NIH Public Access
Author Manuscript
Brain Lang. Author manuscript; available in PMC 2008 April 10.

Published in final edited form as:
Brain Lang. 2007 January ; 100(1): 53–68.

Widening the temporal window: Processing support in the treatment of aphasic language production

Marcia Linebarger[a], **Denise McCall**[b], **Telana Virata**[b], and **Rita Sloan Berndt**[b]

a *Psycholinguistic Technologies, Inc., Jenkintown, PA*

b *University of Maryland School of Medicine, Baltimore, MD*

Abstract

Investigations of language processing in aphasia have increasingly implicated performance factors such as slowed activation and/or rapid decay of linguistic information. This approach is supported by studies utilizing a communication system (*SentenceShaper*™) which functions as a "processing prosthesis." The system may reduce the impact of processing limitations by allowing repeated refreshing of working memory and by increasing the opportunity for aphasic subjects to monitor their own speech. Some aphasic subjects are able to produce markedly more structured speech on the system than they are able to produce spontaneously, and periods of largely independent home use of *SentenceShaper* have been linked to treatment effects, that is, to gains in speech produced without the use of the system. The purpose of the current study was to follow up on these studies with a new group of subjects. A second goal was to determine whether repeated, unassisted elicitations of the same narratives at baseline would give rise to practice effects, which could undermine claims for the efficacy of the system.

Keywords

aphasia; agrammatism; assistive technology; augmentative; computer; treatment; self-monitoring; narrative; multiple baseline

INTRODUCTION

There is increasing evidence that performance limitations – as opposed to loss of linguistic competence or inability to compute specific types of linguistic representations– figure

---

Address correspondence to: Marcia Linebarger, Psycholinguistic Technologies, Inc., 3 Old York Road, Suite 1#447, Jenkintown, PA 19046, E-mail: Linebarger@psycholinguistic-technologies.com, Telephone: 215.525.9905.

The communication system described in this paper uses methods and computer interfaces covered by U.S. Patent No. 6,068,485 (Linebarger & Romania, 2000) owned by Unisys Corporation and licensed to Psycholinguistic Technologies, Inc. (www.sentenceshaper.com). A potential conflict of interest arises because ML serves as Director of Psycholinguistic Technologies, which has released *SentenceShaper* as a commercial product. Therefore, ML has not participated in testing, scoring or analysis of raw data in the study reported here.

This work was supported by NIH grant R01-DC05629 from the National Institute on Deafness and Other Communicative Disorders to the University of Maryland School of Medicine. The authors are grateful to the administration and staff of the James Lawrence Kernan Hospital for their continued support of this research, to Dr. Hiram Brownell for advice on data analysis, and to three anonymous reviewers for valuable comments and suggestions. Data from this paper were presented at the 42nd meeting of the Academy of Aphasia, Chicago, IL, in October 2004. Correspondence should be addressed to Dr. Marcia Linebarger, Psycholinguistic Technologies, Inc., 93 Old York Road, Suite 1#447, Jenkintown, PA19046. Email: Linebarger@psycholinguistic-technologies.com.

prominently in aphasic language production disorders. The primary type of evidence for this claim comes from the variability of aphasic speech across different tasks (e.g., Isserlin 1922; Kolk and Heeschen 1992). These and other reports indicate that the same speaker's linguistic productions may vary quite markedly from occasion to occasion, suggesting that performance rather than linguistic knowledge is at issue (see also Bastiaanse 1994).

This performance limitations view has been articulated most fully by Kolk and colleagues, who sought to explain both the variability of patients' symptoms across tasks, as well as the apparent overlap of morphological symptoms in agrammatic and paragrammatic speech (Kolk & Heeschen, 1992). The Temporal Window Hypothesis (Kolk, 2006; Kolk & van Grunsven, 1985), holds that rapid decay and/or slow retrieval of linguistic information prevent the agrammatic speaker from holding sentence elements simultaneously in working memory, as would be necessary in order to integrate them into a larger structure. The hypothesis addresses patterns of impaired comprehension as well as production; we focus here upon the latter.

Individual symptoms, and especially their variability within a patient, are attributed to the means that patients adopt to overcome the processing limitation. In some cases (e.g., in tasks with great resource requirements) they may rely on the use of simplified syntactic frames, impoverished in morphology and/or phrase structure, thus resulting in fragmented, telegraphic speech that conveys an appropriate message. They may also attempt to employ self-monitoring before or during speech production, to prevent and/or correct errors.

Self- monitoring is not limited to error prevention. It may also lead to self-priming, because repeated attempts to produce the target utterance may increase the activation level for the relevant words and structures. When the utterance is finally initiated, this increased activation may allow more information to be recruited simultaneously, enlarging the temporal window. Furthermore, monitoring of material already produced may activate additional material through a "cloze" effect. Self-monitoring may therefore increase the structural elaboration and well-formedness of agrammatic speech, but may also be associated with decreased fluency as the speaker hesitates or fills the silence with repetitions or filler words while assembling an utterance.

The idea that the strategies patients adopt to overcome their temporal processing limitations may result in a tradeoff between structure and speech rate can be extended to other sorts of tradeoffs. The temporal window hypothesis also predicts a tradeoff between structural and semantic/conceptual information as a function of task demands. The literature on agrammatism contains numerous reports of reduced sensitivity to semantic features when the task requires the integration of syntactic information with semantic or conceptual information, both in production (Hartsuiker, Kolk, & Huinck, 1999) and in comprehension (e.g., Frazier & Friederici, 1991; Linebarger, 1989; Mauner, Fromkin, & Cornell, 1993).

Such a tradeoff may also underlie the phenomenon that is the focus of this research: aphasic individuals who are able to produce full sentences when describing single pictures are often found to produce more fragmentary and ill-formed utterances when attempting multi-sentence productions (e.g., Lesser, 1989; Mitchum & Berndt, 1994; Weinrich, Shelton, McCall, & Cox, 1997). One explanation for this effect, congruent with the Temporal Window Hypothesis, is that the increased conceptual complexity of multi-sentence discourse reduces the resources available for syntax, with a resulting degradation of structure. This phenomenon is not limited to agrammatic speakers but has been described for other types of patients as well. In fact, there seems to be no *a priori* reason to assume that resource limitations of one type or another are not widespread among aphasic speakers.

One source of evidence supporting the Temporal Window Hypothesis has come from studies of the impact of a "processing prosthesis," a computerized communication system that provides processing support but no linguistic information (Linebarger, Schwartz, Romania, Kohn, & Stephens, 2000). *SentenceShaper*™ supports language production by allowing the user to record words or phrases in her own voice, and by linking these recorded utterances to visual icons ("shapes") which can be clicked to replay the recorded segment, or dragged across the computer screen to be ordered into larger units such as sentences and narratives.[1] Because the system allows the user to replay utterances that she has already produced, information in working memory can be refreshed repeatedly. *SentenceShaper* may thus be seen as artificially enlarging the temporal window for language production by allowing the aphasic user to overcome limitations caused by slowed activation or exacerbated decay.

Improved maintenance of information in working memory may allow users to monitor their emerging sentences both before and after overt production (Levelt, 1989: pre-articulatory versus post-articulatory monitoring). Previously produced utterance fragments can be replayed by clicking their associated shapes, thus removing time limits on post-articulatory monitoring. In the absence of such external support, the rapid decay of linguistic information characteristic of aphasia would normally make it difficult if not impossible for the user to internally replay his previous utterances. *SentenceShaper* also increases the opportunities for pre-articulatory monitoring, because the user can rehearse a planned utterance *ad lib* before turning on the sound recorder. Most importantly, the system allows the user to employ both types of monitoring in tandem while producing a larger structure: by recording the structure in small segments initially, the user can replay these segments to keep them activated (post-articulatory monitoring) while he rehearses the production of the full structure prior to turning on the sound recorder (pre-articulatory monitoring). Both types of monitoring are observed in aphasic subjects using *SentenceShaper*: extended rehearsal (*sotto voce* or silent) before turning on the sound recorder, and repeated replays of previously recorded material.

As reported in Linebarger et al. (2000), aphasic speakers produced markedly more structured utterances while using *SentenceShaper* than they were able to produce without it, a kind of efficacy that we will term the "aided effect". Comparison of six subjects' narratives produced with and without the communication system revealed that utterances in the aided narratives were significantly longer and more sentential than in the unaided versions for five of the six subjects. These aided effects provide compelling support for the Temporal Window Hypothesis. No linguistic information whatsoever – no word finding assistance[2] or grammatical template – was provided by the system or by clinicians during the creation of these aided narratives. Thus the structural superiority of the aided productions provides a data point on the variability continuum that could not be observed under naturalistic conditions.

The use of this processing prosthesis has also been linked to treatment effects --specifically, to changes in connected **unaided** speech -- when the system has been used over a period of time for largely independent home practice (Linebarger, McCall, & Berndt, 2004; Linebarger & Schwartz, 2005; Linebarger, Schwartz, & Kohn, 2001). For some of the aphasic subjects in these studies, unaided narratives following treatment (narratives produced without use of *SentenceShaper*) contained longer, more sentential, and/or more grammatically well-formed utterances than the same narratives elicited prior to treatment. The nature and extent of these changes varied across subjects, ranging from no impact on language production to quite marked gains in both structural and content measures. It is notable that these treatment effects have been found in multi-sentence narrative tasks such as retellings of short, silent film clips. As

[1]The generic acronym *CS (Communication System)* was used in previous studies (e.g., Linebarger et al., 2000) to denote a research prototype of the communication system discussed in this paper. The final version of the program is termed *SentenceShaper*. For clarity, both versions of the program are referred to here as *SentenceShaper*, given their identical functionality. In the current study, three subjects (ER, HD, KR) used the prototype, while the other three subjects (MR, SL, NA) used the final version of the program.

