Mapping therapy for sentence production impairments in nonfluent aphasia

Article

Published Version

It is advisable to refer to the publisher's version if you intend to cite from the work.

www.reading.ac.uk/centaur

CentAUR

Ψ Psychology Press
Taylor & Francis Group

Mapping therapy for sentence production impairments in nonfluent aphasia

Elizabeth Rochon, Laura Laird, Arpita Bose, and Joanne Scofield
University of Toronto, Toronto, Canada

This study investigated a new treatment in which sentence production abilities were trained in a small group of individuals and nonfluent aphasia. It was based upon a mapping therapy approach which holds that sentence production and comprehension impairments are due to difficulties in mapping between the meaning form (thematic roles) and the syntactic form of sentences. We trained production of both canonical and noncanonical reversible sentences. Three patients received treatment and two served as control participants. Patients who received treatment demonstrated acquisition of all trained sentence structures. They also demonstrated across-task generalisation of treated and some untreated sentence structures on two tasks of constrained sentence production, and showed some improvements on a narrative task. One control participant improved on some of these measures and the other did not. There was no noted improvement in sentence comprehension abilities following treatment. Results are discussed with reference to the heterogeneity of underlying impairments in sentence production impairments in nonfluent patients, and the possible mechanisms by which improvement in sentence production might have been achieved in treatment.

INTRODUCTION

A number of cognitive treatment studies have targeted sentence processing impairments in patients with nonfluent aphasia (e.g., Byng, Nickels, & Black, 1994; Schwartz et al., 1994). As a group, these have been termed "mapping therapies", as the theoretical underpinning for the approach to treatment has been the hypothesis of a "mapping deficit" (Schwartz, Linebarger, Saffran, & Pate, 1987) in the patients. In general, mapping therapy studies have directed treatment at the comprehension impairment, anticipating gains in verbal production as well (Byng, 1988; Byng et al., 1994; Jones, 1986; Le Dorze, Jacob, & Coderre, 1991; Marshall, Pring, & Chiat, 1993; Nickels, Byng, & Black, 1991; Schwartz et al., 1994). In this study, we investigated a treatment aimed directly at the production impairment in nonfluent aphasia, using a mapping therapy approach.

Agrammatic production

The sentence production impairment seen in nonfluent aphasia is often characterised by agrammatism, the absence of grammatical markers and syntactic structure. A number of different patterns and dissociations have been observed (e.g., Friedmann & Grodzinsky, 1997; Miceli, Silveri, Romani, & Caramazza, 1989; Miceli & Mazzuchi, 1990; Nespoulous et al., 1988; Saffran, Berndt, & Schwartz, 1989), including a distinction between morphological versus constructional impairments (Tissot, Mounin, & Lhermite, 1973; see Saffran, Schwartz, & Marin, 1980a, 1980b).

Difficulties in encoding relations for speech production, such as those seen in agrammatism can be accounted for in current models of normal sentence production (e.g., Bock & Levelt, 1994; Garrett, 1980, 1988). In Garrett's model, several separate levels of representation must be computed in order for sentence production to occur. In particular, two levels of grammatical encoding are critical prior to the final realisation of the words in a sentence by the speech motor system. The first is the "Functional level", where semantic aspects of words are represented. It is here that thematic roles, which determiner "who does what to whom" in a sentence, are assigned. The second is the "Positional level" where syntactic and phonological aspects of a sentence are assigned. It has been suggested that patients with agrammatism have difficulty mapping relations between the abstract functional level and the surface syntax at the positional level (Schwartz et al., 1987, 1994).

Asyntactic comprehension

Patients with agrammatic Broca's aphasia tend to have relatively spared comprehension abilities in comparison to the severity of their verbal production impairment (Goodglass, 1976). However, many of these patients

have difficulty comprehending semantically reversible and syntactically complex sentences (Berndt, 1987; Berndt, Mitchum, & Haendiges, 1996; Caplan & Futter, 1986; Caplan, Baker, & Dehaut, 1985; Caramazza & Zurif, 1976; Grodzinsky, 1986; Saffran & Schwartz, 1988; Schwartz, Linebarger, & Saffran, 1985). A sentence is semantically reversible when both actors in the sentence could conceivably carry out or be affected by the action of the verb. For instance, in a sentence such as, *The man is hugging the woman*, both the man and the woman can be the "doer" (i.e., the agent) of the action hugging, or the receiver (i.e., the theme) of the hugging. Knowing "who is doing what to whom" in a sentence requires understanding the thematic roles, in order to determine which actor is the agent and which is the theme of the action. The thematic roles must then be linked onto the syntactic positions of subject and object in a sentence. Reversibility contributes to comprehension difficulty especially where sentences have noncanonical thematic role assignment as is found in complex structures such as passive sentences. For instance patients often interpret a sentence such as *The woman is hugged by the man* as meaning that the woman is doing the hugging, presumably because they ascribe agent status to the first noun in the sentence (i.e., *the woman*) and theme status to the second noun (i.e., *the man*: see Caplan & Hildebrandt, 1988). Despite the above difficulties, agrammatic patients have demonstrated retained abilities to make accurate grammaticality judgements about sentences (Linebarger, Schwartz, & Saffran, 1983; Schwartz et al., 1987; but see Grodzinsky & Finkel, 1998).

It should be noted that not all agrammatic patients show the asyntactic comprehension pattern described above, and patients with different kinds of aphasia can also show this pattern (Berndt, 1991; Caplan & Hildebrandt, 1988). Various linguistic interpretations have been advanced (see Bradley, Garret, & Zurif, 1980; Caplan & Futter, 1986; Caramazza & Berndt, 1985; Grodzinsky, 1986, 2000), in addition to processing resource (Caplan & Hildebrandt, 1988) or working memory limitations (Miyake, Carpenter, & Just, 1994), as explanations for this pattern of comprehension impairment.

The mapping deficit hypothesis and mapping therapies

The mapping deficit hypothesis (Schwartz et al., 1987) holds that asyntactic comprehension is due to difficulty in assigning thematic roles to the parsed constituents in a sentence, thereby not achieving integration between sentence form and sentence meaning. In essence, the linking, or, "mapping" from the syntactic position (e.g., subject, object, etc., at Garrett's Positional level) to the thematic roles determining who does what to whom (e.g., agent, patient/theme, etc., at Garrett's Functional level), is faulty. Schwartz et al.'s explanation is a "procedural" variant of the mapping deficit, focused as it is on procedures that link thematic role information with

syntactic information in a sentence, and based largely on evidence from complex noncanonical sentences (Saffran & Schwartz, 1988; Schwartz et al., 1987, 1994). A lexical variant of the hypothesis focuses on impaired verb knowledge. In this form of a mapping deficit, patients are impaired even on syntactically simple (e.g., active) sentences (Byng, 1988; see Mitchum, Greenwald, & Berndt, 2000).

Mapping treatment studies have aimed to enable patients to utilise processing operations believed to the common to both comprehension and production; this is the "mapping" between the Functional and Positional levels as discussed above. On the view that there may be a shared common processing component between comprehension and production (i.e., thematic role mapping), mapping therapies have focused on treating the comprehension impairment, with predicted gains in production as well.

Effects of mapping therapy on comprehension and production of sentences. For most patients, improvements in sentence comprehension abilities have been found for trained sentence structures (usually reversible active sentences: see Mitchum et al., 2000, for an extensive review of findings from mapping therapy studies). In general, there has been little evidence for improvement of structures not explicitly targeted in treatment (but see Byng, 1988 and Schwartz et al., 1994, Patients IC, JH, GR, EW). Fink, Schwartz, and Myers (1998) have demonstrated a lack of generalisation to untreated structures regardless of whether treatment is blocked by sentence type (active, passive or object cleft sentences), or by thematic role query (e.g., agent or theme). In contrast, some patients have improved even on complex, noncanonical passive sentences, when these structures have been explicitly trained (Berndt & Mitchum, 1997; Haendiges, Berndt, & Mitchum, 1996; Jones, 1986; Schwartz et al., 1994, Patients EW, IC, GR, JH; Mitchum, Haendiges, & Berndt, 1995).

Some studies have investigated stimulus generalisation performance on the trained sentence types, but in situations or tasks other than those used in treatment (Wambaugh & Thompson, 1989). Results have been variable, with some studies reporting generalisation of treatment effects on novel tasks in some patients (Beveridge & Crerar, 2002; Byng, 1988; Haendiges et al., 1996; Jones, 1986; Mitchum et al., 1995; Patients EW and GG, Schwartz et al., 1994) and others reporting limited or no generalisation (Patient JG, Byng, 1988; Byng et al., 1994; Fink et al., 1998; LeDorze et al., 1991; Nickels et al., 1991; Patients IC, JH, GR, Schwartz et al., 1994).

