Meinzer, Djundja & Rockstroh, 2008) compared two intensive therapies that were communication based, one in which language was limited to verbal output and another in which additional communication modalities were used (e.g., writing). They found that on most measures the two treatments were equally effective. With regard to intensity, there is significant evidence that more intense therapy yields more robust short-term

benefits, although the long-term advantage of intense therapy is less clear (Cherney, Patterson, Raymer, Frymark & Schooling, 2008).

The specific contribution of functionality in CIAT has not been directly addressed. The goal of the present study was to determine the relative importance of functionality in the effectiveness of aphasia therapy. Towards that end, we conducted two single-subject studies in which therapy was provided to individuals with moderately severe non-fluent aphasia. Both participants received drill-based and communication-based treatments in sequence to ascertain the relative efficacy of each treatment type on measures of word, sentence and narrative production.

METHODS

Participants

Two women diagnosed with non-fluent aphasia participated in the study. Each had sustained a single left hemisphere thrombo-embolic stroke. P1 was 54 years old and two years post-onset and P2 was 45 years old and seven years post onset at the time of treatment. Both participants are right-handed and both demonstrated mild-moderate right-hemiparesis. Both participants completed high school and worked prior to their strokes. Following the stroke, both received speech-language therapy; neither was engaged in individual therapy other than the experimental treatment reported here for the duration of the study.

The language production of both participants' consisted primarily of one-word utterances, few function words, and slow speech production with many pauses. Neither exhibited dysarthria or marked apraxia of speech. Their comprehension skills surpassed their production skills and appeared good in conversational context. At the onset of the study, P1's Western Aphasia Battery (WAB, Kertesz, 1982) aphasia quotient was 75.9 and P2's was 65.2 (see Table 1). Neither participant exhibited hearing or visual impairment. The participants volunteered for the study and signed appropriate consent forms.

Treatment protocol

Each participant completed 60 hours of individual treatment. Two different types of treatment were administered sequentially, each for 30 hours, approximately 7.5 hours per week. Due to their difficulty producing verbs and grammatical sentences, the goal of both treatment protocols was identical: production of appropriate verbs in single sentences. The two protocols differed in the way in which the production was elicited and in what the acceptable targets of production were. Specifically, they differed in two crucial respects: whether the targets were drawn from a finite list of verbs and whether the interactions were communication-based. For the sake of explication, the two protocols are labeled "Drill" and "Generative."

The Drill protocol targeted a set of 32 verbs, selected from the action naming protocol presented in prior work on verb production (Almor et al., 2009) and used in the present study as one of the outcome measures (see outcome measures below for further description of the verb set). An additional set of 32 comparable filler verbs were included in treatment to yield a larger practice set of 64 verbs, but only the 32 targeted items were tested before and after treatment. The Drill Protocol provided the participant with highly structured practice of the treated list of verbs. During each session a subset of the target verb list was used in six activities: sentence repetition, picture description, a picture map task, choral reading, a memory game, and sentence completion. During a given activity, multiple sets of 4–10 spoken or written sentences or picture cards each depicting an action, were presented to the participant. Prior to each activity, the participant rehearsed the items in a three-step process by listening to the clinician's model, repeating the production in unison with the clinician,

and then repeating after the clinician's model. Following this rehearsal, the words were practiced in the activity. Production was not elicited in a communicative context; rather, in each case, the clinician knew the target and worked through scaffolding to elicit the correct verb and structure. On average each verb was practiced approximately 40 times in the course of 30 hours of Drill treatment. Novel pictures depicting the target action verbs were introduced frequently. Therefore there were more than 64 different pictures used to depict the 64 actions/verbs.

In contrast, the Generative treatment did not involve a limited set of preselected verbs. Here, the participant was encouraged to generate single sentences containing verbs in grammatical structures to communicate the description of the picture they selected. The Generative protocol relied on verbal games and informative communication-based interactions. Therapy activities included card games, such as Go Fish, Memory, and a Map task, in which the participant and the clinician took turns generating sentences that described actions depicted in picture cards. In these language games, all verbal interactions adhered to conversational maxims in that the clinician did not know which card the participant was attempting to describe in a given turn. This was accomplished by use of a physical barrier and duplicate materials. Crucially, in the Generative treatment no clear targets were pre-determined; any verb in a sentence, which communicated relevant information, was accepted. The verbs and sentence structures were not rehearsed; modeling was limited to the clinician's productions within the turn taking context of the activities. The emphasis in this therapy was for the participants to develop the habit of generating diverse verbs in sentences in functional contexts, rather than to practice a limited set of pre-selected verbs. To reinforce the habit of producing verbs, the participant and clinician produced two different but appropriate responses per turn.

Shaping and scaffolding were used throughout both treatments. If the participant did not produce an appropriate verb in a grammatical sentence, the clinician requested the missing information (in the case of omitted verb or other element) and modeled an accurate response as needed. Both participants quickly internalized the rules of the games and attempted to produce verbs in sentences in each of their productions. When participants achieved 80% accuracy during in-session probing, the target structure was modified. For example, if a participant was able to produce simple S-V-O structures, the elicited structure was advanced to S-V-O+PP. None of action pictures from the outcome measure were used as stimuli during treatment. See the Appendix for more details of treatment tasks and parameters.