[2]Versions of *SentenceShaper* used in treatment studies have incorporated some limited word-finding assistance, as described in this study; such assistance was disabled in the version of the system used in Linebarger et al. (2000).

noted above, aphasic speech often degrades in multi-sentence contexts; treatment interventions for aphasia rarely “scale up” from constrained tasks such as single picture description to more demanding and naturalistic tasks such as multi-sentence narrative production.

The current study investigates the impact of *SentenceShaper* use on the unaided narrative production of six non-fluent aphasic individuals. No attempt was made to select only patients with agrammatic speech, both because there are no universally-accepted criteria for this designation (see, for example, Saffran, Berndt & Schwartz, 1989) and because our goal was to investigate further the types of patients with production impairments who might benefit from use of the system. The impact of *SentenceShaper* use has been investigated in subjects with non-fluent aphasia, although not all of these subjects have shown characteristics of agrammatism (see Linebarger & Schwartz (2005), Linebarger et al., 2004). As noted above, the kinds of processing limitations that *SentenceShaper* was designed to ameliorate are unlikely to occur only in agrammatism. And because the system works in such a general way, by supporting the retention and monitoring of speech, its impact may be interestingly heterogeneous across different kinds of language impairments. For example, subjects with primarily LEXICAL as opposed to structural difficulties might use the system to monitor the semantic appropriateness of their utterances, and the impact of this self-monitoring might be revealed by measures of content rather than by more structural measures such as the length or syntactic complexity of their utterances.

One issue to be addressed here is the hypothesized tradeoff between conceptual and structural processing. Prior to the introduction of *SentenceShaper*, the same two narrative samples are elicited over an extended baseline period. We assume that repeated elicitation (accompanied by repeated presentation of the target material) should reduce the conceptual processing demands of the task by increasing familiarity with the specific story lines of the two narratives. Under the tradeoff hypothesis outlined above, this reduction of conceptual processing demands might be expected to allow for greater syntactic elaboration.

These repeated narrative elicitations at Baseline also allow us to assess an issue of considerable methodological importance: the feasibility of using repeated elicitations of the same narratives to assess treatment gains. If narratives are elicited with the same materials both before and after treatment is applied, what appear to be treatment effects may actually represent practice effects for these repeated stories. Interestingly, few studies have addressed this point, and those that have examined the impact of repeated elicitations (e.g., Coelho, Liles, & Duffy, 1990; Ernest-Baron, Brookshire, & Nicholas, 1987) have focused upon information content rather than structural properties of the samples produced. The extended baseline period in the current study, using the same two narratives, is designed to address this issue.

METHODS

Subjects

Six participants with chronic aphasia are reported in this study. All had been aphasic for at least one year (range=1–9.5 years) and had completed inpatient and outpatient speech therapy. All patients were non-fluent speakers (wpm > two standard deviations below control mean. Demographic and language performance information for these subjects is presented in Table 1. Patients are roughly ordered in the table based on the degree of apparent structural impairment (Proportion of Words in Sentences, Mean Sentence Length, Proportion of Sentences Well-formed and Ratio of Closed class to Open class Words) in narrative speech at Baseline. These measures are derived from the Quantitative Production Analysis (Saffran et al., 1989) methods applied to a baseline retelling of *The Snowman* (Briggs, 1989). A description of the elicitation method, scoring criteria, and reliability data are presented below. Four of the patients were left-handed premorbidly, and two of these four had experienced an intracranial hemorrhage (ICH), on the right for ER and on the left for NA. All remaining patients had suffered left hemisphere infarctions.

Assessment of the participants' auditory comprehension and spoken production abilities was completed at the outset of this study using the Philadelphia Comprehension Battery (Saffran, Schwartz, Linebarger, Martin, & Bochetto, 1988), an action/object naming test (Berndt, Mitchum, Haendiges, & Sandson, 1997), and narrative speech samples. An estimate of short-term memory was also obtained for four of the patients using a word repetition task (Martin & Saffran, 1997).

Appendix A contains a transcription of the first three minutes of each subject's initial narrative sample at Baseline, as well as a comparable transcript from a normal control subject.

Materials

The communication system—The main work screen of *SentenceShaper* is shown in Figure 1. To operate *SentenceShaper*, the user records utterances by touching the **On/Off Buttons** that control a sound recorder. Each recording is associated with a unique icon, a colored **Shape,** which immediately appears in the area above these buttons. In Figure 1, there are three shapes in what we will term the **Work Area**. Touching a shape causes its associated utterance to be replayed. Shapes can be assembled into sentences by dragging them from the Work Area into one of the ordered slots in the **Sentence Assembly Area**. In the Figure, there are five shapes, and two unfilled slots, in this area. Unwanted recordings can be discarded by dragging them to the **Trash.** The shapes in the Sentence Assembly Area can be played back to allow the user to evaluate the sequence and re-record or re-order the sentence elements. The component shapes in these "sentences" can then be concatenated and linked to an oval "bean" icon in order to create multi-sentence productions in the **Narrative Assembly Area** using a similar incremental procedure. In the Figure, there are two "beans" in this area.

The version of *SentenceShaper* used in Linebarger et al. (2000) did not provide word-finding support, because the goal of that study was to test the hypothesis that processing support alone would produce an aided effect. For use as an augmentative communication device and for treatment studies, however, the system does provide assistance with word-finding. In the current study, lexical support was provided by up to 36 labeled **Side Buttons**. When touched, these buttons play the words or phrases whose text appears on the buttons; the user must record the spoken word himself in order to incorporate it into his production. These buttons can be customized for individual subjects, but in all treatment studies with *SentenceShaper* they have contained prepositions, high-frequency, semantically flexible verbs, and (in some cases) pronouns or subordinating conjunctions. In the study reported here, the Side Buttons were the only word-finding support provided by the system.

Elicitation materials—The narrative elicitation method used in this study required participants to watch short, silent video stories and then to retell the stories in their own words, without correction or feedback. The videos used included a silent video adaptation of a children's book, *The Snowman*, edited to approximately nine minutes; an adaptation of a children's book, *The Velveteen Rabbit* (Williams, 1987), and a short animated film, *The Christmas Story* (Ilja Novak, private distribution), each edited to approximately five minutes in length.

Procedures

Aphasic participants—A schematic of the timeline of elicitation and testing procedures for the aphasic subjects is shown in Figure 2.

1. Elicitation of narratives prior to *SentenceShaper* training: One video, *The Snowman*, was viewed and described by the aphasic subjects at the outset of the study (partial data from that administration can be found in Table 1), and was not seen again until all training had been completed. To examine the impact of repeated, unaided narrative production prior to the introduction of *SentenceShaper*, participants produced spoken narratives to describe the other two videos (*The Velveteen Rabbit, The Christmas Story)* weekly for eight consecutive weeks. The order of presentation of the two videos was alternated across sessions, and no feedback was given as patients produced their narratives. Responses were tape recorded for later analysis.

2. Training in use of *SentenceShaper*: Participants received weekly 1–1½ hour sessions of *SentenceShaper* training in the lab. The total number of training sessions varied across participants, ranging from six to 13, depending on the ease with which they learned the system. (The relationship between ease of learning the system and outcome is discussed below.) Training focused on instructions regarding the basic mechanics of using *SentenceShaper*, including how to record, manipulate, and replay productions. Instruction was also provided regarding system features such as using the Side Button vocabulary to expand, correct or complete partial sentences (see Linebarger & Schwartz, 2005, for discussion of the use of Side Buttons to facilitate sentence construction). Participants produced short narratives using *SentenceShaper* to describe a variety of stimuli including multi-action pictures, sequential wordless picture books, and a number of short (one to five minute) videos that had been edited from various silent films. No feedback was provided regarding missing story elements or text coherence in terms of the order of events, use of correct referents, and so forth. Lab training was discontinued when subjects were able to perform the basic *SentenceShaper* mechanics and consistently used system features that improved their productions.

3. Home use of *SentenceShaper*: Following lab training, subjects were provided with a computer to use at home and were encouraged to use *SentenceShaper* on a daily basis to produce short narratives. Stimuli for home practice included the same types of materials used in training, with an increasing emphasis placed on more complex materials such as video clips and retellings of television shows viewed at home. Subjects were permitted to review and revise productions across several days. Caregivers were asked to avoid providing assistance or giving feedback regarding the content or structural accuracy of the user's productions. Participants continued to attend weekly lab visits during the period of home use. The narratives that they had recorded on *SentenceShaper* at home were collected using a zip disk. During lab visits, subjects were asked to demonstrate use of *SentenceShaper* to assure continued use of relevant system features. Suggestions and feedback designed to refine their use of the system were provided as needed.

Spoken narratives (the same stories as used at Baseline) were elicited without use of the system every four to five weeks during the time period when participants were using *SentenceShaper* at home, with the goal of collecting at least four narratives that could be compared to Baseline. These assessments thus provided multiple data points collected at intervals during the period of *SentenceShaper* home practice, not only at the end of such practice. We refer to these assessments here as "post-*SentenceShaper*".