Improvements in production also have been described subsequent to comprehension training for many patients (Byng, 1988; Byng et al., 1994; Marshall et al., 1993; Schwartz et al., 1994). Since there exist few tasks that elicit sentence production in a constrained fashion, investigators have been largely limited to examining production abilities on unconstrained

narrative tasks. Changes on narrative tasks have been reported for a variety of structural and morphological measures (Byng, 1988; Byng et al., 1994; Schwartz et al., 1994), but not necessarily for the production of the sentence types trained in therapy, nor with respect to the operation of thematic role assignment. Furthermore, in studies that did not allow spoken production by the patient in treatment, no improvements in production have been noted despite gains in comprehension (Berndt & Mitchum, 1997; Haendiges et al., 1996; Mitchum et al., 1995). It appears that claims regarding improved production abilities after comprehension training are subject to two concerns. First, there are difficulties in measuring improvements in sentence production relative to the treatment received in comprehension. Second, it also appears possible that the spoken practice allowed in comprehension training pro- grammes could have been responsible for the production gains (see Mitchum et al., 1995), leading Mitchum et al. (2000) to argue that to achieve gains in production, explicit treatment of the production impairment is most likely needed. This also raises the possibility that the demonstration of a functional dissociation between comprehension and production abilities could undermine the hypothesis that thematic mapping operations are shared between production and comprehension (Berndt, 1991, 1998; Byng et al., 1994; Mitchum et al., 2000).

Indeed several treatment studies and approaches have demonstrated the benefit of treating sentence production in patients with aphasia (e.g., Ballard & Thompson, 1999; Marshall, Pring, & Chiat, 1998; Thompson & Shapiro, 1994; Thompson, Shapiro, & Roberts, 1993; Thompson et al., 1997; Wambaugh & Thompson, 1989). Of the few studies that have had a similar theoretical underpinning to mapping (i.e., the link between the Func- tional and the Positional level), results with a small number of patients have provided evidence of structure-specific improvement with training on active sentences (Mitchum & Berndt, 1994; Mitchum, Haendiges, & Berndt, 1993) and passive sentences (Mitchum, Greenwald, & Berndt, 1997) using an approach called grammatical frames therapy. Weinrich, Boser, McCall, and Bishop (2001) have also successfully trained one patient (but not another) to produce passive sentences on an iconic computer-based communication system (C-VIC). We have found improved active and passive sentence pro- duction in one patient, but only after training on each sentence type separately and with no generalisation from one sentence type to the other (Rochon & Reichman, 2003).

As in the comprehension-based mapping therapy studies, there is evidence for structure-specific improvement in treatment studies that have targeted sentence production, but the question of reciprocal gains in sentence compre- hension remains unclear (Weinrich et al., 2001). For instance, we found no improvement in comprehension abilities in our patient whose production abil- ities successfully improved (Rochon & Reichman, 2003), and Jacobs and

Thompson (2000) have found cross-modal generalisation for comprehension training, but not for production training.

To summarise, comprehension-based mapping therapy studies have documented acquisition of trained sentence types, yet generalisation to untrained sentence types, to untrained tasks, and to an untrained modality (i.e., production) has been variable (see Fink, 2001, for review). Furthermore, despite the expectation of improvement in sentence production with a mapping approach, there has been no explicit mapping therapy study that has focused on training sentence production. Of the few studies that have had a similar theoretical underpinning to mapping, results with a small number of patients have also provided evidence of structure specific improvement with sentence production training (Mitchum & Berndt, 1994; Mitchum et al., 1993; Rochon & Reichman, 2003; Weinrich et al., 2001). However, stimulus generalisation (i.e., to untrained tasks) has rarely been examined and generalisation to comprehension has been found to be variable (Weinrich et al., 2001; Rochon & Reichman, 2003).

The aims of this study were threefold. First, we wished to investigate the effectiveness of a mapping therapy approach to directly target sentence production impairments. Given the variable success that has been found with generalisation to untrained structures in previous treatment studies, we were primarily interested in whether we could demonstrate successful acquisition of trained sentence types. We adapted and modified a well-established mapping therapy protocol for sentence comprehension impairments (Fink et al., 1998)[1], for the treatment of sentence production impairments. Semantically reversible sentences that had both canonical and noncanonical thematic role assignment were included. In contrast to most production studies, where only one, or at most two sentence structures were trained (but see Thompson et al., 1997), we used two canonical structures (actives and subject cleft) and two noncanonical (passive and object cleft) to examine the replicability of our results. The two cleft structures as well as the more standard active and passive structures were targeted for three reasons. First, active and passive sentences have been used extensively to test for and to treat syntactic comprehension and production impairments (e.g., Schwartz et al., 1994; Jones, 1986). Second, patients with aphasia have been found to perform in a similar manner on actives and subject clefts and on passives and object clefts (Caplan & Hildebrandt, 1988). Finally, object and subject cleft sentences have been targeted in previous mapping therapy protocols (Fink et al., 1998; Schwartz et al., 1994). We predicted acquisition of trained sentence types on both types of canonical and noncanonical sentences.

A second aim was to examine stimulus generalisation. To do this, we examined patients' abilities to produce the same sentence types trained in

treatment on tasks that were similar to, but different from, the treatment task. As mentioned, gains in production have been difficult to assess after mapping therapy, because generalisation measures were not always designed to measure the appropriate dependent variables and/or had very different task demands from the therapy task. Accordingly, we used constrained sentence production tasks that were designed to elicit the production of specific sentence types. To our knowledge, there exist very few constrained sentence production tasks. We chose two from the literature whose normative properties were well documented (Caplan & Hanna, 1998; Fink et al., 1994, 1995), whose elicitation procedures were different from the treatment task and from each other, and which contained some (but unfortunately not all) the sentence types to be treated in our protocol. Importantly, control participants have been shown to produce the sentence structures used in these two tasks equally well (Caplan & Hanna, 1998; Fink et al., 1994)[2]. We predicted gains on both these tasks for sentence types that had been treated in our therapy (i.e., actives, passives, object clefts), and on some sentence types that were not treated but were morphosyntactically similar to the treated sentences. Stimulus generalisation was also examined using a less constrained task (story narration), similar to previous studies. As in Schwartz et al. (1994), we predicted change on structural measures, but not on morphological or lexical measures.

The third aim was to examine changes in sentence comprehension abilities after treatment for production. The literature suggests that gains in production after mapping therapy for comprehension are difficult to evaluate (Mitchum et al., 1995, 2000) and gains in sentence comprehension abilities after production training rarely occur (Jacobs & Thompson, 2000; Rochon & Reichman, 2003). This question is important for the theoretical foundation of mapping therapy, which predicts cross modality gains after treatment.

METHODS

Participants

Five individuals participated in the study. Participants were referred by speech-language pathologists in hospitals and two aphasia centres in the Toronto area. All participants had suffered a single, left-hemisphere cerebrovascular accident. Participants were between 2 and 9 years post-onset. All participants were female. Three participants were randomly assigned to receive the treatment and two participants to the controls[3]. All participants

---

[2] Caplan and Hanna (1998) report that their test included subject–object relatives, however this structure was not reliably elicited from control participants. We did not include subject–object relatives in our use of the test.

[3] Unfortunately a third control participant could not be found within the timeframe of the study.

were native speakers of English, and all passed a hearing screening test. Table 1 summarises the participant characteristics.

As advocated by Schwartz, Fink, and Saffran (1995), we wished to determine whether our mapping therapy would be effective for patients presenting with a range of agrammatic symptoms. Accordingly, participants for this study were required to have chronic aphasia, to demonstrate a profile of symptoms consistent with Broca's aphasia on the Boston Diagnostic Aphasia Examination (BDAE: Goodglass & Kaplan, 1983), and to produce enough speech to be analysed on a number of measures described below. Table 2 summarises the participant characteristics and presents selected scores from the BDAE. In addition, performance on a number of other tasks was compared for each participant at baseline (i.e., pre-treatment) and at post-treatment to determine whether there was generalisation of treatment effects. These tests are outlined in the next section. Table 2 also shows data for two of these tests, one that examines comprehension (PCB: Philadelphia Comprehension Battery, Saffran et al., 1988) and another that examines sentence production (Caplan & Hanna, 1998). The data presented for these two tasks in Table 2 were collected in an initial screening session. They were subsequently re-administered with the other tasks described below for pre-treatment performance measures. The number of narrative words produced in a narrative retelling task (i.e., the Cinderella story; Berndt et al., 2000) is also included to show that participants had enough output to do this and other production tasks.

The profile for speech characteristics for all patients fell within the range of Broca's aphasia on the BDAE. As can be seen in Table 2, all participants were rated at a severity level of 1 or 2, due to the severity of their verbal production difficulties. The phrase length measure ranged from 2.5 to 4 for the study participants, with 4 being the upper end of the range usually found in Broca's

TABLE 1
Participant characteristics

Participant	Age (years)	Handedness	Education (years)	Years Post CVA	Aetiology
SM	74	R	9	5	L MCA infarction
QO	36	R	13	9	L MCA infarction
NS	31	R	16	2	L AVM
MJ[+]	32	R	16	4	L MCA infarction
PM[+]	82	R	12	2	L MCA infarction
Mean	51.0		13.2	4.4	
Median	36.0		13.0	4.0	
SD	24.9		2.9	2.9	

[+] Control participants with aphasia.