Treatment was provided by student clinicians under direct supervision of the authors. Treatment fidelity was assured by regular discussions to plan and review the protocol and by observation of the sessions.

Pre- and post-testing (outcome measures)

Two measures were administered prior to and following a 3-week period of no treatment (to establish a stable baseline) and before and after each treatment block. One measure assessed verb naming and the other examined narrative production. At each testing time, data were collected over three sessions, with a third of the outcome data collected in each of the three sessions. Testing over three sessions allowed us to calculate effect sizes to assess change. A standardized measure, the WAB, was administered once before and after each treatment block.

Verb naming

A set of 96 action pictures was used to elicit verb naming in sentences (Almor et al., 2009). All of the verbs were highly imageable, familiar, and with high (87%) overall naming

agreement. The verbs were divided into three comparable lists; one list was treated during the Drill protocol. The outcome measure for the verb-naming task was the number of target verbs accurately produced. We counted as correct only those verbs that were expected as targets in the Almor study (e.g., for a picture of a girl “balancing” on a beam, the verb “practicing”, although describing the picture, was not counted as correct). Verbs that are homophones of nouns (e.g., to iron-an iron) were only counted as verbs if there was evidence in the sentence to support this (e.g., in a picture of a woman ironing a shirt, a single word production “iron” was not counted as a verb; the production of “iron” in: “the woman iron the shirt” was). For Drill treatment, we examined naming accuracy for treated and untreated items.

Narrative production

Three personal narratives (a vacation; a happy moment for you or someone in your family; and a movie, book or TV show that you enjoyed) were elicited prior to and following each treatment block. Narratives were elicited with a minimum of prompting. The participants were prompted to tell a different narrative each time so they did not tell the story of the same vacation, happy moment or movie on each testing. The recorded narratives were transcribed, checked by another investigator, and divided into utterances. To identify each utterance, we modified Foster, Tonkyn & Wigglesworth’s (2000) notion of an Analysis of Speech Unit (AS-Unit), defined as any independent clause (that minimally includes a finite verb) as well as any sub-clausal unit, “which can be elaborated to a full clause by means of recovery of ellipted elements from the context of the discourse....” Identifying each individual AS-Unit took into consideration intonation contour, pauses, meaning, and syntactic construction. We modified the original definition to apply to utterances as minimal as a single word, even in instances when the missing elements were difficult to infer from the context.

The outcome measures for the narratives included lexical, sentence, and discourse variables as follows. Because treatment focused on verb production, our lexical measures were Total verbs (the total number of verbs produced), Verb diversity (the number of different verb types) and Verb form (the number of correctly conjugated verb forms). The sentence level variables included the Lexical density of utterances (proportion of content words of total words) as well as whether or not the utterances were complete and grammatical. Utterance completeness was rated using the AS-Unit scale described above, taking into account both the complexity of the structure and presence of obligatory elements. Note that utterances could be complete if they contained all the obligatory elements even if the end result was not grammatical due to poor grammatical form (e.g., “The birds is eating the kitten”). To determine sentence completeness, AS-Units were identified and then rated on a 5-point scale: 1 = incomplete sentence; 2 = simple and complete, containing all obligatory elements; 3 = complex sentence containing an incomplete subordinate clause; 4 = complete coordinate structure; 5 = complete containing a subordinate structure. Grammatical utterances contained all obligatory elements and adhered to grammatical rules of English.

Discourse variables included two measures of productivity: Total number of words (not including exact repetitions, false starts, interjections or formulaic expressions) and Total number of utterances; and three measures of narrative structure: Local coherence, Global coherence, and Story line ratio. Local coherence was gauged by whether each utterance was directly related to the prior utterance by elaboration, sequencing, focus, etc. Global coherence refers to the relevance of each utterance to the general topic. For the coherence measures we used a 3-point scale (1= unrelated, 2= possibly related, and 3=clearly related), modified from Coelho and Flewellyn (2003). The final discourse measure, Story line ratio, was the number of utterances that described a clear story line divided by the number of utterances that contained a description of the setting or background for the narrative (Longacre, as reported by Olness, 2006). The WAB subtests related to spontaneous speech,

auditory-verbal comprehension, repetition, and naming and word finding were administered before and after each treatment block and an aphasia quotient was calculated.

ANALYSIS AND RESULTS

All testing sessions were recorded and transcribed for later analysis. The transcripts were coded for variables of interest by individuals blind to treatment conditions. Approximately 20% of each transcript was re-coded by a second individual, with inter-rater reliability ranging between 93% and 99% for each transcript; disagreements were resolved by discussion with the authors. In the results that follow, we evaluated change using effect sizes based on means and standard deviations of results from the three testing sessions prior to and following each treatment. Effect sizes were calculated by subtracting the pre-treatment mean from the post-treatment mean, divided by the pre-treatment standard deviation. Although there has been debate about how to determine whether an effect size represents a meaningful change (e.g., Beeson & Robey, 2006; Robey, 1998), we follow Raymer and colleagues (e.g., Raymer et al., 2007; Rodriguez, Raymer & Gonzalez Rothi, 2006) in interpreting as clinically significant improvement a positive effect size greater than 2.5 with a change of greater than 20% from baseline. Three of the narrative variables reported below (AS units, Local and Global coherence) were measured on a limited nominal scale (1, 2 or 3 rated for individual utterances). For these, pre- to post-treatment change was assessed using non-parametric Mann-Whitney comparisons based on all items across the three testing sessions before or following each treatment rather than means of each testing session, due to the lack of variability inherent in the measure and brevity of some of the narratives. For these variables, we used an alpha level of .05. We report below the results for the standardized testing and our two experimental tasks: Verb naming and Narrative production.