The period of *SentenceShaper* home practice ranged across patients from 11 to 23 weeks. Motivation and performance factors were considered in determining when to discontinue home training. Home use of the system was discontinued if a subject reported that he was no longer using the system at home on a regular basis, or if few narratives were recorded on *SentenceShaper* at home between lab visits, or if performance measures remained unchanged over three post-*SentenceShaper* assessments. (The relationship of amount of home use and outcome is discussed below). Subjects who continued to use the system at home, and/or who continued to improve on the repeated performance measures, were allowed to retain the computer for several additional months. These subjects consequently contributed more than the targeted four data points to the evaluation of the effect of *SentenceShaper* use on performance. One exception to this protocol was Subject KR, who relocated and had to leave the study after the third post-*SentenceShaper* assessment. The number of post-*SentenceShaper* assessments contributed by each patient is indicated below in Table 2.

Elicitation of narratives from normal subjects—All three videos were presented to a group of ten neurologically unimpaired control subjects (mean age 56; mean education 14.6 years) over four sessions. *The Snowman* was described in sessions 1 and 4; *The Velveteen Rabbit* and *The Christmas Story* were described repeatedly, across all four sessions. After viewing each video, participants were asked to re-tell the story, and were allowed to continue until they indicated that they were finished. No feedback regarding the accuracy or content of the story was provided. Participants' spoken responses were tape-recorded for later analysis.

Scoring and Analysis of narrative samples

All narratives (from aphasic subjects and controls) were analyzed using a revised and shortened version of the Quantitative Production Analysis (QPA; Saffran et al., 1989) and the Correct Information Unit (CIU) Analysis (Nicholas & Brookshire, 1993). These two instruments provide quantitative information on three aspects of production: **structure** (Proportion of Words in Sentences, Mean Sentence Length, Proportion of Sentences Well-formed; **content** (Proportion of Correct Information Units (CIUs) to Total Words); and three **speech rate** measures (Words per Minute, Narrative Words per Minute, CIUs per Minute).

QPA analysis—Detailed procedures for elicitation, transcription, segmentation and analysis (Berndt, Wayland, Rochon, Saffran, & Schwartz, 2000) were followed for all samples. Most QPA measures are based on "narrative" words, defined as words intended to advance the narrative, excluding comments to the tester regarding the task, false starts, unintended repetitions and material that is repaired (see Appendix A for examples). The QPA defines sentences minimally as a subject plus verb; copular constructions (e.g., "the snowman was…") are counted as sentences only if they are completed (e.g., "the snowman was hot"). Analyses are based on a sample of 150 narrative words, which is divided on the basis of structure and prosody into utterance segments for further scoring. Inter-scorer reliability was assessed on the segmentation of the narrative into utterances, which has been shown to be the aspect of QPA scoring that is most difficult (Rochon, Saffran, Berndt, & Schwartz, 2000). Reliability was assessed on a subset of 26 randomly-chosen samples (36% of the total samples collected for all subjects, at least four samples per subject). Baseline narrative samples as well as samples obtained during the home practice period were included. Mean inter-judge reliability for all selected samples ranged from 82% to 96%.

CIU analysis—The CIU analysis is designed to assess changes in the semantic content of subjects' connected speech samples. This analysis, which is based on the total number of words produced in the full sample (rather than on narrative words), counts as "Correct Information Units" those words that are judged to be accurate, relevant and informative in relation to the story being told. The analysis was carried out on a subset of 32 randomly chosen samples (25% of the total samples collected for all subjects) following published procedures. At least three samples for each subject were randomly selected, and baseline samples as well as samples obtained during the home use period were included. Mean inter-judge reliability for number of words and number of CIUs for all selected samples ranged from 87% to 99%.

RESULTS

1. Searching for “practice effects”: Establishment of a stable Baseline

Since the effect of *SentenceShaper* use was to be assessed through repeated elicitations of the same stories, it was necessary to establish that repeated retelling of the same narrative does not BY ITSELF result in reliable improvements (“practice effects”) in the structure and content of that narrative. It was also critical that a stable Baseline of performance be established prior to introduction of *SentenceShaper*. Therefore, each subject produced eight narrative retellings of the same two video samples, and these baseline narratives were analyzed as described above.

The C-statistic (Tryon, 1982) is frequently used to determine whether or not there is a statistically significant change in a time series. It has been shown to be reliable over as few as eight cases, and is often used to establish a stable Baseline. However, it has been argued that this procedure produces an unacceptably high probability that an effect will be found where none exists (i.e., a Type 1 error). Crosbie (1995) has demonstrated, using Monte Carlo simulations, that many types of statistical tests are invalid if time series data are positively autocorrelated, i.e., if data points that are contiguous in a series are more highly correlated with each other than they are with the mean of the series. Positively autocorrelated data are frequently a problem in longitudinal research with single subjects (Barlow & Hersen, 1984), because they indicate that the data points are not independent. Because the issue of independence is critical for other analyses to be presented below, we first analyzed the baseline data to determine the degree to which the data were autocorrelated.

For each subject, autocorrelations were computed over the eight baseline sessions for each of the seven measures generated by the QPA and CIU procedures. Autocorrelation coefficients at lag 1 ranged from −.583 to .353, which are smaller than the values required to indicate significant autocorrelation at $p<.05$, for both positive and negative correlation coefficients (Anderson, 1942).

Since the baseline data were not significantly autocorrelated, we computed the C-statistic separately for each language measure and each subject across the series of eight baseline sessions. The significance of this statistic is assessed for each subject by generating a z-score for each measure, requiring a $z > 1.64$ to achieve significance at $p<.05$. Of the 42 data sets analyzed, data from two subjects showed evidence of improvement at Baseline in all three of the speech rate measures. Patients HD and MR improved significantly in Words per Minute, Narrative Words per Minute and CIUs per Minute across the eight baseline testings. All other measures for all subjects exhibited a stable Baseline using this measure.

Control participants—Narrative samples for the same elicitation videos from the ten control participants were analyzed using the QPA and CIU procedures described above. Control speakers produced the same two stories four times over a period of several weeks to assess the stability of the measures employed here for normal subjects. Mean and variance estimates for the first session (to be presented below) were generally similar to data presented previously for the QPA (Berndt et al., 2000) and CIU systems (Nicholas & Brookshire, 1993). To determine whether or not repeated production of the same story affected normal performance, paired two-sample t-tests were computed for each measure that assessed the difference between the means for the group in the first (T1) versus the fourth (T4) session. There were no significant differences between test sessions for any of the structural or content measures, or for one of the rate measures (all t-values <1). However, the control mean of 163 Words per Minute at T1 differed significantly from the mean of 178 at T4 ($t(8)=-2.18$, $p=.03$, one-tailed); the mean for Narrative Words per Minute at T1 (143) was also significantly slower than the mean at T4 (160; $t(8)=-2.19$, $p=.03$, one-tailed).

Summary and Discussion—The QPA and CIU measures collected across the eight baseline sessions were generally stable. The only exceptions involved measures of rate of speech, which increased for two aphasic subjects during baseline testing. Interestingly, control speakers as a group also showed an increase in rate of speech over four sessions of producing the same narrative. It appears that practice in telling a specific story, without any other intervention, may increase speech rate in retelling that specific story (even for two of six aphasic speakers). This finding is consistent with the idea that enhanced conceptual level processing (as speakers became more familiar with the story lines over repeated testings) may result in improved speech fluency, at least for some patients. However, it also appears that the structure and content of utterances do not change simply from repeatedly producing the same story, suggesting that familiarity with the story, by itself, does not free up processing resources sufficiently to affect those aspects of production. This finding is also methodologically significant, as it indicates that the results of aphasia treatment studies are not likely to be confounded with improvement of structure and content stemming from repeated elicitation of the same stories.

2. Narrative production following introduction of *SentenceShaper:* The two Baseline narratives

The same two narratives employed at Baseline were elicited from the subjects (without use of *SentenceShaper*) approximately every four weeks during the period of home practice. We calculated autocorrelation coefficients for the post-*SentenceShaper* data for all seven QPA and CIU measures for the number of data points contributed by each of the six subjects. To accommodate the unequal Ns for the post- *SentenceShaper* assessments across patients, the significance of these coefficients was evaluated by using confidence limits of two standard errors, calculated separately for each data set. All coefficients at lag 1 fell within the confidence limits, indicating that the post-*SentenceShaper* data were not autocorrelated.

Because neither the Baseline nor post- *SentenceShaper* assessment data showed significant positive autocorrelation (thus indicating that the data points are independent), we compared the mean for each measure across eight baseline testings to the mean of the post-*SentenceShaper* assessments that were available for each subject, using an independent samples t-test. The number of assessments contributing to the post-*SentenceShaper* means for each participant is given in Table 2.

Structural, content and rate measures—Table 2 shows the Baseline and post-*SentenceShaper* assessment mean and variance for each subject for Mean Sentence Length, Proportion of Words in Sentences, Proportion of Sentences Well-formed, Proportion of Correct Information Units (CIUs) to Total Words, and three measures of speech rate. Means and variances for the Control group at Baseline are also shown in the Table (reported p-values are one-tailed).

Four subjects showed a significant increase following the introduction of *SentenceShaper* in the **proportion of words that were produced as part of sentences**, with Subjects HD and KR moving into the normal range on this measure. Patient SL was performing in the normal range even at Baseline. Three of the subjects showed an improvement in **sentence length**, with KR again moving into the normal range. Three patients showed no change on this measure. Finally, two subjects improved significantly in the **well-formedness** of their sentences with HD approaching normal levels.