TABLE 2
Selected BDAE results and percentage correct scores for all participants
on selected tests from the screening battery

Test	Participants				
	SM	QO	NS	MJ[+]	PM[+]
BDAE[1]					
Severity (0–5)	2	1	1	1	2
Phrase length (0–7)	4	2.5	3	4	4
Auditory comprehension (mean percentile)	90	80	55	65	77
Philadelphia Comprehension Battery[2]					
Lexical comprehension (n = 44)	95	100	97	98	91
Grammaticality judgements (n = 60)	82	85	83	87	86
Sentence comprehension					
Lexical distractors (n = 60)	92	97	87	90	95
Reversible distractors (n = 60)	83	77	58	70	63
Sentence Production (Caplan & Hanna)[3]					
Sentence type					
Active (n = 5)	0	20	60	0	80
Passive (n = 5)	0	0	0	0	0
Dative (n = 5)	0	0	20	40	60
Dative Passive (n = 5)	0	0	0	0	0
Number of narrative words produced (QPA)[4]	155	153	111	104	177

[1] Boston Diagnostic Aphasia Examination (Goodglass & Kaplan, 1983).
[2] Philadelphia Comprehension Battery (Saffran et al., 1988).
[3] Sentence Production (Caplan & Hanna, 1998).
[4] Quantitative Production Analysis (QPA; Berndt et al., 2000).
[+] Control participants with aphasia.

aphasia. Participants also demonstrated a range in their composite auditory comprehension score, again, all within the range of Broca's aphasia. The two other tasks included in Table 2 provide a more detailed characterisation of the patients' sentence comprehension and sentence production abilities in particular. On the PCB, all participants showed grammaticality judgement and lexical comprehension abilities that were superior to their sentence comprehension abilities. Patients performed better in sentence comprehension when foil sentences contained lexical rather than reversible distractors, although this difference was minimal in patient SM. All were unable to produce noncanonical sentences on the Caplan and Hanna (1998) test (i.e., passive and dative passive sentences), although they were able to produce simple active sentences and dative sentences to varying degrees. Finally, except for NS and MJ, patients were able to produce the requisite 150 narrative words required for analysis of the Cinderella speech sample (Berndt et al., 2000). NS and MJ were included as they met all other criteria and it was deemed possible to analyse their narrative output.

Assessment battery for measurement of generalisation

Generalisation to two constrained sentence production tasks

Two constrained sentence production tasks were administered, which varied as to the method of elicitation. These two tests were Caplan and Hanna's Sentence Production Test (Caplan & Hanna, 1998) and the Picture Description with Structure Modeling Test (PDSM; Fink et al., 1995). Despite the fact that neither test elicited the production of subject cleft sentences, we deemed it worthwhile to use these measures because of their constrained nature.

Caplan and Hanna's Sentence Production Test.   This is a 20-item test in which five exemplars of each of four sentence structures were elicited. The target sentences were: active, passive, dative and dative-passive sentences. Stimuli were line drawings of actions and events. The test was designed to elicit particular syntactic structures and thematic roles, collapsed across a variety of verb argument structures.

Details of administration and scoring can be found in Caplan and Hanna (1998). Briefly, pictures were annotated with arrows designating all items in the picture that were to be mentioned in a sentence. With the presentation of each picture, the experimenter also provided the root form of the verb to be used in the sentence. A detailed scoring system accorded points for each correct noun, each correct thematic role, and correct verb morphology. In addition, points were given for use of the by-phrase in passive and dative passive sentences, as well as the locative phrase in dative passive sentences. To achieve full marks for a given target sentence all the necessary elements had to be present and correct. For instance, for the passive target sentence, *The cat was chased by the dog*, two points were awarded for each of the nouns, two points for each of the thematic roles conveyed, one point for correct verb morphology (auxiliary plus *-ed*), and one point for the by-phrase. If a patient said: *The dog chased the cat*, although they did not receive full marks for this sentence, they nevertheless received two points for the correct nouns and two points for correct thematic role assignment but no points for verb morphology or the by-phrase. If a patient said *Dog chasing by the cat* for the passive target above, they received two points for the correct nouns and one point for the by-phrase. In contrast, if a patient said *Dog chases by the cat* for an active target sentence (i.e., The dog chases the cat), they would be awarded two points for the correct nouns and one point for correct verb morphology. We tallied overall correct scores on this task by sentence type for each patient, and we also scored all of the sentence constituents separately for each sentence type.

Two of the same sentence structures as were treated in mapping therapy were elicited (actives and passives), and there were two sentence structures

that were not included in the treatment programme (datives and dative passives). This allowed us to track improvement on this test post-treatment on both treated and untreated sentence structures. Importantly, the scoring system allowed us to investigate whether change was evident on verb morphology scores, on thematic role scores, and/or on any of the other scores. On this task, we predicted improved sentence production of sentence structures trained in treatment. We did not have strong predictions with regards to change on the untreated dative and dative passive sentences. Improvement might have been expected on dative passives if participants acquired the passive rule in treatment. However, other variables, such as longer sentence length and different verb argument structures than those used in treatment might argue against expecting changes on these sentence structures.

Picture Description with Structure Modeling Test (PDSM).   This is a picture description task in which there were 10 exemplars of each of 6 different sentence structures. Three of the sentence structures were from among those trained in this study (active, passive and object cleft sentences). Three of the sentence structures were not trained in this study but were designed to be morphosyntactically similar to the mapping therapy sentences (nonreversible actives, truncated passives and embedded sentences).

The elicitation procedure entailed sequences of models of the targeted sentence structure presented by the examiner, followed by the test trials in which it was the participant's task to produce a target sentence. Trials were blocked by sentence structure. Details of the procedure and scoring can be found in Fink et al. (1995). Administration of three separate but equivalent forms of the test were randomised and balanced across all presentations.

Performance on this task post treatment allowed us to measure stimulus generalisation on novel exemplars of the same sentence structures as were targeted in treatment (except subject clefts) as well as on sentence structures that were not treated. We expected generalisation to untrained exemplars of the sentence structures trained in mapping therapy. In addition, as in Fink et al. (1995) we predicted generalisation to the untrained sentence structures, since they were morphosyntactically similar to the sentence targeted in treatment.

Generalisation to an unconstrained narrative retelling task

Narrative Production Task.   A story narration task was analysed pre- and post-treatment using the Quantitative Production Analysis (QPA; Berndt et al., 2000). A narrative speech sample was elicited by asking patients to retell a well-known fairytale ("Cinderella" after looking through a wordless picture book. Using the QPA, the narrative sample was then analysed for various lexical, syntactic and morphological elements. As in Schwartz et al. (1994), we predicted change on structural measures (e.g., "the

proportion of words in sentences", "the structural elaboration of sentences"), but not on morphological or lexical measures (e.g., "auxiliary complexity index", "inflection index") "percentage of closed class words").

Generalisation to comprehension of noncanonical sentence structures

Two sentence comprehension tasks were administered, which varied as to the method of elicitation. Both tests contained some or all of the sentence structures trained in treatment. One test also contained novel sentence structures, not targeted in treatment. These two tests were the Sentence Comprehension subtest from the Philadelphia Comprehension Battery (PCB; Saffran et al., 1988), and the Picture Comprehension Test, developed for this study.

Sentence comprehension subtest from the PCB.   The sentence comprehension test in this battery examines patients' comprehension of six different sentence structures. The structures are: active, active with a terminal phrase, passive, locative, object-relative clause, and subject-relative clause. There were 20 exemplars of each structure. Target sentences were semantically reversible. Each sentence occurred twice in the test: once with a syntactic (i.e., reversed-role) distractor, and once with a lexical distractor. On each trial, participants heard a sentence and were required to point to one of two pictures that correctly depicted the spoken sentence. As above, we predicted improved performance in comprehension of reversible sentences after production treatment.

Picture Comprehension Test.   We developed a sentence picture matching test to measure comprehension of the sentence structures treated in our sentence production therapy. This is a 60-item test comprising 15 exemplars of each of the four sentence structures treated in therapy: actives, subject cleft, passives, and object cleft sentences. Sentences were chosen from the pool of sentences that was used in our treatment programme, and from the PDSM test.