Standardized testing

Performance on the WAB showed minimal increase following each treatment type (see Table 1). The extent of change did not exceed five points and therefore is not considered clinically significant (Katz & Wertz, 1997).

Verb naming

The results of the verb naming are presented in Table 2. Overall, mean naming accuracy for the 96 items did not change significantly for either participant following either treatment (P1: pre-treatment $M=18.3$ $SD=2.5$; post-treatment $M=22.0$; $SD=2.6$; effect size = 1.5; P2: pre-treatment $M=21.0$; $SD=7.9$; post-treatment $M=22.3$; $SD=7.4$; effect size = .17). When we examined the change in the 32 treated items before vs. after Drill treatment, we observed a numerical increase in number of correctly produced verbs for both participants (from 18 to 21 for P1 and from 18 to 28 for P2). This increase can be taken as direct treatment effect: following Drill treatment, P1 named all the verbs she had named prior to the treatment plus three additional verbs she had not named accurately before; P2 named all the verbs she had named before Drill treatment except one and named 11 additional verbs that she had not named before the treatment. Those 11 verbs were practiced on average during treatment as often as other verbs among the 32 trained items. On the 64 untreated items, the two participants had different patterns: P1 increased naming accuracy (from 37 to 45) while P2 decreased in naming accuracy (from 45 to 39).

Following Generative treatment there was no improvement in total accuracy (P1 decreased in total accuracy from 65 to 57; P2 named 68 correct verbs pre- and 69 post Generative treatment). In the Generative treatment, no verbs were targeted for treatment so a comparison of treated vs. untreated items is not applicable. Nor was there change when the test was administered following a period of no treatment (P1: Pre $M=18.0$ $SD=0$; Post M

=18.0; SD=2.6; effect size not calculable due to pre-treatment SD of 0; P2: Pre M=12.7 SD=2.9; Post M =14.0; SD=3.0; effect size = .46).

Narrative production

We report below the results for variables at the lexical, sentence, and discourse levels derived from the narrative production. Overall, the participants produced better narratives after Generative but not Drill treatment. The narrative results are presented in Table 3.

Lexical variables

The narratives of both participants contained a greater number of verbs, greater range of verbs (verb diversity), and more accurate verb morphology following Generative but not Drill treatment. Specifically, the lexical measures showed positive albeit modest effect for P1 (increase of 20% from baseline on all three measures but effect sizes smaller than 2.5) and medium-to-large effect sizes (between 2.89 and 7.5 with increases of at least 20%) on all three measures for P2.

Sentence variables

Overall, sentence quality improved following Generative but not Drill treatment. Both participants produced a significantly greater number of complete utterances (see AS units in Table 3) and more grammatical utterances following Generative treatment. P1 produced a greater number of grammatical utterances following Generative treatment (effect size = 5.2) despite the fact that she did not produce more utterances overall following either treatment. P2 also produced more grammatical utterances following Generative treatment, but an effect size could not be calculated because she produced no grammatical utterances prior to Generative treatment (a pre-treatment SD of 0 cannot be used to calculate an effect size). The only sentence-level measure that showed no positive change for either participant following either treatment was Lexical density (i.e., proportion of content words compared to all words) within utterances.

Discourse variables

The discourse measures overall showed improvement following Generative but not Drill treatment. On the measures of overall narrative productivity – the number of words and utterances each participant produced before and after treatment – P1 did not change following either treatment; P2 produced substantially longer narratives following Generative (but not Drill) treatment. All three measures of internal narrative structure changed towards typical values following Generative but not Drill treatment. Specifically, both participants produced more utterances that clearly related to the overall narrative topic (Global coherence) and more utterances related to immediately preceding utterances (Local coherence) after Generative treatment. Moreover, the post-Generative narratives contained more utterances relating punctiliar events (story-line elements) that were judged to move the narrative forward than the pre-Generative narratives. The same was not seen comparing performance prior to and following Drill treatment.

SUMMARY AND DISCUSSION

Two women with chronic non-fluent aphasia participated in language treatment designed to improve verb production in sentences. Each participant received two treatment protocols, provided in sequence. The two protocols contrasted drill-based vs. communication-based interactions between the participant and the clinician. Outcome measures, administered before and after each treatment block, examined the production of verbs in an action picture description task and production of personal narratives. Drill treatment, which relied on repetition, had a positive albeit small effect on participants' verb naming accuracy. The

Generative treatment had a pronounced positive effect on participants' use of verbs in self-generated narratives as well as on their sentence and narrative structure.

Aphasiologists acknowledge that many treatment approaches in aphasia can bring about change in measured language behaviors that are directly related to those addressed in treatment. However, a more challenging task has been to demonstrate consistent generalization from language produced during treatment to language use outside of treatment (e.g., Edmonds, Nadeau, & Kiran, 2009; Links et al., 2010; Peach & Reuter, 2010; Rose & Sussmilch, 2008).