Improvement in the content of patients' narratives, indexed by the **proportion of CIUs to total words**, was somewhat less striking than results for the structural measures (no patient approached normal), although three subjects showed statistically significant changes.

Two patients showed robust improvement of all three **rate of speech** measures, although again no aphasic speaker approached normal values. Recall that HD and MR had improved on the rate measures at Baseline; MR continued this improvement into the post-*SentenceShaper* assessment period, while HD did not. KR, who had shown stable rates during Baseline, demonstrated strong rate gains after using *SentenceShaper*. These effects are summarized in Table 3.

The effect of ease of learning the system, and amount of system home usage, on outcome—As noted above, there was substantial variability in both the time it took patients to learn to use the system effectively and in the amount of home use that patients experienced. Table 4 shows for each patient the number of hours required to learn the system, and number of hours of home use estimated from system logs. It is possible that the differential outcome across patients (summarized in Table 3) is related in some way to this variability. Also indicated in Table 4 is an index of outcome for each patient: the number of measures (0–6) on which the patient made significant improvement. Note that all patients except ER were able to operate the system with their premorbidly preferred hand (see Table 1).

With regard to the system learning rates, two patients stand out in terms of time required to learn the system. HD, who entered the study an experienced computer user, learned to use the system very easily; this patient also had a good outcome, showing increases in four measures after the home use period. In contrast, patient NA, who showed no gains following system usage, took considerably longer than any other subject to become proficient with the system. Aside from these two patients, ease of learning does not appear to be strongly related to outcome, with MR and KR (good outcomes, each improving on six measures) taking about the same number of hours to learn the system as ER and SL (moderate outcomes, improving on only one or two measures).

As mentioned earlier, two versions of *SentenceShaper* were used in this study. Both systems tracked subjects' usage patterns, but in somewhat different ways. The earlier version (used by ER, HD and KR) recorded only the start and stop of work sessions. The final version of the program (used by MR, SL and NA) recorded all user actions and therefore allowed for removal of periods of idle time (>5 min with no user input). Even though usage records from the earlier version were inspected to ensure that they included no long periods of idle time, it is possible that the usage data from the earlier version is slightly inflated compared to that from the later system.

With this caveat in mind, we can consider how amount of home use might predict outcome. As noted above, the time that patients were allowed to continue using the system at home was largely (except for KR) determined by patients' continued usage of, and/or apparent improvement on, the system. So patient ER, with by far the greatest number of hours of home use, remained engaged and active on the system for many weeks, while patients NA and SL practiced for shorter periods before discontinuing regular use at home. Since all three of these patients showed at best moderate improvement (making gains on 1, 0, and 2 measures, respectively) following *SentenceShaper* usage, it does not appear that amount of usage, by itself, predicts outcome. Likewise, only two of the three patients with better outcomes (making gains on 4–6 measures) showed substantial amounts of usage, with KR (who left the study early) having the second lowest usage. It is likely that individual subject characteristics, such as motivation and untested cognitive abilities, contributed importantly to both the difficulty patients had learning the system, and to the amount that patients used the system at home, and also had some effect on the ultimate impact of system use. This suggestion remains hypothetical, however, and future research will collect more precise data on both ease of learning and patterns of home usage.

Summary of effects—As summarized in Table 3, use of *SentenceShaper* over a period of home use was associated with significant improvements in the narrative production of two stories repeated eight times at Baseline, and a minimum of three times during the home use period. These results included significant gains in structural, content and rate measures for some of the subjects. Two patients showed strong and consistent gains in the three structural measures following *SentenceShaper* practice. HD and MR improved on all three structural measures, while ER improved only on proportion of words in sentences. Interestingly, KR, who was near normal levels at Baseline, and who left the study early and had relatively little exposure to the system at home, showed significant effects in all structural, content, and rate measures except well-formedness.

Patients NA and SL did not show structural improvements. However, SL showed significant change on the content measure and the content-related rate measure (both of which started from a very low Baseline). This individual was already producing a normal proportion of sentences in his connected speech at Baseline, and his difficulties appeared to be largely lexical/semantic in nature (see naming data in Table 1). Thus, it is not surprising that the system seemed to produce the most improvement on measures of content. NA demonstrated no gains. Some proposals about the outcomes for these two patients are offered below.

Although there was substantial variability across patients both in the time it took them to learn the mechanics of the system, and the number of hours they used the system at home, these differences did not appear to be major predictors of outcome.

3. Narrative production following *SentenceShaper* use: Performance with a non-practiced narrative

The improvements noted above were found after many repetitions of the same stories during the baseline period and after the start of *SentenceShaper* use. The absence of structural or content gains across eight baseline assessments in these repeated narratives demonstrates that practice alone could not have sufficed to bring about the impressive changes in structure or content observed after *SentenceShaper* use began. In order to assess the generality of these positive gains to a narrative that was not practiced in this way, we elicited the *Snowman* story at the very end of the study and compared the results to the baseline values for the same story. These samples were analyzed using an expanded set of QPA measures that included assessment of changes in lexical production (closed class words and obligatory determiners). The CIU analysis was also carried out on these narratives. Pre-post measures expressed as proportions (e.g., Proportion of Words in Sentences) were analyzed using 2 x 2 chi-square tests (e.g., words in sentences vs. words not in sentences); Mean Sentence Length was compared using paired t-tests carried out on equal-length pre-post lists of number of words in sentences. P-values reported are one-tailed.

Table 5 shows the *Snowman* Baseline vs. post-*SentenceShaper* data for all six subjects. Values in bold font indicate that performance is within one standard deviation of the normal mean. The performance of NA and SL on the unpracticed narrative was similar to their performance on the practiced narratives: NA made no gains whatsoever, even though several measures (well-formedness, determiner use and proportion CIUs) were far below normal. SL continued to show improvement on the content measure, but also improved on the unpracticed narrative in the proportion of words produced in sentences (reaching ceiling).

Of most interest is the performance of the patients who had shown significant structural improvements on the repeated narratives. Data from HD and MR on this unpracticed narrative mirror the data presented above (summarized in Table 3) for the practiced narratives. These two patients also showed significant improvement in the content of their narratives. KR, who had shown structural and content improvements with the practiced narratives, showed gains on this testing only for Proportion of Words in Sentences. However, KR's Mean Sentence Length declined only slightly from a Baseline that was within the normal range. All three of these patients also showed improvement on QPA indices measuring use of closed class words and/or obligatory determiners, providing information about factors that may be associated with some aspects of their structural improvement, especially well-formedness. Five of the six patients showed a significant change from Baseline in Proportion of Words in Sentences, a measure that seems especially to benefit from *SentenceShaper* practice.

These quantitative changes are illustrated in Appendix B for HD and MR, who made the most consistent gains on the unpracticed narrative. From these short samples, it is clear that both patients produced a larger proportion of utterances that were sentences, and produced longer and a larger number of well-formed sentences. [3]

GENERAL DISCUSSION

The data presented here support the argument that practice using *SentenceShaper* at home, without direct clinical supervision, is associated with structural and content improvements for some patients.[4] It is especially encouraging that these improvements can be demonstrated even in the challenging context of multi-sentence narrative (descriptions of videotaped sequences), and even for a narrative sample that was not practiced (through repeated elicitations) during the period of the study. Several other results are of interest:

Practice effects

Repeated elicitations (N=8) of the same narratives without the support provided by *SentenceShaper* did not result in any structural or informational changes in subjects' speech, even for the practiced narratives themselves. However, five of the six subjects did make gains in the structure and/or content of their unaided narratives AFTER the period of *SentenceShaper* use began, and these treatment effects were quite marked for several of the subjects.

The lack of structural or informational practice effects from the repeated narrative elicitations at Baseline – for both aphasic subjects and normal controls – is quite encouraging from a methodological standpoint. Since all but one of the aphasic subjects did make gains on these measures after treatment commenced, we can conclude that repeated, unaided elicitations of the same materials do not lead to structural or informational practice effects, even for patients with some capacity for improvement. This would appear to eliminate the need for the

---

[3]It might be argued that the increases in mean sentence length do not provide evidence of the kind of increase in syntactic complexity predicted by the temporal window hypothesis. Such an increase could be effected by structurally irrelevant verbiage, such as noun phrase conjoining ("**the boy and the girl and the snowman and the dog** went outside"). It is for this reason that we have included Appendix B, listing the first 20 utterances produced pre versus post-*SentenceShaper* for the two subjects with the most marked structural gains. The length increase observed in sentences produced after *SentenceShaper* use derives from many different kinds of grammatical elaboration, including: increased use of articles, subordinate clauses expressing purpose and temporal information ("while the parents are in the house"), infinitival direct objects, prepositional phrases serving as both adjuncts and arguments, and conjoined verb phrases. Note that conjoined verb phrases (e.g., "this boy comes in the house **and sits down**"), which increased following SentenceShaper use in several subjects, reflect the ability to maintain a long-distance relationship between the subject NP and the second VP conjunct ("the boy …. sits down") in order to inflect the verb appropriately. Unfortunately, however, the sheer heterogeneity of structural changes observed across subjects makes it unlikely that any single measure, such as an analysis of argument structure or embedding, could be used to capture this increased complexity with statistically significant results.