On each trial participants heard a sentence, spoken by the examiner, and were required to indicate ("yes" or "no") whether the sentence they heard matched the picture displayed. Pictures were paired with either a correct target sentence or with a syntactic or lexical foil sentence. Syntactic foils were semantically reversible sentences which depicted reversed-role distractors (e.g., a sentence such as *The girl was splashed by the clown* was presented with a picture of a girl splashing a clown). Lexical foils were sentences in which one lexical item had been changed. Sentences with lexical foils were not semantically reversible. Eighty percent of the foil sentences were syntactic in nature and 20% were lexical. There were three separate but equivalent

forms of this test. Administration of the different forms was randomised and balanced across all presentations.

The test was normed on 10 normal control participants, whose mean age was 65.9 years ($SD = 8.5$), with a mean of 15.4 years of education ($SD = 2.7$). The mean score of the control participants on the test was 98% correct. In keeping with the hypothesis that production and comprehension share thematic mapping operations, we predicted that participants' performance on this test should improve after treatment.

General overview and rationale of the design of the study

Briefly, the aim of the treatment programme was to help patients produce a variety of different sentence structures. Four levels of the treatment programme varied as to whether an Agent Cue or a Theme Cue was provided and which combination of four sentence structures was elicited. Pretest probes were also administered in every treatment session to track acquisition. The treatment programme is outlined in the Appendix. Progress through the treatment programme was measured by "between level" probe tests administered in between each level of treatment. Acquisition of all four trained sentences was measured by comparing performance for each patient on the between level tests administered before treatment began (i.e., baseline testing) to performance at the end of the programme.

As previously described, an assessment battery was administered before and after the treatment programme. It contained tasks that measured stimulus generalisation. These included two constrained sentence production tasks, a narrative story retelling task, and two sentence comprehension tasks. One constrained sentence production task (Fink et al., 1995), one sentence comprehension task and a video narration task (Rochon et al., 1994) were also administered with the between level probe tests in between each level of treatment. Lastly, follow-up of treatment effects was assessed four weeks after the end of the treatment programme.

Due to the difficulties of using traditional single-subject design with a mapping therapy protocol, we adapted a design used by Schwartz et al. (1994), with some modifications. We incorporated a number of features into the design of the study to achieve experimental control. First, to demonstrate treatment effects, we tracked performance on trained and untrained items with regular probes (i.e., the between level tests). In addition, the two control participants received all pre/post-testing and all between level probe tests that the treated participants did. Each control participant was yoked to a treated participant for the scheduling of the testing. The repetitive nature of the treatment and probe testing allowed for the possibility of unwanted learning effects due to repeated exposure of the same stimuli. For this reason, the sentences used in the pre-test probes, the between level

probes and in treatment were all from different stimuli sets, even though they represented the same sentence structures (the characteristics of the stimuli are explained below). Note that this made the acquisition measure quite conservative as we were measuring performance not just on trained sentences but on novel exemplars of trained sentence structures. In addition we used a non-parametric statistical test (McNemar test; see Siegel & Castellan, 1988) to test the significance of changes before and after treatment on our measures of acquisition and generalisation.

Stimuli

The four sentence structures trained in this programme were: active, subject cleft, passive, and object cleft sentences. Canonical (i.e., active and subject cleft) and noncanonical sentences (i.e., passive and object cleft) were also matched as closely as possible for length: active and passive sentences were matched, and subject and object cleft sentences were matched.

The sentences came from a large sentence bank of over 500 semantically reversible sentences (Mapping Therapy Sentence Bank; Schwartz, Fink, Rochon, & Myers, 1992). Sentences for all phases of the study were drawn from this set of stimuli. Stimuli used in different phases were distinguished by the verbs used in the sentences. Verbs were matched for frequency across sets. The nouns used in all the sentences were common, picturable nouns. The same nouns were used in all phases of the treatment programme, albeit in different combinations.

There were 144 sentences in the treatment phase, in which a set of six transitive verbs was used. The verbs were: *call, chase, hug, push, shoot, slap*. The mean frequency of the verbs was 146.8 (Francis & Kucera, 1982). Verbs in active and subject cleft sentences were in the present progressive (*-ing*) or simple present (*-es*) or past tense (*-ed*). Verbs in passive sentences were in the past tense. Verbs in object cleft sentences were in the past tense or in a present tense. Approximately 35% of sentences contained an adjective. Some examples of the sentence stimuli can be found in the Appendix.

The stimuli used to elicit sentence production were colour photographs. All the sentences were photographed with the actors in the sentence (i.e., subject and object) clearly distinguishable one from the other (actors wore costumes to achieve this) and with a clearly depicted action (e.g., *The cashier slapped the doctor*). Although some of the sentences were somewhat implausible, lexical items were chosen to minimise any confusion in the depiction of "who did what to whom", to maximise the likelihood that patients' attempts at production would be correctly scored vis-à-vis thematic role assignment (i.e., to avoid situations such as: "the woman slapped the woman"), and to keep the patients engaged in a task that was very repetitive. All photographs were taken against a plain backdrop.

Two small, black and white icons were used to reinforce the difference between the agent and the theme in a target sentence. The icons were different from each other (one with a circular pattern, the other with a linear pattern). The examiner placed each one above the person in the photograph who was identified as conveying each role, as the role was identified.

For ease of administration, the sentences (and photographs) were divided into six separate sets of 24 sentences each. The order of presentation of the sets of sentences was randomised such that each participant received a different random order of sets of sentences at every level of treatment. In addition, the random order of sentence sets was different for every participant across the entire treatment programme. The six verbs were balanced within and across sets at all treatment levels, and sets were balanced as closely as possible for the other lexical items that appeared in the sentences. The same method of randomising sets of stimuli was adopted for the between level probe testing, the pre-test probes and follow-up sentence stimuli.

Section 1. Treatment Programme

The rationale for employing the cueing approach and for the hierarchy of the cues was derived from previous experience with the sentence query approach in mapping therapy for comprehension (Fink et al., 1998). In Level 1 of our mapping therapy the participant was first trained to recognise and use only one cue, the agent cue, to produce a sentence (only active and subject cleft sentences were elicited at this level). In Level 2 the participant was again trained to recognise and use only one cue, but this time it was the theme cue (passives and object clefts were elicited at this level). As a preparatory step to the last level of treatment, Level 3 introduced the identification of both roles, but in a fixed order. Here, although role identification was in a fixed order, random role order was used in requesting participants to begin a sentence (i.e., agent or theme), which allowed for elicitation of all four sentence structures. Only at Level 4, the last phase of treatment, was role identification varied and randomised such that either an agent or a theme cue could be provided to elicit a target sentence of one of the four sentence structures being trained. At this level, the lexical item corresponding to the role not identified was also provided (e.g., "The other person is . . .", see part D of the Appendix) so that the participant was provided with all the lexical items with which to produce the target sentence. However, she was always asked to begin her sentence with the identified role. It was our expectation that the gradual introduction of the two different roles, in this stepwise procedure, would allow the participants to recognise and use the necessary thematic role information for correct production of a given sentence structure. On the other hand, it also meant that not until the last level of treatment were participants provided with cues about both thematic roles in random order.

For this reason, although the between level probes were administered between every level of treatment, we deemed the important measure of acquisition to be a comparison of performance before Level 1 of treatment and after Level 4 of treatment.

Procedures

Treatment sessions.   Each training session consisted of 12 sentences at Levels 1 and 2 (where only two sentences were elicited, respectively), and 24 sentences at Levels 3 and 4 (where all four sentence structures were elicited). There were thus six exemplars of each sentence structure treated in every session of treatment. To successfully complete a given level of treatment, participants had to achieve a predetermined criterion that was based upon a combination of their performance in the treatment session and on pre-test probes.

   Treatment and control sessions occurred twice a week for approximately one hour. All sessions were audiotaped. Participants' responses were scored for accuracy relative to the target sentence structure. Detailed scoring procedures were developed. Because the focus of treatment was on training of thematic role relations, participants were not penalised for changes or errors of verb morphology. For instance, for the target sentence *It is the woman that the waitress slaps*, we accepted the response: *It is the woman that the waitress slapped.* Omission of the auxiliary verb was also allowed (e.g., *The patient calls by the lawyer* was accepted for *The patient is called by the lawyer*). Note that in the latter case there is some ambiguity as to whether correct thematic role assignment is expressed. Although in adopting such a convention we ran the risk of inflating patients' abilities to correctly convey thematic role assignment especially in passive sentences, both constrained measures of generalisation required that the auxiliary verb be correctly employed in their scoring conventions.

   Control participants received all the same testing (between level probes, follow-up, pre–post-testing) as the treated participants. Controls were tested at the same time intervals as one of the treated participants. Control MJ was yoked to patient QO in this way, and control PM to patient SM. Control participants came twice a week for approximately one hour, for unstructured conversation about current events, etc. Every second session controls were given a narrative retelling task (Rochon et al., 1994).

Between Level Test.   This test represented the probes used to measure performance before treatment began as well as the test given in between levels of treatment. One hundred and forty four sentences were used in this test. The set of six transitive verbs employed was: *choke*, *follow*, *help*, *hit*,

kiss, *pull*[4]. The mean frequency of the verbs was 151.8 (Francis & Kucera, 1982). The other lexical items were the same as were used in the treatment programme.