In the present study, improved verb naming for items that were practiced during treatment would be evidence of a direct treatment effect, whereas increased naming accuracy for verbs that were not practiced in therapy would be considered generalization to untreated items; improvement on narrative production would be taken as generalization to an untrained task. One of our participants (P2) demonstrated a direct treatment effect following Drill therapy, as her naming accuracy increased on treated items. She showed little evidence of generalization to untreated items. This is similar to other studies that have reported a lack of generalization to unpracticed verbs following various verb training protocols (e.g., Raymer et al., 2007; Webster, Morris, & Franklin, 2005). The other participant (P1) showed evidence of improvement on verb naming overall as she named more treated and untreated items following Drill treatment. We conclude that overall, the verb naming measure demonstrated generally positive outcomes. That is, practicing the production of specific verbs during treatment led to increased ability to produce these verbs and, to some degree, to increased ability to produce other verbs in a task similar to one used during treatment. As may be expected, neither participant named more target verbs in the action description task following the Generative treatment, as the participants did not directly practice any of the verbs used in the outcome measure during this treatment. However, following Generative treatment, both participants improved their verb production in narratives.

One of our primary questions was whether treatment improved language production that is more typical of language use outside of therapy. To gauge general language use, we analyzed personal narratives the participants produced before and after each treatment. This measure differed from the treatment tasks in that telling personal narratives required participants to generate multiple linked utterances without visual support of pictures. Personal narratives are an appropriate assessment of functional language use as they occur in many discourse contexts from reporting daily events, telling jokes and stories, and can contribute to an individual's ability to understand life experiences, present oneself to others, and participate in various communal and communicative events. Indeed, it might be argued that one ultimate goal of aphasia treatment is functional production and comprehension of narratives.

The results of our narrative analyses were consistent and striking: The only marked changes in these measures for either participant were seen after Generative but not Drill treatment. While there were some differences in the strength of findings for the two participants, they both showed improvements on measures of lexical diversity, sentence structure and discourse in their narratives following Generative but not the Drill treatment. These results are discussed below in the context of other similar findings in the literature.

First, our findings of greater verb diversity and more accurate verb morphology in narratives following Generative treatment are consistent with a recent study reporting treatment generalization to elicited connected speech (Links et al. 2010). Links and colleagues trained verb and sentence production using a program they developed and termed ACTION, in which verbs are elicited at four levels (single verbs, filling in infinitives, filling in finite

verbs, and sentence construction). They examined a number of variables to assess change in discourse production following treatment and reported positive change for some of their participants on two lexical measures comparable to ours, namely, those pertaining to verb diversity and verb morphology. They also found, similar to the improvement we documented for AS Units for our participants, that nine of their 11 participants produced longer sentences following treatment.

Second, the improvement in sentence structure we documented (more complete and more grammatically correct utterances) reflects a better ability to retrieve the required elements of the sentence, and is similar to the generalization findings reported in two recent studies. Edmonds et al. (2009) trained the production of verbs in conjunction with thematic roles and then used a picture-description task to elicit connected speech to examine treatment generalization. They found that, following treatment, their participants produced both the subject and the verb in the sentence more often than prior to treatment and that three of their four participants produced more complete utterances following treatment. Similarly, Webster et al. (2005) treated one person's verb production deficit using activities that required verb production, noun-verb association, and sentence generation and demonstrated that their participant produced a greater variety of argument structures and omitted fewer obligatory items in connected speech following treatment.

Finally, to our knowledge, aspects of discourse structure have not been examined in prior studies of verb therapy in aphasia. Our results of impaired narrative coherence support Coehlo and Flewellyn's (2003) finding that macrolinguistic abilities are impaired in aphasia, and suggest that therapy that incorporates basic discourse principles can be helpful in improving these elements of narrative production.

Why was Generative treatment more effective than Drill treatment for improving narrative production? In order to address this question, we discuss here how the two treatments were similar and how they differed. Both treatments were similar insofar as they adhered to two principles familiar from descriptions of Constraint Induced Aphasia Therapy (CIAT): massed practice and constrained language use (Goral & Kempler, 2009; Maher et al., 2006; Pulvermüller et al., 2001). Massed practice was implemented by administering both treatments with a relatively intense schedule of 7.5 hours a week for four weeks. Constraint was implemented in both therapies by emphasizing the use of verbal modality and the expected use of verbs in every sentence. The therapeutic use of only spoken language and frequent verb production were motivated by the theory of learned nonuse in which a less impaired skill becomes a new standard as an individual with a brain injury learns to avoid a relatively impaired function and rely on a relatively spared function (Lillie & Mateer, 2006; Pulvermüller et al., 2001; Taub, Uswatte, Mark & Morris, 2006). From this perspective, persons with aphasia come to rely on non-verbal communication modalities, and, if they produce verbal language, they rely on those language components that are most easily accessible to them. For individuals with non-fluent, agrammatic aphasia, this means they often rely on the production of nouns in verbal communication rather than struggle to produce clear and complete sentences containing verbs. The sole use of the verbal modality was exercised here to avoid compensatory use of less-impaired communication modalities, such as gesture and drawing; the requirement for production of verbs was motivated to reduce an over-reliance on nouns. Since both treatments required participants to produce a spoken verb on every turn despite their difficulty in doing so, the simple act of producing verbs cannot explain the different outcomes of the two treatment protocols.