[4]It should be noted that this kind of processing support need not be provided by a computer program. For example, treatment gains were reported from a somewhat similar but non-computerized protocol (Peach & Wong, 2004). In this study, aphasic subjects retold narratives to a clinician who transcribed and then read aloud the patient's narrative, providing explicit feedback and affording an opportunity for the patient to correct the transcribed productions. However, computer-based support seems preferable for this purpose, as it is more cost-effective, allows unlimited time for practice, fosters independence, allows the user to work with his own recorded speech rather than written transcripts, and can be configured to provide lexical support such as the Side Buttons. Furthermore, it is not clear how the aphasic subject could convey to the clinician exactly what material he wished to have read back.

development of “comparable” materials to be used in treatment research, which would be very difficult to achieve in any event.

Trade-off with conceptual enhancement

As discussed above, it is frequently reported that treatment-induced improvement in the production of single sentences fails to “scale up” to narrative production. One explanation is that the added processing demands imposed by the conceptual complexity of multi-sentence production may narrow the temporal window available for syntactic processing. Thus it might be expected that increased familiarity with the story line (due to the repeated viewings of the videos and retellings of the stories) would reduce the processing load and allow larger structures to be assembled, with the result that structure should be more elaborated or well-formed. However, our failure to obtain STRUCTURAL practice effects from repeated unaided elicitations does not support this hypothesis. Over the course of these repeated elicitations, subjects developed considerable familiarity with the specific story lines of the two baseline narratives, yet no patient showed any trend toward structural improvement in the Baseline period. Rather, enhanced familiarity with the stories appeared to be reflected only in the increased SPEECH RATE of two of the aphasic subjects and the group of normal controls.

It is also somewhat surprising that enhanced familiarity with the stories over the eight Baseline assessments did not result in improvements in story CONTENT, as indexed by the CIU analysis. In contrast, use of *SentenceShaper* did produce content improvements for several subjects, both with the practiced narratives and with the *Snowman* story that was not practiced. Of particular interest are the gains made by SL, who produced only 29% relevant content words in Baseline testing. This subject’s speech was characterized by considerable semantic jargon and often included words of unknown origin. These types of errors were still found at the eighth Baseline testing, suggesting that familiarity with the plot did not exert any beneficial impact on his expression of content. After working with *SentenceShaper*, however, he demonstrated significant improvement (increasing to 46%). Some possible contributors to this outcome are discussed below.

Information maintenance and structural gains: the role of monitoring

The central tenet of the Temporal Window Hypothesis is that aphasic speakers fail at sentence production either because information is not activated quickly enough, or decays too quickly, to be effectively formulated into a structural whole. Limited memory testing was carried out only for four members of our group (see Table 1), and all of them showed a reduction of word span. One of the functions of *SentenceShaper* is to improve speakers’ ability to maintain information by allowing unlimited “refreshment” of working memory through replaying of completed fragments. It is not entirely clear how subjects take advantage of this opportunity for improved maintenance when the system is not being used, i.e., in unaided production. Four of the subjects made structural gains, albeit on a continuum from marked (e.g., MR) to minimal (ER). All of these speakers started the study with grossly reduced speaking rate, which is interpreted within the Temporal Window Hypothesis as a reflection of the patients’ attempts to covertly generate and repair utterances prior to articulation. It has been proposed that non-fluent subjects rely on such pre-articulatory monitoring rather than on post-articulatory monitoring, primarily because the latter requires maintenance of on-going spoken utterances in memory (see also Ooman, Postma, & Kolk, 2001). It is likely that one of the effects of *SentenceShaper* use is to improve speakers’ post-articulatory monitoring by largely overcoming this memory limitation during the construction of aided narratives. As a result, the “habit” of attempting to monitor output during sentence production may develop over a period of *SentenceShaper* use and may carry over into unaided speech.

The qualitative improvements shown in Appendix B support the interpretation that these gains resulted because speakers monitored their productions, noting and correcting errors in their speech and using previously uttered material to self-cue word retrieval. For example, patient MR, who achieved robust structural improvement, entered the study with very fragmented speech (see Appendix A and Table 1), producing the most reduced mean sentence length and closed class word production of any of the patients. As can be seen in Appendix B, his productions post-treatment were more sentential, longer and much better formed; and function words were frequently present. Note that this change was achieved without *any* explicit instructions about how the initial utterances needed to be revised.

One question raised by this account is why all subjects did not achieve the same types of gains from *SentenceShaper* use. One possibility is that some aphasic speakers may be so severely impaired in terms of working memory that the system cannot engender consistent improvement. That may be the case for patients ER and NA, whose memory span was greatly reduced. Clearly, more rigorous memory testing is required to determine if a specific pattern of memory impairment is a predictor of success using the system. It may be that the more severely impaired patients cannot benefit from *SentenceShaper* training, at least with the training protocol utilized in this and earlier studies.

Another possibility is that a patient may have difficulty with self-monitoring because of an inability to identify errors in general, even in the speech of others. The ability to detect grammatical violations in others' speech may be a prerequisite for successful use of the system. Patient NA, the sole participant to show absolutely no treatment gains, differed quite markedly from the other subjects in one respect: she was largely unable to detect violations of structure in spoken sentences (see Table 1). A patient who does not notice omissions and mis-selections of elements when listening to speech is unlikely to use *SentenceShaper* effectively. The system essentially allows the user to monitor her own speech as she would monitor the speech of another individual, i.e., it allows her to carry out post-articulatory monitoring without the attentional demands of doing so simultaneously with her own speech production. Interestingly, the one subject in the Linebarger et al. (2000) study who failed to show an aided effect from *SentenceShaper* was also the only one of six subjects to perform poorly on a grammaticality judgment task. If the ability to detect violations in the speech of another person is a prerequisite for successful use of the device, then this supports the argument that post-articulatory monitoring plays a critical role in *SentenceShaper*'s treatment effects.

There are some indications from our data that the type of post-articulatory monitoring that is possible with *SentenceShaper* (i.e., listening to one's own voice and detecting errors without the simultaneous demands of speech production) differs substantially from post-articulatory monitoring as originally conceived (i.e., listening to one's own voice and correcting errors while in the process of speaking.) This distinction seems most pertinent in understanding the outcome for patient SL, who showed no content gains during Baseline testing, but did demonstrate significantly improved content following system use. It was clear during the Baseline testing that SL was not "monitoring while speaking", because he produced very few pauses, repairs, clarifications or repetitions (see Appendix A). In fact, he seemed completely unaware of his often egregious semantic errors. When using *SentenceShaper*, however, he frequently replayed his recorded segments, rejected them, and recorded another attempt, indicating that he could detect errors in the speech he had recorded on *SentenceShaper*. A similar finding was reported by Maher, Gonzalez Rothi, & Heilman (1994) who described a jargonaphasic subject with good auditory comprehension and poor awareness of the errors in his speech. This patient, like SL, could detect errors when he listened to a tape-recording of his own voice but not while he was in the process of speaking. These findings presumably indicate another effect of resource overload: monitoring for errors while in the process of speaking may require more (or different kinds of) resources than the aphasic speaker can muster, even when the bare ability to detect speech errors is not impaired. The important point concerning SL's outcome results is that practice in "monitoring" without simultaneous production carried over to improvement in content in post-testing, when the system was not being used.

Clearly, questions remain concerning the abilities that support speech monitoring of various sorts; research using processing support systems such as *SentenceShaper* may contribute to finding some answers. In summary, the results of this study suggest that narrative production with computerized processing support represents an effective approach to the treatment of aphasic language production disorders. The strong likelihood that *SentenceShaper* works its effects primarily by strengthening self-monitoring provides additional support for the view that aphasic language production disorders reflect performance issues arising out of the slow retrieval and/or rapid decay of linguistic information.

References

Anderson RL. Distribution of the serial correlation coefficient. The Annals of Mathematical Statistics 1942;13(1):1–23.

Barlow, DH.; Hersen, M. Single Case Experimental Designs: Strategies for Studying Behavior Change. New York: Pergamon Press; 1984.

Bastiaanse R. Broca's aphasia: A syntactic and/or a morphological disorder? A case study. Brain and Language 1994;48:1–32. [PubMed: 7712146]

Berndt RS, Mitchum CC, Haendiges AN, Sandson J. Verb retrieval in aphasia: Characterizing single word impairments. Brain and Language 1997;56:68–106. [PubMed: 8994699]

Berndt, RS.; Wayland, S.; Rochon, E.; Saffran, EM.; Schwartz, M. Quantitative Production Analysis: A Training Manual for the Analysis of Aphasic Sentence Production. Hove, UK: Psychology Press; 2000.

Briggs, R. The Snowman. New York: Random House; 1989.

Coelho CA, Liles BZ, Duffy RJ. Contextual influences on narrative discourse in normal young adults. Journal of Psycholinguistic Research 1990;19:405–420. [PubMed: 2286935]

Crosbie, J. Interrupted time-series analysis with short series: Why it is problematic; How it can be improved. In: Gottman, JM., editor. The Analysis of Change. Mahwah, NJ: Lawrence Erlbaum; 1995. p. 361-395.

Ernest-Baron CR, Brookshire RH, Nicholas LE. Story structure and retelling of narratives by aphasic and non-brain-damaged adults. Journal of Speech and Hearing Research 1987;30:44–49. [PubMed: 3560898]

Frazier L, Friederici A. On deriving the properties of agrammatic comprehension: Syntactic structures and task demands. Brain and Language 1991;40:51–66. [PubMed: 2009447]

Hartsuiker RJ, Kolk HHJ, Huinck P. Agrammatic production of subject-verb agreement: The effect of conceptual number. Brain and Language 1999;69:119–160. [PubMed: 10447988]

Isserlin M. Uber agrammatismus. Zeitschrift fur die gesamte Neurologie und Psychiatrie 1922;75:322–416.