Baseline testing occurred over four sessions prior to the beginning of treatment. Twenty-four sentences were administered at each of Levels 1 and 2, and 48 sentences were administered at each of Levels 3 and 4, for a total of 12 exemplars of each sentence structure. In all between level probes, half the sentence stimuli assessed production for the level of treatment that the participant had just terminated, and half the sentences assessed production for the next level of treatment. Limited assessment of trained items was also measured during each of the between level probe sessions in between every level of treatment.

Follow-up.   Four weeks after the completion of the treatment programme (i.e., after all post-testing), participants were tested on 32 sentences (8 exemplars of each of 4 sentence structures) to assess maintenance of treatment effects.

Reliability

Twenty percent of responses from the treatment sessions, the pre-test probes, and the between level probe tests were scored by both the primary examiner and an independent scorer. Point-to-point agreement between the two scorers was 98%. Twenty percent of Cinderella narratives were also scored for inter-rater reliability of both utterance segmentation (proportion of narrative words assigned to the same utterance) and scoring of the narrative samples based upon the QPA analysis system. Inter-rater reliability for utterance segmentation was 85%, and reliability for scoring of various measures on the QPA was 95% or greater. Inter-rater reliability was also calculated for scoring of the PDSM and the Caplan and Hanna sentence production tests. Scoring agreement was 99% for the PDSM and 98% for the Caplan and Hanna test.

RESULTS

Participants completed the treatment programme in 19 sessions on average, or in approximately 2.5 months. Between level probes took an average of 17 sessions per participant and all pre- and post-treatment testing took an average of 11 sessions. Thus, participants participated in the study for an average of 47 sessions, or about 6 months.

---

[4] Weinrich et al. (1997) have shown that patients may have difficulty generalising the past tense rule for irregular verbs. Because we were balancing our sets of verbs for frequency, picturability, etc., we had difficulty eliminating all irregular verbs from our stimuli. However, each set contained only one irregular verb.

Section 1: Treatment

Did mapping therapy result in the acquisition of all four sentence structures trained?

Successful acquisition of all sentences entered into treatment was noted for all treated participants at each level of treatment entered into, with little accompanying improvement noted for untrained structures at that level of treatment. Performance of the two control participants did not show the same pattern of overall improvement. Only one treated participant, NS, was entered into all four levels of treatment. Patient SM completed Levels 1 and 2 of treatment, after which baseline testing for Levels 3 and 4 revealed that treatment was not further indicated. Patient QO completed three levels of treatment (Levels 1, 2, and 4).

Table 3 compares performance on the between level probe test (i.e., on novel exemplars of trained sentence structures), broken down by sentence structure, before treatment, after treatment, and at follow-up. As can be seen in Table 3, all treated participants improved on treated structures after treatment. The McNemar change test was used to evaluate changes in scores before and after treatment. When collapsed across sentence types, the change in performance was significant for all treated participants, SM, $\chi^2 = 13.5$, $p < .001$; QO, $\chi^2 = 8.9$, $p < .01$; NS, $\chi^2 = 30$, $p < .001$, but not for either of the two control participants. Although not significant in the overall analysis, it is important to note that one control participant, MJ, improved on subject cleft sentences, and to an even greater extent on passive sentences. The second control participant did not improve on any sentence structures.

Follow-up data indicate that treatment gains were generally maintained four weeks after the treatment ended. As expected, there were no significant changes between post-treatment and follow-up data for any participants with the exception of NS ($\chi^2 = 5.4$, $p < .05$), whose score on object cleft sentences declined from 100% to 0 at follow up[5]. Examination of the data for NS's production of object cleft sentences revealed that all her attempts at producing object clefts were of the same form: she incorporated a by-phrase into her sentence (e.g., for the target sentence *It is the mother that the daughter calls*, NS produced *It is the mother that calls by the daughter*). Although it is possible that NS was attempting to convey the correct under-lying thematic knowledge, we did not give her credit for any of these sentences.

---

[5] NS's follow-up testing was delayed three weeks due to illness. It is possible that her performance would have been different at an earlier point in time.

TABLE 3
Performance of trained sentence structures (percentage correct) on between level tests pre- and post-treatment and at follow-up

Participants	Active			Subject cleft			Passive			Object cleft		
	Pre n = 12	Post n = 12	F/U n = 8	Pre n = 12	Post n = 12	F/U n = 8	Pre n = 12	Post n = 12	F/U n = 8	Pre n = 12	Post n = 12	F/U n = 8
SM	92	90	100	0	66	88	66	92	100	0	100	63
QO	83	92	100	83	92	88	33	83	50	0	50	63
NS	83	92	88	42	100	100	8	92	100	0	100	0
MJ+	92	83	88	33	67	100	0	92	100	17	0	0
PM+	100	100	75	50	8	88	75	50	75	0	0	0

Pre = Pre-treatment; Post = Post-treatment; F/U = Follow-up; + Control participants with aphasia.

Section 2: Measures of generalisation

Did treatment gains generalise to two other constrained sentence production tasks?

Caplan and Hanna Sentence Production Test.   Table 4(A) shows performance for each participant, collapsed across treated, untreated, and all sentence structures. As can be seen, the McNemar test showed that the change in performance after treatment was significant for two of the treated participants, SM, $\chi^2 = 7.0$, $p < .01$; QO, $\chi^2 = 9.0$, $p < .001$. Table 4 also shows the data for each participant collapsed across treated and untreated structures before and after treatment. When treated sentence structures (i.e., active and passives) were examined separately, the same two participants showed a significant difference before and after treatment, SM, $\chi^2 = 7.0$, $p < .01$; QO, $\chi^2 = 8.0$, $p < .01$. When untrained sentence structures were examined,

TABLE 4
Percent correct pre- and post-treatment on treated and untreated structures
on Caplan and Hanna Sentence Production Test

A: Overall score across all the sentences

Participant	Treated structures		Untreated structures		Total correct	
	Pre *n = 10*	*Post* *n = 10*	*Pre* *n = 10*	*Post* *n = 10*	*Pre* *n = 10*	*Post* *n = 10*
SM	0	70**	0	0	0	35**
QO	0	80**	0	10	0	45**
NS	30	40	0	10	15	25
MJ[+]	20	30	20	0	20	15
PM[+]	50	30	40	40	45	35

B: Sentence constituents on Caplan and Hanna Sentence Production Test

Participant	Nouns		Thematic role		Verb morphology		By-phrase	
	Pre *n = 20*	*Post* *n = 20*	*Pre* *n = 20*	*Post* *n = 20*	*Pre* *n = 20*	*Post* *n = 20*	*Pre* *n = 10*	*Post* *n = 10*
SM	90	80	75	65	15	60**	0	50
QO	80	85	60	70	0	70***	0	60*
NS	75	85	25	45	15	50	0	40
MJ[+]	75	75	35	45	45	40	0	70**
PM[+]	85	85	70	65	65	55	10	20

n = number of items; *significant at .05 level; **significant at .01 level; ***significant at .001 level; [+]control participants with aphasia.

none of the participants showed a significant difference before and after treatment. Interestingly, participants QO and NS in particular showed non-significant trends in the right direction, however the control participants did not.

The scoring system in this test allowed us to examine whether change had occurred on noun production, on thematic role assignment, on verb morphology scores, or on correct use of the by-phrase. These data are presented in Table 4(B). When analysed across all sentence structures using McNemar's test, there were no significant changes in noun production or in thematic role assignment for any of the participants. Two treated participants showed significant improvement on verb morphology scores, SM, $\chi^2 = 10.0$, $p < .01$; QO, $\chi^2 = 14.0$, $p < .001$, with participant NS showing a marginally significant difference, $\chi^2 = 4.45$, $p < .06$. One treated participant showed significant change in her use of the by-phrase, QO, $\chi^2 = 6.0$, $p < .05$, as did one control participant, MJ, $\chi^2 = 7.0$, $p < .01$. Participant SM showed a marginally significant difference, SM, $\chi^2 = 5.0$, $p = .06$. When treated sentences were examined separately, one treated patient showed a significant change in her verb morphology scores, QO, $\chi^2 = 8.0$, $p < .01$. Participant SM and control MJ showed a marginally significant difference, SM, $\chi^2 = 5.0$, $p = .06$; MJ, $\chi^2 = 5.0$, $p = .06$. This change in verb morphology scores was also seen in one treated patient for untreated sentences as well, QO, $\chi^2 = 6.0$, $p < .05$, and was marginally significant in the other two treated patients, SM, $\chi^2 = 5.0$, $p = .06$; NS, $\chi^2 = 5.0$, $p = .06$. Examination of the data showed that the verb morphology changes appeared to hold across all sentence structures, with only minimal variation, whereas the by-phrase changes were accounted for by changes in the use of the by-phrase in passive but not dative passive sentences.