The two treatments diverged with regard to a third principle of CIAT – the one we set out to examine in the present study – functionality. Specifically, the tasks used in the two treatments differ in whether or not they were designed to communicate new information.

The Drill treatment relies on clinician modeling of target responses and the participant repeating, reading aloud, describing action pictures and participating in choral reading. Therefore the Drill therapy does not involve exchange of new information. In contrast, in the Generative treatment, there are no target responses modeled by the clinician and the language productions are relatively unscripted. On each turn, there are a variety of possible and acceptable responses. Each turn is designed to transmit relatively novel information, capitalizing on preserved pragmatics of cooperative discourse such as turn taking and maintaining contextual relevance, in structured language games designed to elicit single sentences. These elements are neither relevant nor called into use during Drill therapy and are, in fact, artificially suspended for the drill-based exercises. We suggest that integrating the discourse principle of informative exchanges into Generative therapy enhanced the generalization of the practiced tasks (production of verbs in single sentences describing action pictures) to another discourse context (production of self-generated narratives consisting of linked utterances without the visual support of pictures).

The integration of novel information exchange in the Generative treatment makes it similar to therapeutic approaches that emphasize the importance of conversational context, such as Supported Conversation and the Life Participation Approach to Aphasia (e.g., Kagan et al., 2008; 2001; Simmons-Mackie, 2008). These approaches emphasize the social-functional context of aphasia and aim to bridge the disparity that often exist between gains that people with aphasia achieve in treatment sessions and their ability to carry over these improvements to other communicative and discourse contexts.

Generative therapy is similar to conversation-based treatments insofar as they both are based on the communication of novel information. However, unlike conversational treatment contexts, (1) our tasks constrained participant responses to verbal productions only (as opposed to encouraging the use a wide array of communication modalities) and (2) required participants to generate single sentences, neither connected to one another nor to a unifying thematic context. This structure allowed us to address our participants' specific deficits and focus on restoring the habit of producing spoken sentences containing verbs. We postulate that the combination of focusing on the use of verbs in spoken language production and communicative exchanges helped the participants to begin undoing the learned nonuse of verbs in sentences.

Another reason Generative treatment might be more likely to generalize is that it required participants to self-generate sentences, which is similar to most other communicative activities including our narrative outcome measures. This is not true of Drill therapy in which responses were modeled for the participant. We note that many of the verbs the participants produced in their post-Generative narratives were not produced more than a few times in the course of the 30 hours of treatment. Of the 14 verbs produced by P2 in the post-Generative narratives, 10 were produced during the therapy sessions fewer than 10 times each and only three were practiced more than 15 times; similarly, half of the verbs produced by P1 in her post-Generative narratives were used during treatment 10 times or fewer. Therefore, it was not the production of specific verbs that improved following Generative treatment but the ability to use verbs more consistently during language production.

Finally, another advantage of the Generative treatment might be the fact that it is more engaging and motivating than the Drill approach simply because the information exchange is more typical of daily language use. Integrating this discourse principle into the therapy technique allows participants to use their intact conversational and pragmatic skills. Whereas we do not deny the value of some drill-based practice in aphasia treatment, we take our results as evidence for a limited carryover afforded by repeated practice of specific items in a non-communicative context.

The results of this study should be interpreted carefully given several limitations. First, the findings should be replicated with additional participants before we can generalize these findings to a wider population. Second, the two treatments we employed were administered sequentially – both of our participants received Drill treatment prior to Generative treatment. Therefore, it is possible that the Generative treatment was more effective in improving narrative production simply because it was the second treatment and the improvement is due to a cumulative effect of hours of treatment, not the type of treatment. We do not believe this to be the case for several reasons: Our participants were both individuals with chronic aphasia, who had shown only modest gains in traditional therapy before this study, and made no progress on our outcome measures during a three-week period of no treatment prior to this study, nor during the intense course of Drill treatment. Therefore we think it is reasonable to attribute the improvements we observed to the properties of the Generative therapy protocol. Similarly, it is possible that improvement on the outcomes measures after the second treatment was due to a practice effect associated with the repeated administration of the measure. However, this appears unlikely as we elicited different specific narratives (i.e., different vacations, different movies) at each testing time, making it unlikely that the participants had merely gotten better at telling a particular story. Also of note, two other data points from our laboratory, not reported here, support these conclusions: (1) a prior published study with a comparable (i.e., chronic, non-fluent) participant demonstrated that Generative treatment elicited significant positive results (Goral & Kempler, 2009) while a drill-based treatment with the same individual following the Generative treatment did not show continued gains; (2) one of the participants in the study reported here (P2) received additional non-generative treatment following the Generative treatment reported here, and did not demonstrate additional functional language gains. Nonetheless, we recommend that future work counterbalance the order of treatments to study a potential contribution of treatment order to the effects reported here.