Kolk, HHJ. How language adapts to the brain: An analysis of agrammatic aphasia. In: Progovac, L., editor. The Syntax of Nonsententials: Multi-Disciplinary Perspectives. London: John Benjamins; 2006.

Kolk HHJ, Heeschen C. Agrammatism, paragrammatism, and the management of language. Language and Cognitive Processes 1992;7:82–129.

Kolk HHJ, van Grunsven MF. Agrammatism as a variable phenomenon. Cognitive Neuropsychology 1985;2(4):347–384.

Lesser, R. Some issues in the neuropsychological rehabilitation of anomia. In: Seron, X.; Deloche, R., editors. Cognitive approaches in neuropsychological rehabilitation. Hillsdale, NJ: Lawrence Erlbaum; 1989.

Levelt, WJM. Speaking: from intention to articulation. Cambridge, MA: MIT Press; 1989.

Linebarger, MC. Neuropsychological evidence for linguistic modularity. In: Carlson, G.; Tanenhaus, M., editors. Linguistic Structure in Language Processing. Dordrecht: Kluwer; 1989.

Linebarger MC, McCall D, Berndt RS. The role of processing support in the remediation of aphasic language production disorders. Cognitive Neuropsychology 2004;21:267–282.

Linebarger, MC.; Romania, JF. System for synthesizing spoken messages. Washington DC: U.S. Patent and Trademark Office; 2000.

Linebarger MC, Schwartz MF. AAC for hypothesis-testing and treatment of aphasic language production: Lessons from a processing prosthesis. Aphasiology 2005;19:930–942.

Linebarger MC, Schwartz MF, Kohn SE. Computer-based training of language production: An exploratory study. Neuropsychological Rehabilitation 2001;11 (1):57–96.

Linebarger MC, Schwartz MF, Romania JF, Kohn SE, Stephens DL. Grammatical encoding in aphasia: Evidence from a "processing prosthesis". Brain and Language 2000;75:416–427. [PubMed: 11112295]

Maher LM, Gonzalez Rothi LJ, Heilman KM. Lack of error awareness in an aphasic patient with relatively preserved auditory comprehension. Brain and Language 1994;46:402–18. [PubMed: 7514943]

Martin N, Saffran EM. Language and auditory-verbal short-term memory impairments: Evidence for common underlying processes. Cognitive Neuropsychology 1997;14(5):641–682.

Mauner G, Fromkin VA, Cornell TL. Comprehension and acceptability judgments in agrammatism: Disruption in the syntax of referential dependency. Brain and Language 1993;45:340–370. [PubMed: 8269330]

Mitchum, CC.; Berndt, RS. Verb retrieval and sentence construction: Effects of targeted intervention. In: Humphreys, GW.; Riddoch, MJ., editors. Cognitive Neuropsychology and Cognitive Rehabilitation. Hove, UK: Lawrence Erlbaum Associates; 1994. p. 317-348.

Nicholas LE, Brookshire RH. A system for quantifying the informativeness and efficiency of the connected speech of adults with aphasia. Journal of Speech and Hearing Research 1993;36:338–350. [PubMed: 8487525]

Oomen CCE, Postma A, Kolk HHJ. Prearticulatory and postarticulatory self-monitoring in Broca's aphasia. Cortex 2001;37:627–641. [PubMed: 11804213]

Peach RK, Wong PCM. Integrating the message level into treatment for agrammatism using story retelling. Aphasiology 2004;18:429–441.

Rochon E, Saffran EM, Berndt RS, Schwartz MF. Quantitative analysis of aphasic sentence production: Further development and new data. Brain and Language 2000;72:193–218. [PubMed: 10764517]

Saffran EM, Berndt RS, Schwartz MF. The quantitative analysis of agrammatic production: Procedure and data. Brain and Language 1989;37:440–479. [PubMed: 2804622]

Saffran, EM.; Schwartz, MF.; Linebarger, MC.; Martin, N.; Bochetto, P. The Philadelphia Comprehension Battery. 1988.

SentenceShaper [Computer software]. Psycholinguistic Technologies, Inc. 2006. http://www.sentenceshaper.com

Shelton, JR.; Martin, RC.; Yaffee, L. Investigating a verbal short-term memory deficit and its consequences for language processing. In: Margolin, D., editor. Cognitive Neuropsychology in Clinical Practice. New York: Cambridge University Press; 1992.

Tryon WW. A simplified time series analysis for evaluating treatment interventions. Journal of Applied Behavior Analysis 1982;15:423–429. [PubMed: 7142062]

Weinrich M, Shelton JR, McCall D, Cox DM. Generalization from single sentence to multisentence production in severely aphasic patients. Brain and Language 1997;58:327–352. [PubMed: 9182751]

Williams, M. The Velveteen Rabbit. New York, NY: Grosset & Dunlap; 1987.

Appendix A

Unedited transcripts of first three minute segment of each aphasic subject’s initial baseline narrative of *The Snowman*, with a two-minute control sample (words in bold print are “Narrative” words [excluding repetitions, repairs, false starts, comments, and conjunctions joining independent clauses]; see Saffran et al., 1989).

Control subject transcript (2 minutes)

WPM (full sample) = 159

It began with a young boy rolling snow apparently **to make a snowman he started with the** with the the **rolled snow** and then **he used a shovel** and **he built a really large snowman** and **he topped it off with another ball of snow and put a hat on him and a scarf and was very pleased with the snowman** and then he **he had to go in the house** and **he went in the house** and well **the house actually just appeared there behind the snowman** and **the mother was calling him to come in he went in** (3sec.) and **when he slept that night he was tossing fitfully not sleeping** and **he got up and looked out the window to look at his snowman** and **when he looked at the snowman he shivered he was either cold or he was thinking the snowman was cold** but **he went back to bed** and then **he still couldn't sleep** so he got **he put a robe on and went down the steps** and **it was just midnight** and he went outside oh no he **he looked outside through the door and saw the snowman** and **he looked at the clock again it was** you know **midnight** and that **that time he looked at the snowman again** and **it changed come over to the snowman** and **it became alive the boy ran out the snowman took him by the hand** and they went off **they went running off together** and then **they began to fly** and **they flew over the town** and **the other snowman waved** at the **as they flew by** and **a couple of them even went up to join them for a while** then **they were lying by themselves** and **they were going a great distance** and **you can see that style of building change looked almost Russian they were going further and further and then saw the Aurora Borealis so you figured they were heading perhaps to the north pole** and **they got there** and **there were a group of** snowman **snowmen all together**

Subject: ER

WPM (full sample) = 17

A little boy (4 sec.) **he tried the** (4 sec.) snow I (6 sec.) mean **snowman** I mean **eyes nose smile and carrots** (12sec.) I mean **a hat** (7 sec.) he slept (8 sec.) **he slept yawning go to bed woken up surprise and snowman is real** (8sec.) **he went in the house** and (3 sec.) **he saw something** (4 sec.) **outside taking a trip they flew and the wind** (8 sec.) **they saw something** (13 sec.)

Subject: MR

WPM (full sample) = 23

One he rolled up and **he rolled up and snowman** (3sec.) **body and face** and **he went inside to show** (7sec.) **mother** and **mother and child heat** heat and (3sec.) **back outside** (3sec.) **he made a** (6sec.) **snowman** ears no (3sec.) **mouth chest** and (3sec.) chest inside (6sec.) I (5 sec.) **he went inside and up to bed** (5sec.) **later snowman came to life** (9sec.) snowman came to life and back when back **snowman and boy back inside** (16sec.) **Christmas tree**

Subject: HD

WPM (full sample) = 22

Boy is is **making a snowman** (5 sec.) then he (4 sec.) **he rest** no rest **twelve o'clock** then **he went downstairs** and (5 sec.) **snowman was light it up** and (6 sec.) then **he brought snowman inside house** (13 sec.) then he **cat screamed** or screamed (5 sec.) then he (3 sec.) cat screamed then (13 sec.) **he remembered how hot it was listening to fire hot** then he (3 sec.) **boy and snowman** (12 sec.) snowman **picture** (9 sec.) **snowman has a picture of** /so/

Subject: KR

WPM (full sample) = 41

It's it is tell it like it is is is is **little boy** (5sec.) **set up a snowman** (3sec.) and **he** (3sec.) **going to have the dinner** and **he come back out to finish building the snowman** (8sec.) and it's built it's built in it's built **snowman is built in** (3sec.) you you I said it once oh you don't okay (5sec.) no anyway (12sec.) **he went** went in **inside to sleep** and **about twelve o'clock he came back out and** and **took** look to to the mat **snowman in the house and** and and and **showed him some** some **of the that he was playing with** (4sec.) then **they come back out and took the** took the took the **snowman out to play** (4sec.) and and **they started back in the house** and then and then **snowman**

Subject: NA

WPM (full sample) = 75

Okay got up **the little boy got up on the side** then **first look the window** then **he went out the doors** and **he walked** and **out there** (5sec.) **he** uh a [bawl] of [wa]…**a big ball of water** then **he pick the ball up** (10sec.) **snow** and then **he did the bottom** and then **he did the top** and then **he put the eyes and the ears and his mouth and his** uh **a smile** the **he put four buttons down the center** then **he went in the house** and **he got a old hat and a** (3sec.) **thing go round your neck** and then **he went back in the house and looked out the window see ifs** is **it's still there** and then **next thing you know he went outside** gotta get gotta get [di] got [dewah] then **outside he was out in the field** so **he** took him **took him outside** and then **pluck Christmas tree** and **they looked at the** [mismas] **Christmas tree** then **they stand with present look at it while** then he put uh **he sit down** (4sec.) sit by /hi/ (3sec.) and then **play for a while** then **they both outside start going through yard** (4sec.) going through yard and then **they went down where a path was at and kept walk on walking** then **they says them** in **in more** (4sec.)