The lack of change on the thematic role score may be somewhat surprising, given the emphasis on thematic role relations in treatment. Close examination of the data revealed that all participants scored at or near 100% for thematic role assignment for active sentences before treatment. Participants differed in their performance on passives, datives and dative passives. SM achieved high marks; QO and PM were variable; and NS and MJ did poorly on thematic role assignment scores on the three sentence structures before treatment.

Picture Description with Structure Modeling (PDSM).   Table 5 shows performance on the PDSM test before and after treatment[6], for each participant, collapsed across treated, untreated, and all sentence structures. As can be seen, the McNemar test showed that the three treated participants

---

[6] Both the PDSM test and the Picture Comprehension test were administered during "between level" tests, as well as before and after all phases of treatment. Examination of the data for the three treated participants for between level performance on both tests revealed variable patterns of improvement across participants for treated and untreated sentence structures after each level of treatment. For both tests, only the data for before and after treatment are reported here.

TABLE 5
Percent correct pre- and post-treatment on treated and untreated structures on
Picture Description with Structure Modeling (PDSM)

Participant	Treated structures		Untreated structures		Total correct	
	Pre $n = 30$	Post $n = 30$	Pre $n = 30$	Post $n = 30$	Pre $n = 30$	Post $n = 30$
SM	20	63***	27	60*	24	62***
QO	30	47	37	53	34	50*
NS	20	53**	27	73***	24	63***
MJ[+]	24	57*	40	60	32	59*
PM[+]	27	37	43	27	35	32

n = number of items; *significant at .05 level; **significant at .01 level; ***significant at .001 level; [+]control participants with aphasia.

improved significantly in their overall performance after treatment, as did one control participant, SM, $\chi^2 = 15.1$, $p < .001$; QO, $\chi^2 = 6.25$, $p < .05$; NS, $\chi^2 = 19.6$, $p < .001$; MJ, $\chi^2 = 11.6$, $p < .05$; PM, $\chi^2 = .17$, $p = .84$. We found that the pattern of improvement after treatment held for two treated participants when sentence structures trained in treatment were examined separately, SM, $\chi^2 = 11.3$, $p < .001$; NS, $\chi^2 = 9.0$, $p < .01$, as well as for one control participant, MJ, $\chi^2 = 8.3$, $p < .05$. The change was marginally significant for participant QO, $\chi^2 = 5.0$, $p = .06$. Changes were also significant for two of the treated participants when untreated sentences were examined, SM, $\chi^2 = 5.0, p < .05$, and NS, $\chi^2 = 10.9, p < .001$. Inspection of the data indicated that the significant improvement on trained sentence stimuli seen in the participants appeared to be accounted for mostly by the improved production of active and passive sentences, as most participants showed little improvement on object cleft sentences in this task after treatment. For the two participants whose performance improved significantly on untrained sentences, their improvement appeared to have been largely on truncated passive and embedded sentences. Inspection of PDSM test results for control participant JM across repeated administrations of the test showed a steady improvement in correct passive sentence production with only minimal improvement in active sentence production and little other change.

Were structural changes seen on an unconstrained narrative retelling task?

Narrative Production Task.   Table 6 shows performance on selected measures of the QPA applied to the Cinderella task. The table also includes values for 12 normal subjects taken from another study (Rochon, Saffran,

TABLE 6
Performance on selected measures from the narrative task

Measures	Normal controls[1] X (SD)	SM Pre	SM Post	QO Pre	QO Post	NS Pre	NS Post	MJ[+] Pre	MJ[+] Post	PM[+] Pre	PM[+] Post
Median utterance length	8.17 (1.39)	4.5	5	2	2	3	4	2	4	5.5	5.5
Structural measures											
% Words in sentences	98 (5)	87	95	35	47	50	75	64	75	97	95
% Well-formed sentences	95 (8)	64	81	38	53	18	77	80	77	88	81
Sentence elaboration index	3.06 (0.64)	1.12	1.18	1.31	1	0.9	0.97	1.7	1.4	1.34	0.98
Morphological measures											
Aux complexity index	1.26 (0.26)	1.43	1.31	0.67	1	0.67	1.05	0.55	0.55	1.21	1.38
Inflection index	0.92 (0.15)	0.8	0.54	0.67	0.55	0.67	0.67	0.43	0.5	0.83	1
% Closed class words	54 (4)	51	49	32	26	50	38	42	41	60	60
Lexical measures											
Proportion verbs	0.48 (0.06)	0.46	0.41	0.27	0.32	0.43	0.43	0.23	0.33	0.51	0.47

[1] data taken from Rochon, Saffran, Berndt, and Schwartz (2000); [+] Control participants with aphasia.

23

Berndt, & Schwartz, 2000). In applying a criterion for improvement beyond two standard deviations seen in the control data for each measure, Table 6 shows that treated participants improved consistently on structural measures (with the exception of the sentence elaboration index score). The median length of utterance (MedLu) remained relatively stable (although control MJ's MedLu doubled), while the percent of words in sentences and the percent of well-formed sentences increased for all treated participants after treatment (SM's improvement on percent words in sentences fell just short of the criterion, however her score and control participant PM's score were high before treatment). Control MJ showed an increase in the percent words in sentences but not the percent of well-formed sentences. In comparison, none of the changes on morphological or lexical measures exceeded what would be expected with normal variability. Table 7 presents eight randomly

TABLE 7
Eight randomly chosen examples of treated participants' utterances
from Cinderella narratives before and after treatment

SM Pre-treatment	SM Post-treatment
1. she sisters	1. the stepmother had three girls
2. they were miserable for her	2. the sisters were invited
3. she had to work hard cleaning up the mess	3. twelve o'clock she slip out
4. she cannot go to the party	4. she had lost a glass slipper
5. pumpkin was a carriage	5. she was happy
6. she has	6. the stepmother was not satisfied
7. she was miserable	7. the fairy godmother appeared because Cinderella was not happy
8. fairy godmother was	8. she was happy in ball

QO Pre-treatment	QO Post-treatment
1. her daughters and her	1. the mailman is giving letter to Cinderella
2. fairy godmother	2. mailman gave it to Cinderella
3. coach	3. everybody was in step Cinderella
4. next	4. beautiful
5. mother step daughter poor	5. fairy godmother dress
6. late	6. wicked stepmother
7. bell	7. stepsister gave rip
8. what kind of	8. mice

NS Pre-treatment	NS Post-treatment
1. Cinderella cleaning again	1. Cinderella was washing the windows
2. Cinderella clean	2. sewing
3. Cinderella has no skirt	3. make a wish
4. the good person	4. Cinderella will go to the party
5. try it on	5. every whole town wears
6. the mother and three sister is bad	6. the prince like to meet him again
7. Cinderella and I was dancing with	7. the prince
8. the weather	8. would you like to wear shoes

selected examples from Cinderella narratives for the three treated partici-
pants, comparing their production before and after treatment.

Was there improvement seen in the comprehension of noncanonical sentence structures?

Sentence Comprehension from PCB and Picture Comprehension Test.   Table 8 shows performance before and after treatment, for each par-
ticipant, on the Sentence Comprehension subtest from the PCB, and on the
Picture Comprehension Test. As can be seen, despite variability among the
patients, performance for all participants on both sentence comprehension
measures changed very little after treatment, with no significant differences
found for any participants using McNemar's test. Close examination of
performance on the PCB test revealed that participants' scores for reversible
sentences were brought down on this test, due not only to poor performance
on passive sentences, but also to poor performance on several of the other
sentence structures included, such as object relatives and locatives. However,
there was little change for any of these sentence structures after treatment.

Examination of the data for the individual participants on the Picture
Comprehension Test revealed that all participants performed better on
active and subject cleft sentences than the passive and object cleft sentences
before treatment, and that for two of the treated participants (SM and QO)
performance on the canonical sentences was nearly at ceiling. Despite their
poor comprehension performance on noncanonical sentences before treat-
ment (53% overall), improvement in production of noncanonical sentences
was not mirrored in comprehension. It is interesting to note that in control
participant MJ, whose production of passive sentences improved over time,
there was no concomitant improvement on this comprehension test for

TABLE 8
Performance pre- and post-treatment on comprehension tasks (percent correct)

Participants	PCB[1] Sentence Comprehension Test (reversible sentences)		Picture Comprehension Test	
	Pre	Post	Pre	Post
SM	76	82	81	82
QO	88	80	83	92
NS	55	53	58	60
MJ[+]	67	63	60	67
PM[+]	63	68	73	73

[1] Philadelphia Comprehension Battery (Saffran et al., 1988); [+] Control participants with aphasia.

passive sentences, even though it too was administered at repeated intervals (like the PDSM test), and even though MJ had one of the lowest scores on passive sentences before treatment, leaving room for improvement over time.