Another question unanswered by this research is whether the effectiveness of Generative therapy depends on the severity and/or language profile of the person with aphasia. Our participants were similar in that they were relatively young and demonstrated chronic, moderate non-fluent symptoms. Despite these similarities, their language production and responses to treatment differed to a degree. P2, who produced shorter and fewer spontaneous utterances with fewer verbs than P1 at the start of the study, showed greater improvement than P1 at the lexical level after Generative treatment. By contrast, P1 included verbs in many of her utterances prior to treatment and her main improvement was characterized by significant changes at the sentence and discourse levels. Therefore it appears that this type of protocol, which can be adjusted to target output of various levels of utterance complexity (verb alone, SVO, SVO+PP), can have a positive effect for participants with a range of verbal output levels. Both our participants were ideal for this research protocol because while their spontaneous speech was severely limited, they had the residual ability to produce target items and structures in a structured context with a minimum of scaffolding. However, it remains to be seen if the efficacy of the Generative treatment is comparable for participants with more or less severe as well as different symptoms.

Finally, we used only one outcome measure of language use (personal narratives). We acknowledge that story telling is only one aspect of typical language use and may not be representative of performance in conversation and other ways in which people use language. Future research will be strengthened by including a wider range of outcome measures, to assess other aspects of functional language, including measures of social validation through of treatment-related change.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank our participants and their families as well as the following research assistants who participated in collecting and analyzing the data: Carmit Altman, Maria Boklan, Caroline Cano, Carol Cayer-Meade, Peggy Conner, Laura Glufing-Tham, Kristen Maul, Youngmi Park, Dagmar Alvarado, Shancia Jarrett, Deepa Sar, Shannon Saul, Vanessa Smith, Tali Swann-Sternberg, Libby Venancio and Stephanie Wolfe. Support for Mira Goral was provided by NIH grant DC-009792.

References

- Albert ML, Sparks RW, Hlem NA. Melodic Intonation Therapy for Aphasia. *Archives of Neurology*. 1973; 29:130–131. [PubMed: 4717723]
- Almor A, Aronoff JM, MacDonald MC, Gonnerman LM, Kempler D, Hintiryan H, Hayes UL, Arunachalam S, Andersen ES. A common mechanism in verb and noun naming deficits in Alzheimer's patients. *Brain and Language*. 2009; 111(1):8–19. [PubMed: 19699513]
- Barthel G, Meinzer M, Djundja D, Rockstroh B. Intensive language therapy in chronic aphasia: Which aspects contribute most? *Aphasiology*. 2008; 22:408–421.
- Beeson PM, Robey R. Evaluating single-subject treatment research: Lessons learned from the aphasia literature. *Neuropsychological Review*. 2006; 16:161–169.
- Belin P, Van Eeckhout Ph, Zilbovicius M, Remy Ph, Francois C, Guillaume S, Chain F, Rancurel G, Samson Y. Recovery from nonfluent aphasia after melodic intonation therapy. *Neurology*. 1996; 47:1504–1511. [PubMed: 8960735]
- Brookshire, RH. *Introduction to Neurogenic Communication Disorders*. St. Louis: Mosby; 2003.
- Chapey, R. *Language Intervention Strategies in Aphasia and Related Neurogenic Communication Disorders*. 5. Philadelphia: Wolter Kluwers/Lippincott, Williams & Wilkins; 2008.
- Cherney LR, Halper AS, Holland AL, Cole R. Computerized script training for aphasia: Preliminary results. *American Journal of Speech-Language Pathology*. 2008; 17:19–34. [PubMed: 18230811]
- Cherney LR, Patterson JP, Raymer A, Frymark T, Schooling T. Evidence-based systematic review: Effect of intensity of treatment and constraint-induced language therapy for individuals with stroke-induced aphasia. *Journal of Speech, Language and Hearing Research*. 2008; 51:1282–1299.
- Coelho CA, Flewellyn L. Longitudinal assessment of coherence in an adult with fluent aphasia: A follow-up study. *Aphasiology*. 2003; 17(2):173–182.
- Edmonds LA, Nadeau SE, Kiran S. Effect of verb network strengthening treatment (VNeST) on lexical retrieval of content words in sentences in persons with aphasia. *Aphasiology*. 2009; 23:402–424. [PubMed: 19763227]
- Foster P, Tonkyn A, Wigglesworth G. Measuring spoken language: a unit for all reasons. *Applied Linguistics*. 2000; 21(3):354–375.
- Goral M, Kempler D. Training verb production in communicative context: Evidence from a person with chronic non-fluent aphasia. *Aphasiology*. 2009; 23(12):1383–1397. [PubMed: 19911035]
- Grice, HP. Logic and conversation. In: Cole, P.; Morgan, JL., editors. *Speech Acts, Syntax and Semantics*. Vol. 3. New York: Academic Press; 1975. p. 41-58.
- Hengst JA, Duff MD, Dettmer A. Rethinking repetition in therapy: Repeated engagement as the social ground of learning. *Aphasiology*. 2010; 24:887–901.
- Helm-Estabrooks, N.; Holland, A. *Approaches to the Treatment of Aphasia*. San Diego: Singular; 1998.
- Howard, D.; Hatfield, FM. *Aphasia Therapy: Historical and Contemporary Issues*. East Sussex: Lawrence Erlbaum Associates; 1987.
- Kagan A, Black SE, Duchan JF, Simmons-Mackie N, Square P. Training volunteers as conversation partners using 'Supported Conversation for Adults with Aphasia' (SCA): A controlled trial. *Journal of Speech, Language and Hearing Research*. 2001; 44:624–638.