Subject: SL

WPM (full sample) = 54

Retelling of the snowman **on Christmas day it was snowing out in Pennsylvania** and **the snowman** (3sec.) **caught the snow bunny in front of the house** and then **he started to go to bed** when Sam when Dad said **it's approximately eight thirty** (3sec.) and **if the snow bunny were right he was looking up at eight o'clock and fall back asleep** (3sec.) **twelve thirty at that time he come up to the snow bunny** (3sec.) **and** (4sec.) **showed** him his face and showed **him he was human being** (3sec.) and then **the snow bunny took him on a ride all the way through the blue sky and taught him he was going to** and then **he gave him** the bunny excuse me **the boy a gift of a** yellow (15sec.) a yellow a **blue scarf** and **the bunny could let him go back** (6sec.) with **him the bunny lifted him rack** and **he got back at which time the snow bunny** it was now twelve o'clock excuse me

Appendix B

Unpracticed, unaided narrative, first 20 utterances counted as narrative words for subjects HD and MR before and after the period of *SentenceShaper* use (utterances in **bold** type were scored as sentences).

Subject HD

Baseline	Post-*SentenceShaper* use
boy is making a snowman	**the kid going to make a stoman while the parents are in the house**

Baseline	Post-*SentenceShaper* use
he rest	**the snowman and the kid goes in the house to play with the toys**
he went downstairs	**he cat screams at the stoman**
snowman was light it up	**the snowman and the kid goes out to play with the other stoman**
he brought snowman inside house	**he flies with the kid over the neighborhood and flies to the other snowman at lights**
cat screamed	**they have a Christmas party for the kid**
he remembered how hot it was listening to fire	**reindeer comes over to**
hot	reindeer
boy and snowman picture	**Saint Nicholas takes the kid to play with the reindeer**
snowman has a picture of nature	**the kid unwraps it**
music is good him for boy	**he likes it**
snowman is	a scarf that is
snowman lots of friends and flying	**the snowman takes the kid back to the house and waits for Christmas to come**
twelve o'clock	**Christmas day the kid comes out for**
gift for a boy	watches the sun come up
scarf	no snowman
he or she snowman	dried up
boy went inside and slept	**the kid is sad**
parents mom and dad were eating breakfast	**he wants to play with the snowman but no more**
boy went outside to check on snowman	**whenever the boy is ready the parents sit inside**

Subject MR

Baseline	Post *SentenceShaper* use
he rolled up and snowman	**the snowman and the boy was out building the snowman**
body and face	**this boy comes in the house**
he went inside to show mother	**this boy comes in the house and sits down**
mother and child heat	**the mother shows him some soup**
back outside	**this boy comes out and builds the snowman**
he made a snowman mouth chest	**the boy probably comes out**
he went inside and up to bed	eyes nose ears mouth and kid gloves and scarf
later snowman came to life	**the boy comes in**
snowman and boy back inside	**the mother and father come in and show the boy the bedroom and go to sleep**
christmas tree	**this boy wakes up and twelve o'clock**
plugged in	**this snowman comes alive**
box	**this snowman and the boy come in the house**
put them in snowman	**this snowman and the cat was livid**
snowman box	**this snowman trips on the**
snowman and him went outside and whisked away north	**this boy and the snowman sit down**
snowman gathered him people	this snowman this coal fire and TV
snowman gathered	**his snowman was livid**
walked up and Santa	**the boy shows him the box where the snowman lays down**
snowmen	the snowman peas
boy and Santa scarf	**this snowman and the boy come out and travel**

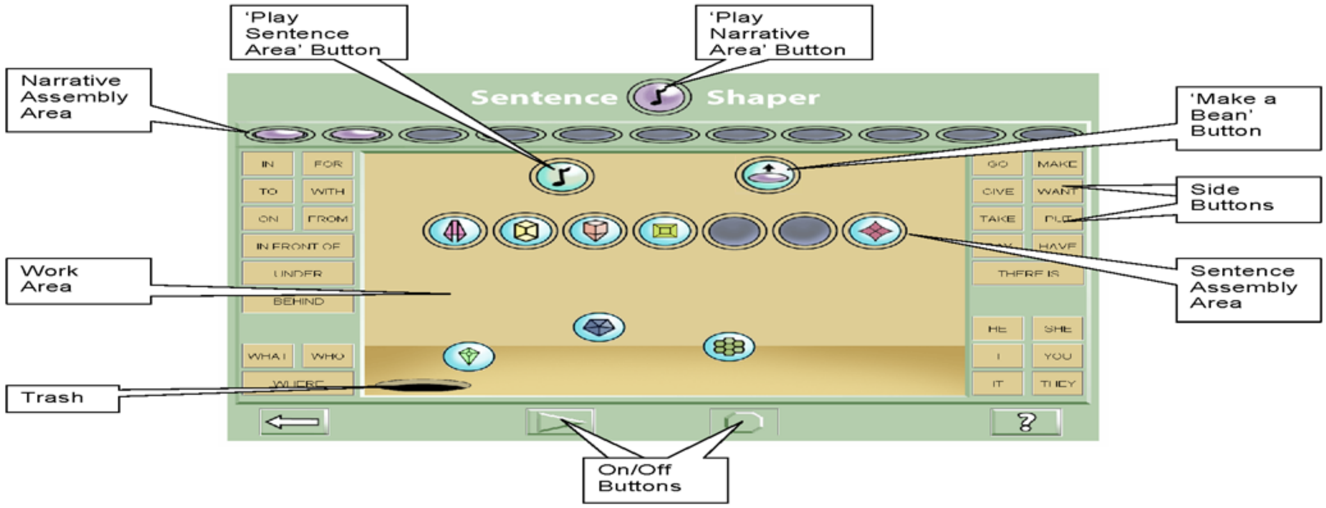

Figure 1.
SentenceShaper screen

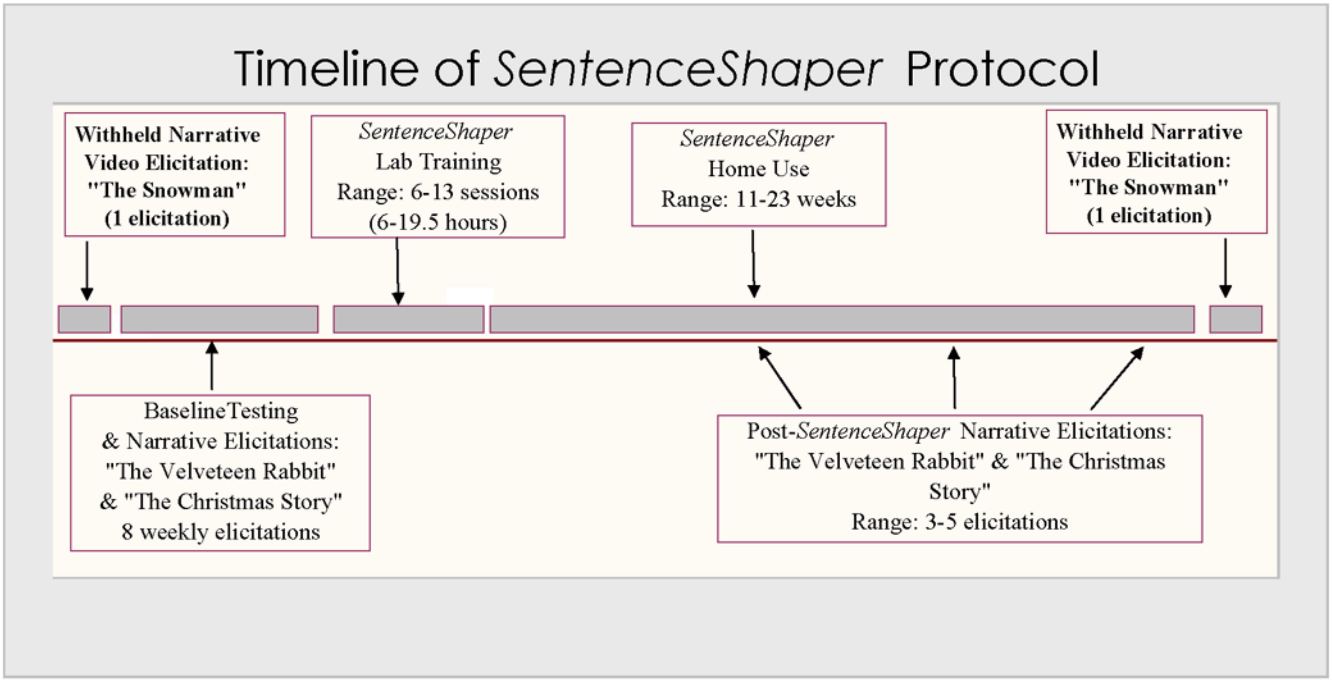

Figure 2.
Timeline for testing and training of aphasic participants

Table 1

Description of participants with aphasia (highlighted values are outside normal range)

	ER	MR	HD	KR	NA	SL
Age /gender	43/F	52/M	36/M	62/M	52/F	54/M
Time post onset (yrs)	9.5	1.5	2.5	1.5	1.5	1
Years education	12	14	16	12	12	18
Premorbid handedness	L	L	R	R	L	L
Lesion information	R ICH	L MCA frontal temporal	L MCA	L frontal temporal	L ICH	L MCA frontal temporal parietal
Hemiparesis	Left	Right	Mild R	Mild R	Right	Right
Hand used on *SentenceShaper*	Right	Left	Right	Right	Left	Left
Prop.Wds/Ss[a]	.57	.66	.82	.90	.93	.84
Mean S length[a]	5.12	4.5	5.59	7.32	7.63	10.77
Prop. Ss WF[a]	.41	.14	.05	.32	.47	.54
Prop. closed /open class[a]	.41	.37	.41	.48	.51	.49
Action/object naming[b]	.57/.93*	.73/.90	.90/.97	.77/.80	.80/.70	.77/.53
Grammaticality judgments[c]	.88	.93	.86	.97	.40	.81
Comprehension of reversible Ss[c]	0.52	0.9	0.92	0.65	0.8	0.88
Est. word span[d]	1.18	3.47	NT	NT	1.18	4.0

[a]Measures calculated from Baseline *Snowman* narrative speech sample using Quantitative Production Analysis procedures (described below). Normal values are presented in Table 5.