DISCUSSION

In this study we targeted sentence production using a mapping therapy approach. In addition, we included constrained sentence production tasks (i.e., the PDSM and the Caplan and Hanna Test) to measure stimulus generalisation. Our findings substantiate earlier studies (Mitchum & Berndt, 1994; Mitchum et al., 1993, 1997; Rochon & Reichman, 2003; Weinrich et al., 2001) in which cognitively-based treatments have been shown to be successful in training sentence production of even complex sentence structures. They extend earlier findings by demonstrating the feasibility of applying the mapping therapy approach to sentence production, and by showing that treatment gains generalise to trained and some untrained sentence structures on novel tasks. Our findings also raise a number of questions. Among these are the improvements seen in one control participant and the degree to which her results qualify the present findings; another relates to the possible underlying mechanisms that may have been at work in the treatment; and a third is the question of the relationship between sentence production and sentence comprehension.

In contrast to mapping studies that focused on comprehension, the gains seen in sentence production in the patients who received treatment in this study appear to be stronger and more consistent across participants than in previous studies. Of course, as mentioned, previous mapping studies for comprehension did not include constrained elicitation tasks to measure gains in stimulus generalisation; therefore it is possible that patients in these studies might have shown improved sentence production for specific, treated sentence structures after treatment, had these sorts of measures been included. However, this seems unlikely, given the relatively modest gains seen in most patients after treatment with comprehension-based mapping therapies (Byng et al., 1994; Fink et al., 1998; Jones, 1986; Nickels et al., 1991; Schwartz et al., 1994). It must be mentioned that a weakness of the present study was that generalisation for subject cleft sentences could not be measured with the two tasks we used, and to a lesser degree that object clefts were measured in only one of the generalisation tasks. As such, we do not know whether improvements seen in treatment would have generalised to other elicitation conditions for subject cleft sentences; nor do we know how patients would have performed on object cleft sentences in the Caplan and Hanna test.

Overall the participants who received treatment showed excellent acquisition of trained structures, generalisation of treatment gains to two different constrained sentence production tasks, of both trained and some untrained

structures, and changes after treatment on a narrative task. One control participant also showed some improvement on some of the same measures. The improvement of the control participant could be viewed as a loss of experimental control, invalidating the treatment approach and the results for the participants who received treatment. However, the strong and consistent results for the treated participants, from a population in whom it is notoriously difficult to document acquisition and generalisation gains (Fink et al., 1998; Mitchum et al., 2000; Thompson, 1989), suggest that the treatment results represent real effects. While the control participant's performance certainly qualifies the findings and conclusions of the study, and requires that this treatment approach be further investigated, we believe that there exist a number of possible explanations for the changes seen in the control participant's performance.

There have been a few patients (see Mitchum et al., 2000) in the literature in whom rapid and impressive gains have been documented after treatment. One of these is Byng's (1988) patient BRB, who was able to comprehend passive sentences and who showed structural gains in narrative production after only two weeks of (home) therapy. Another is patient EW (Schwartz et al., 1994), who showed almost errorless comprehension of noncanonical sentences following treatment. It is possible that MJ in our study was another such patient. She may have been able to benefit from minimal exposure to either mapping operations or to structural information (as provided in the PDSM, for example), to produce passive sentences. For instance, as recently demonstrated in another patient by Inglis (2003), she may have had strong analytical abilities that allowed her to abstract the by-phrase "rule" necessary for correctly producing a passive sentence. That she employed a superficial learned strategy to produce passive sentences is supported by the fact that her sentence comprehension for passive sentences did not improve significantly, even on a test to which she was exposed repeatedly (Picture Comprehension Test). This pattern of performance is very much like that reported by Haendiges et al. (1996) for patient EA. Another possibility is that her repeated exposure to the sentence stimuli may have contributed to enhanced event perception of the picture stimuli, resulting, in turn, in improved sentence production (Marshall et al., 1993, 1998). That this participant may have arrived at her correct sentence production in a manner different from the treated participants is suggested by her complete lack of improvement in the post-treatment results on object cleft sentences. Predictions for improved thematic role mapping following treatment were dependent on improvement on both passive and object cleft sentences as they share noncanonical thematic role assignment. All treated participants demonstrated improvement on both sentence structures; however the control participant improved only on passive sentences. In addition, significant improvement was noted in

this participant on the PDSM test on trained but not untrained sentence structures, and whereas in two of the three treated participants change on the Caplan and Hanna Test was significantly different, it was not for this participant. That improvement is noted on the PDSM test but not the Caplan and Hanna test is interesting. Whereas the PDSM, with its modelling technique, provides the patient with a template of the syntactic form to be produced, the Caplan and Hanna Test provides no cues as to which syntactic form to produce. In the absence of a modelled template, control MJ made only small gains, whereas two of the three treated participants showed significant gains.

Another possibility is that control MJ may have acquired the ability to achieve NP-movement but not *wh*-movement (e.g., Ballard & Thompson, 1999; Thompson et al., 1997). Thompson and colleagues (1997) have shown that training object cleft sentences did not generalise to passive sentences and training passive sentences did not generalise to object cleft sentences in two agrammatic patients. Object cleft sentences require *wh*-movement and passives require NP-movement (Chomsky, 1993). It must also be noted that although our treated participants demonstrated improved production of both passive and object cleft sentences after treatment, the dissociation between passives and object clefts for two participants at follow-up testing, as well as the relative lack of generalisation to object cleft sentences on the PDSM suggests that the linguistic properties of sentences may have played a role in the patterns of recovery and generalisation in our participants. This will require further investigation in future studies.

Analysis of post-treatment scores on one of the generalisation tests, the Caplan and Hanna Test, provides clues as to what may have benefited the participants in treatment. The finding that the thematic role score did not improve significantly in any of the treated participants may raise questions about whether the treatment successfully targeted thematic role assignment, and we will take up an alternative possibility below. However, these findings also relate to the issue of patient selection criteria for mapping therapy studies. Our selection criteria were comparable to those used in previous mapping therapy studies (e.g., Schwartz et al., 1994; Byng et al., 1994). However, our inclusion of the Caplan and Hanna Test in this study allowed us to analyse the underlying nature of patients' sentence production impairments with a scoring system that examined sentence constituents separately. Examination of the results from this test showed that patients who could by other criteria be characterised as having difficulties with thematic role mapping actually emerged with widely varying abilities as measured by their thematic role scores on this test, even before therapy. This is an important, if unexpected finding, and suggests that selection criteria for mapping therapy studies might benefit from revision in the future.

The most consistent finding after treatment was of improved verb morphology and by-phrase scores. This is interesting for a number of reasons.

In our mapping therapy, we provided the morphologically correct form of the verb, without otherwise focusing treatment, providing feedback, or scoring for correct morphology. Nevertheless, it would appear that this minimal information may have been sufficient to effect a change in verb morphology scores. One possibility is that the morphological information contained in the verb in particular, aided participants in assembling the necessary information at the functional level for accurate translation to the positional level, and, ultimately, for correct sentence production. Indeed, Weinrich and colleagues (Weinrich, Shelton, Cox, & McCall, 1997; Weinrich, Boser, & McCall, 1999) have demonstrated the benefit of training tense morphology on improved verb production. Their patient EA, whose sentence production and comprehension improved after training on passive sentences (Weinrich et al., 2001), had previously undergone training for tense morphology. Grammatical frames therapy (Mitchum & Berndt, 1994; Mitchum et al., 1993; Rochon & Reichman, 2003) also focuses on training morphological elements.

It can be difficult to disentangle measurement of correct thematic role assignment, and production grammatical elements and syntactic forms. It could be argued that these are different ways of measuring the same things (see Caplan & Hanna, 1998). They are at least partly separable however, as can be seen when patients produce simple active forms for passives (such as was seen in some of our participants even before treatment). Correct thematic role information can be provided without the correct grammatical elements and syntactic forms. The improved verb morphology and by-phrase scores lead us to wonder whether there was another, additional, mechanism at work that led to improvement in this treatment programme.

One possibility is that we may have inadvertently obtained an effect of syntactic priming (Bock, 1986; Bock & Loebell, 1990). Syntactic priming refers to the finding that speakers tend to repeat previously presented syntactic structures. This phenomenon has been found in normal participants (e.g., Bock, 1986; Bock & Loebell, 1990; Hartsuiker, Kolk, & Huiskamp, 1999) and in aphasic patients (Hartsuiker & Kolk, 1998; Saffran & Martin, 1997). It should be noted that aphasic patients make many errors in priming, including using the wrong preposition for "by", and making tense errors in the passive form (Saffran & Martin, 1997). That this structural priming is syntactic in nature comes from the findings of Bock and Loebell (1990) who found that sentences that share syntactic structure prime each other even if they differ in thematic role assignment (e.g., intransitive locatives and passives both prime passives). Hartsuiker and Kolk (1998) have claimed that syntactic priming works by overcoming a resource limitation in aphasics, at least temporarily. Chang, Dell, Bock, and Griffin (2000) have recently claimed that structural priming involves a form of implicit learning and that priming may derive from multiple sources or routes. In addition to syntactic structure (Bock & Loebell, 1990), priming has been found to be influenced by the verb

in the sentence, irrespective of tense, aspect or number (Pickering & Branigan, 1998), by word order (Hartsuiker et al., 1999), and by thematic roles (Hare & Goldberg, 1999, but see Chang et al., 2000).