- Kagan A, Simmons-Mackie N, Rowland A, Huijbregts M, Shumway E, McEwen S, Threats T, Sharp S. Counting what counts: A framework for capturing real-life outcomes of aphasia intervention. *Aphasiology*. 2008; 22 (3):258–280.
- Katz R, Wertz RT. The efficacy of computer provided reading treatment for chronic aphasic adults. *Journal of Speech, Language, and Hearing Research*. 1997; 40:493–507.
- Kempler, D. *Neurocognitive Disorders in Aging*. Thousand Oaks: Sage Publications; 2005.
- Kertesz, A. *Western Aphasia Battery*. New York: Grune & Stratton; 1982.
- Lillie R, Mateer CA. Constraint-based therapies as a proposed model for cognitive rehabilitation. *Journal of Head Trauma Rehabilitation*. 2006; 21(2):119–130. [PubMed: 16569986]
- Links P, Hurkmans J, Bastiaanse R. Training verb and sentence production in agrammatic Broca's aphasia. *Aphasiology*. 2010; 24 (11):1303–1325.
- Maher LM, Kendall D, Swearingin JA, Rodriguez A, Leon SA, Pingel K, et al. A pilot study of use-dependent learning in the context of constraint induced language therapy. *Journal of the International Neuropsychological Society*. 2006; 12:843–852. [PubMed: 17064447]
- Olness GS. Genre, verb, and coherence in picture-elicited discourse of adults with aphasia. *Aphasiology*. 2006; 20:175–187.
- Peach R, Reuter K. A discourse-based approach to semantic feature analysis for the treatment of aphasic word retrieval failures. *Aphasiology*. 2010; 24:971–990.
- Pulvermüller F, Neininger B, Elbert T, Mohr B, Rockstroh B, Koebbel P, et al. Constraint-induced therapy of chronic aphasia after stroke. *Stroke*. 2001; 32:1621–1626. [PubMed: 11441210]
- Raymer AM, Ciampitti M, Holliway B, Singletary F, Blonder LX, Ketterson T, Anderson S, Lehnen J, Heilman KM, Gonzalez Rothi LJ. Semantic-phonologic treatment for noun and verb retrieval impairments in aphasia. *Neuropsychological Rehabilitation*. 2007; 17:244–270. [PubMed: 17454696]
- Raymer AM, Singletary F, Rodriguez A, Ciampitti M, Heilman KM, Rothi LJG. Gesture training effects for noun and verb retrieval in aphasia. *Journal of the International Neuropsychological Society*. 2006; 12:867–882. [PubMed: 17064449]
- Robey RR. A meta-analysis of clinical outcomes in the treatment of aphasia. *Journal of Speech, Language, and Hearing Research*. 1998; 41(1):172–187.
- Rodriguez AD, Raymer AM, Gonzalez Rothi LJ. Effects of gesture+verbal and semanticphonologic treatments for verb retrieval in aphasia. *Aphasiology*. 2006; 20:286–297.
- Rose M, Sussmilch G. The effects of semantic and gesture treatment on verb retrieval and verb use in aphasia. *Aphasiology*. 2008; 22:691–706.
- Simmons-Mackie, N. Social approaches to aphasia intervention. In: Chapey, R., editor. *Language Intervention Strategies in Aphasia and Related Neurogenic Communication Disorders*. NY: Lippincott, Williams & Wilkins; 2008. p. 290-318.
- Taub E, Uswatte G, Pidikiti R. Constraint-induced movement therapy: A new family of techniques with broad application to physical rehabilitation--a clinical review. *Journal of Rehabilitation Research and Development*. 1999; 36 (3):237–51. [PubMed: 10659807]
- Taub E, Uswatte G, Mark VW, Morris DM. The learned nonuse phenomenon: Implications for rehabilitation. *Europa Medicophysica*. 2006; 42(3):241–256. [PubMed: 17039223]
- Webster J, Morris J, Franklin S. Effects of therapy targeted at verb retrieval and the realization of the predicate argument structure: A case study. *Aphasiology*. 2005; 19:748–764.
- Wertz, RT. A case of aphasia, apraxia of speech, and apraxia of phonation. In: Helm-Estabrooks, N.; Holland, A., editors. *Approaches to the Treatment of Aphasia*. San Diego: Singular Publishing Group; 1998. p. 25-38.
- Whitworth, A.; Webster, J.; Howard, D. *A Cognitive Neuropsychological Approach to Assessment and Intervention in Aphasia*. New York: Psychology Press; 2005.