[b]Frequency and length matched words, N=30 per word type (Berndt et al., 1997).

[c]Subtests of Philadelphia Comprehension Battery (Saffran et al., 1988). Grammaticality judgments = proportion ill-formed sentences rejected (N=42); Comprehension of reversible sentences = sentence/picture matching using active, passive, locative and relative clause sentences (N=60).

[d]Word span estimated (following procedures of Shelton, Martin, & Yaffee, 1992), from repetition of two, three and four-word lists controlled for frequency and imageability (Martin & Saffran, 1997). HD and KR were not tested.

*$\chi^2(1) = 10.76$, $p<.01$

Table 2

Mean and variance for each patient on structural, content and rate measures tested at Baseline (n=8 sessions) and post-*SentenceShaper* training on repeated narratives (n=3–5 assessments) *The Christmas Story* and *The Velveteen Rabbit* (Data from ten normal control subjects on first elicitation of the same narratives)

	Mean S length	**Prop. Wds/Ss**	**Prop. Ss WF**	**Prop. CIUs**	**WPM**	**Narrative WPM**	**CIUs per min.**
CONTROL SUBJECTS mean(var)	10.33 (11.54)	.96 (.03)	.81 (.03)	.84 (.002)	164 (114)	146 (670)	129 (404)
ER							
mean (var) baseline	4.95(.40)	.71(.01)	.55(.01)	.45(.002)	25.75(19.31)	19.09(15.70)	10.94(9.77)
mean (var) post (N=5)	4.63(.09)	.82(.002)	.59(.02)	.48(.01)	29.30(23.02)	23.42(25.60)	13.67(21.50)
t(df)	t=1.02(11)	t=−2.16(11)	t=−.51(11)	t=.60(11)	t=−1.37(11)	t=−1.73(11)	t=−1.28(11)
p value	0.16	**0.03**	0.31	0.28	0.1	0.06	0.11
MR							
mean (var) baseline	5.44(.32)	.61(.010)	.16(.009)	.51(.008)	28.79(12.89)	20.96(11.13)	14.28(15.76)
mean (var) post (N=4)	6.85(.17)	.85(.006)	.59(.002)	.57(.0009)	37.42(6.51)	27.61(2.30)	20.88(2.18)
t(df)	t=−4.39(10)	t=−4.12(10)	t=−8.49(10)	t=−1.30(10)	t=−4.25(10)	t=−3.73(10)	t=−3.15(10)
p value	**0.0007**	**0.001**	**<.0001**	0.11	**0.0008**	**0.002**	**0.005**
HD							
mean (var) baseline	6.00(.23)	.87(.01)	.58(.01)	.66(.001)	27.86(12.52)	22.53(14.30)	18.57(4.79)
mean (var) post (N=5)	7.12(.33)	.96(.002)	.77(.02)	.76(.003)	25.85(5.52)	22.04(7.47)	18.21(25.96)
t(df)	t=−3.80(11)	t=−1.77(11)	t=−2.51(11)	t=−4.05(11)	t=1.12(11)	t=.25(11)	t=.18(11)
p value	**0.001**	**0.05**	**0.01**	**0.0009**	0.14	0.40	0.43
KR							
mean (var) baseline	8.23(.28)	.85(.003)	.39(.008)	.41(.006)	26.30(17.13)	14.92(10.17)	10.89(15.95)
mean (var) post (N=3)	10.44(.22)	.95(.004)	.46(.02)	.53(.004)	39.29(9.91)	26.23(14.97)	22.26(3.78)
t(df)	t=−6.30(9)	t=−2.55(9)	t=−1.01(9)	t=−2.29(9)	t=−4.87(9)	t=−4.99(9)	t=−4.62(9)
p value	**<.0001**	**0.02**	0.17	**0.02**	**0.0004**	**0.0004**	**0.0006**
NA							
mean(var) baseline	7.21(1.11)	.87(.003)	.46(.01)	.63(.004)	75.24(126)	60.46(96.39)	47.49(47.59)
mean(var) post (N=4)	6.62(.09)	.84(.002)	.54(.002)	.64(.003)	70.03(49.74)	55.00(80.10)	41.27(50.82)
t(df)	t=1.07(10)	t=.82(10)	t=−1.27(10)	t=−.10(10)	t=.84(10)	t=.93(10)	t=1.46(10)
p value	0.15	0.22	0.12	0.46	0.21	0.19	0.09
SL							
mean(var) baseline	8.67(1.35)	.95(.003)	.62(.02)	.29(.02)	59.58(49.72)	47.21(58.59)	16.79(62.87)
mean(var) post (N=4)	8.79(3.88)	.96(.001)	.69(.008)	.46(.02)	63.23(23.02)	52.53(66.94)	27.72(74.06)
t(df)	t=−.14(10)	t=−.34(10)	t=−1.02(10)	t=−2.27(10)	t=−.92(10)	t=−1.11(10)	t=−2.19(10)
p value	0.45	0.37	0.17	**0.02**	0.19	0.15	**0.03**

Table 3

Summary of significant effects of *SentenceShaper* training on structure, content and rate of speech for two repeated narratives: Comparison of means for Baseline vs. period of *SentenceShaper* use

	ER	MR	HD	KR	NA	SL
Structural Measures						
Mean Sentence Length	No effect	Effect	Effect	Effect	No effect	No effect
Prop. Words in Sentences	Effect	Effect	Effect	Effect	No effect	No effect
Prop. Sentences Well-formed	No effect	Effect	Effect	No effect	No effect	No effect
Content Measure						
Prop CIUs/Total Words	No effect	No effect	Effect	Effect	No effect	Effect
Rate Measures						
Words/minute	No effect	Effect	No effect	Effect	No effect	No effect
Narrative words/minute	No Effect	Effect	No effect	Effect	No effect	No effect
CIUs/minute	No effect	Effect	No effect	Effect	No effect	Effect

Table 4

System learning time and amount of system home usage

Patient	Hours to learn system	Number of measures on which patient made significant gains	Hours of home use
ER	13	1	69.8 (may include some idle time)
MR	10.5	6	26.4 (usage only)
HD	6	4	36.6 (may include some idle time)
KR	13	6	13.5 (may include some idle time)
NA	19.5	0	19.3 (usage only)
SL	11	2	11.0 (usage only)

Table 5

Means for QPA and CIU indices from unpracticed (*Snowman*) narrative for each subject before and after period of *SentenceShaper* use. Values in bold are within one standard deviation of the controls' mean.

QPA Index		ER	MR	HD	KR	NA	SL	Controls*
Words in sentences/total	baseline	0.57	0.66	0.82	**0.9**	**0.93**	0.84	.95 (.09)[&]
	post	.72[a]	.85[a]	**.91**[a]	**.99**[a]	**0.89**	**1.00**[a]	
Mean sentence length	baseline	5.12	4.5	5.59	**7.32**	**7.63**	**10.77**	9.55(2.3)[&]
	post	4.19	7.47[a]	**8.88**[a]	**7.19**	6.09	**10.33**	
Prop sent well-formed	baseline	0.41	0.14	0.05	0.32	0.47	0.54	.85 (.12)[&]
	post	0.65	**.76**[a]	.63[a]	0.33	0.55	0.67	
Prop closed class/total	baseline	0.41	0.37	0.41	0.48	**0.51**	0.49	.54(.04)
	post	0.49	.49[b]	**.56**[a]	**.57**[a]	**0.51**	**0.50**	
Obligatory determiners	baseline	0.7	0.11	0.33	0.76	0.85	0.86	.99 (.02)
	post	0.67	**.97**[a]	**.97**[a]	**1**	0.86	0.96	
Median length utterances	baseline	4	3.5	5	6	**7**	**7**	8.17(1.39)
	post	4	**7**	6	5.5	5	**7**	
CIUs/total words	baseline	0.49	0.48	0.58	0.45	0.58	0.46	.84(.05)[&]
	post	0.54	.59[a]	.76[a]	0.43	0.66	.64[a]	

*Control data from Rochon et al., 2000 (N=12); figures marked '&' represent new data collected for this paper (N=10)

[a] post *SentenceShaper* performance significantly improved over baseline, p<.01

[b] post *SentenceShaper* performance significantly improved over baseline, p<.05