Table 1

Western Aphasia Battery (WAB) results

P1	Drill		Generative	
	Pre	Post	Pre	Post
Spontaneous speech	14	15	14	15
Auditory verbal comprehension	7.3	7.9	8.3	8.6
Repetition	8.1	7.6	6.6	6.8
Naming & word finding	8	8.8	8.5	8.6
Aphasia Quotient	74.8	78.6	74.8	78

P2	Drill		Generative	
	Pre	Post	Pre	Post
Spontaneous speech	11	12	12	14
Auditory verbal comprehension	10	9.95	9.95	10
Repetition	5.1	4.4	4.4	4.2
Naming & word finding	7.8	8.5	8.2	8.2
Aphasia Quotient	67.8	69.7	69.1	72.8

Table 2

Verb naming results

P1	Drill			Generative	
	# items named correctly before tx	# items named correctly after tx	# items named correctly after tx that were not named before tx	# items named correctly before tx	# items named correctly after tx
All (n=96)	55	66	12	65	57
Treated (n=32)	18	21	3	n/a	n/a
Untreated (n=64)	37	45	9	n/a	n/a

P2	Drill			Generative	
	# items named correctly before tx	# items named correctly after tx	# items named correctly after tx that were not named before tx	# items named correctly before tx	# items named correctly after tx
All (n=96)	63	67	18	68	69
Treated (n=32)	18	28	11	n/a	n/a
Untreated (n=64)	45	39	7	n/a	n/a

Table 3
Comparison of narrative outcomes following Drill and Generative treatments for P1 and P2.

P1	DRILL TREATMENT			GENERATIVE TREATMENT		
	Pre Mean (SD)	Post Mean (SD)	Effect size (ES) or Mann-Whitney (p)	Pre Mean (SD)	Post Mean (SD)	Effect size (ES) or Mann-Whitney (p)
Lexical Variables						
Total verbs	6.0 (0)	4.67 (3.79)	NA	7.33 (2.52)	13.33 (2.21)	ES=2.38*
Verb diversity	4.67 (1.53)	3.33 (3.21)	ES<1	5 (2.0)	9.00 (3.46)	ES=2.00*
Verb form	3.67 (2.52)	4.33 (3.21)	ES<1	6.33 (3.21)	9.67 (4.04)	ES=1.04*
Sentence variables						
Lexical density	0.65 (.01)	0.55 (.30)	ES=7.29	0.41 (.07)	0.51 (.03)	ES=1.50
AS units	1.85 (.12)	1.83 (.15)	p>1	1.48 (.16)	1.77 (.39)	p=.03*
Grammatical utterances	2.67 (2.08)	3.00 (1.73)	ES<1	3.33 (.58)	6.33 (2.89)	ES=5.20*
Discourse variables						
Total words	24.67 (2.89)	16.67 (12.90)	ES=-2.77	37 (16.52)	48 (11.36)	ES<1
Total utterances	6.33 (1.53)	5.67 (4.04)	ES<1	10.67 (4.16)	12.33 (3.06)	ES<1
Local coherence	2.26 (.05)	2.45 (.48)	p=.066	1.89 (.62)	2.44 (.36)	p<.001*
Global coherence	2.58 (.52)	2.18 (1.02)	p=.075	2.49 (.46)	2.86 (.07)	p=.001*
Story line ratio	2.22 (2.41)	.22 (.38)	ES<1	0.47 (.46)	3.21 (4.16)	ES=5.95*

P2	DRILL TREATMENT			GENERATIVE TREATMENT		
	Pre Mean (SD)	Post Mean (SD)	Effect size (ES) or Mann-Whitney (p)	Pre Mean (SD)	Post Mean (SD)	Effect size (ES) or Mann-Whitney (p)
Lexical variables						
Total verbs	1 (1.73)	0 (0)	ES<1	1.67 (1.12)	5.0 (2.0)	ES=2.89*
Verb diversity	1 (1.73)	0 (0)	ES<1	0.67 (.58)	5.0 (2.0)	ES=7.51*
Verb form	.33 (.58)	0 (0)	ES<1	0.33 (.58)	2.67 (3.06)	ES=4.04*
Sentence variables						
Lexical density	0.39 (.05)	.40 (.04)	ES<1	0.52 (.22)	.44 (.05)	ES<1
AS units	1 (0)	1 (0)	p>1	1 (0)	1.67 (.23)	p=.002*
Grammatical utterances	0.33 (.58)	0 (0)	ES<1	0 (0)	1.67 (1.53)	NA
Discourse variables						

	DRILL TREATMENT			GENERATIVE TREATMENT		
	Pre Mean (SD)	Post Mean (SD)	Effect size (ES) or Mann-Whitney (p)	Pre Mean (SD)	Post Mean (SD)	Effect size (ES) or Mann-Whitney (p)
P2						
Total words	9 (.57)	6.33 (1.53)	ES<1	10.33 (2.08)	37.33 (12.58)	ES=12.97 *
Total utterances	3.33 (1.53)	2.67(1.15)	ES<1	2.33 (1.53)	7.33 (.58)	ES=3.27
Local coherence	2.33 (.58)	2.00 (0)	p>.1	1.44 (1.26)	2.89 (.19)	p=.049 *
Global coherence	2.5 (.50)	2.67 (.58)	p>.1	1.44 (1.26)	2.89 (.19)	p=.049 *
Story line ratio	0.5 (.87)	0(0)	ES<1	0.17 (.29)	3.39 (3.26)	ES=11.16 *

NA = Effect size could not be calculated due to a pre-treatment standard deviation of 0;
 * improvement of more than 20% but effect size < 2.5;
 * **& Bold** = effect size greater than 2.5 and > 20% improvement or p < .05 (Mann-Whitney)