While we cannot determine conclusively from our findings that priming occurred in our mapping therapy, the possibility exists that some combination of syntactic structure, information about verb morphology and thematic roles led to implicit learning which the impaired system was then able to implement in sentence production. If we are correct, our findings suggest, along with those of Saffran and Martin (1997), that syntactic priming in aphasic patients can be relatively long-lasting. This will require further investigation, but may prove to be a fruitful avenue for future treatment approaches.

Unlike other mapping therapy studies (e.g., Byng et al., 1994; Marshall et al., 1993; Schwartz et al., 1994) where production improved after comprehension training, or others (Weinrich et al., 2001, Patient EA), where comprehension improved after production training, we found no cross-modality gains in this study: Improvements in sentence production were not accompanied by gains in sentence comprehension. This may have been due to a number of factors. On the PCB Test, participants were found to have difficulties on some of the other noncanonical sentence structures included on this test (e.g., object relatives), which are known to be differentially difficult for patients (see Caplan et al., 1985; Schwartz et al., 1987), and upon which performance has been found to vary in treatment (Schwartz et al., 1994). On the other hand, such a lack of improvement cannot account for the lack of improvement seen on the Picture Comprehension Test, which included only the sentence structures trained in treatment. It is possible that having to respond "yes" or "no" to only one picture proved to be a cognitively challenging task for the participants. However, Cupples and Inglis (1993) found a one picture sentence comprehension test to be slightly easier for one of their participants. Our results are more in keeping with those of Mitchum and colleagues (Berndt & Mitchum, 1997; Haendiges et al., 1996; Mitchum et al., 1995) who showed improvements specific to the modality being treated, and of Jacobs and Thompson (2000), who found cross-modal generalisation for comprehension training, but not for production training.

The findings of this study raise the issue of whether previous gains in sentence production seen after treatment for comprehension were at least in part due to a production component of the comprehension treatment. In addition, our findings suggest the possibility that the mapping operations used in comprehension and production may be distinct and may require treatment targeting each modality separately (Mitchum et al., 2000). Alternatively, mapping operations may be multi-faceted and/or may not be completely bi-directional. The operations involved in successfully mapping from the positional to the functional level (i.e., in comprehension) may be sufficient to aid, in turn, in successful mapping from the functional to the positional level

(i.e., in production). However, as seen in Berndt and Mitchum (1997), Jacobs and Thompson (2000), and others, the reverse may not hold. Indeed, Mitchum and Berndt (2001) hold that lexically stated mapping rules (i.e., those that have stressed the importance of the verb in sentence comprehension/sentence production) "are relevant to both production and comprehension, whereas procedural operations exercise over modality-specific domains" (p. 567). It must be mentioned that some comprehension abilities were necessary in order for patients to understand the instructions and move through the steps of our production treatment programme. In that sense, our production programme required comprehension abilities, but these were not sufficient to improve participants' comprehension of semantically reversible sentences.

Our findings suggest that a mapping therapy approach can be successfully applied to targeting the sentence production impairment in nonfluent aphasia directly. The results are qualified by the performance of one control participant. However, the acquisition of trained sentence structures in treatment, and, especially, the generalisation of gains to other tasks shows that most participants who received treatment improved as predicted. Whether a mechanism such as syntactic priming may have influenced performance raises intriguing questions both theoretically and clinically, as does the relationship between the mapping operations underlying the production and comprehension of sentences.

REFERENCES

Ballard, K. J., & Thompson, C. K. (1999). Treatment and generalisation of complex sentence production in agrammatism. *Journal of Speech, Language and Hearing Research*, *42*, 690–707.

Berndt, R. S. (1987). Symptom co-occurrence and dissociation in the interpretation of agrammatism. In M. Coltheart, G. Sartori, & R. Job (Eds.), *The cognitive neuropsychology of language* (pp. 221–223). Hove, UK: Lawrence Erlbaum Associates Ltd.

Berndt, R. S. (1991). Sentence processing in aphasia. In M. T. Sarno (Ed.), *Acquired aphasia* (pp. 223–270). San Diego: Academic Press.

Berndt, R. S. (1998). Sentence processing in aphasia. In M. T. Sarno (Ed.), *Acquired aphasia* (pp. 229–267). San Diego: Academic Press.

Berndt, R. S., & Mitchum, C. C. (1997). An experimental treatment of sentence comprehension. In N. Helm-Estabrooks & A. Holland (Eds.), *Approaches to the treatment of aphasia* (pp. 91–111). San Diego, CA: Singular Publishing Group.

Berndt, R. S., Mitchum, C. C., & Haendiges, A. N. (1996). Comprehension of reversible sentences in "agrammatism": A meta-analysis. *Cognition*, *58*, 289–308.

Berndt, R., Wayland, S., Rochon, E., Saffran, E., & Schwartz, M. (2000). *Quantitative production analysis: A training manual for the analysis of aphasic sentence production.* Hove, UK: Psychology Press.

Beveridge, M. A., & Crerar, M. A. (2002). Remediation of asyntactic sentence comprehension using a multimedia microworld. *Brain and Language*, *82*, 243–295.

Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, *18*, 355–387.

Bock, J. K., & Levelt, W. J. M. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945–984). San Diego, CA: Academic Press.

Bock, J. K., & Loebell, H. (1990). Framing sentences. *Cognition*, *35*, 1–39.

Bradley, D. C., Garrett, M. F., & Zurif, E. B. (1980). Syntactic deficits in Broca's aphasia. In D. Caplan (Ed.), *Biological studies of mental processes* (pp. 269–286). Cambridge, MA: MIT Press.

Byng, S. (1988). Sentence processing deficits: Theory and therapy. *Cognitive Neuropsychology*, *5*, 629–676.

Byng, S., Nickels, L., & Black, M. (1994). Replicating therapy for mapping deficits in agrammatism: Remapping the deficit. *Aphasiology*, *8*, 315–341.

Caplan, D., Baker, C., & Dehaut, F. (1985). Syntactic determinants of sentence comprehension in aphasia. *Cognition*, *21*, 117–175.

Caplan, D., & Futter, C. (1986). Assignment of thematic roles to nouns in sentence comprehension by an agrammatic patient. *Brain and Language*, *27*, 117–134.

Caplan, D., & Hanna, J. E. (1998). Sentence production by aphasic patients in a constrained task. *Brain and Language*, *63*, 184–218.

Caplan, D., & Hildebrandt, N. (1988). *Disorders of syntactic comprehension.* Cambridge, MA: MIT Press.

Caramazza, A., & Berndt, R. S. (1985). A multicomponent view of agrammatic Broca's aphasia. In M. L. Kean (Ed.), *Agrammatism* (pp. 27–64). New York: Academic Press.

Caramazza, A., & Zurif, E. B. (1976). Dissociation of algorithmic and heuristic processes in language comprehension: Evidence from aphasia. *Brain and Language*, *3*, 575–582.

Chang, F., Dell, G. S., Bock, K., & Griffin, Z. M. (2000). Structural priming as implicit learning: A comparison of models of sentence production. *Journal of Psycholinguistic Research*, *29*, 217–229.

Chomsky, N. (1993). A minimalist program for linguistic theory. In K. Hale & S. J. Keyser (Eds.), *The view from building 20.* Cambridge, MA: MIT Press.

Cupples, L., & Inglis, A. L. (1993). When task demands induce "asyntactic" comprehension: A study of sentence comprehension in aphasia. *Cognitive Neuropsychology*, *10*, 201–234.

Fink, R. (2001). Mapping treatment: An approach to treating sentence level impairments in agrammatism. *Special Interest Division*, *2*(11), 83–87.

Fink, R. B., Schwartz, M. F., & Myers, J. L. (1998). Investigations of the sentence-query approach to mapping therapy. *Brain and Language*, *65*, 203–207.

Fink, R. B., Schwartz, M. F., Rochon, E., Myers, J. L., Socolof, G. S., & Bluestone, R. (1995). Syntax stimulation revisited: An analysis of generalisation of treatment effects. *American Journal of Speech-Language Pathology*, *4*, 99–104.

Fink, R. B., Schwartz, M. F., Rochon, E., Myers, J. L., Socolof, G. S., Bluestone, R., & Sobel, P. R. (1994). *Picture Description with Structure Modeling (PDSM): A procedure for measuring syntactic generalisation.* Poster presented at the Academy of Aphasia, Boston.

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

Friedman, N., & Grodzinsky, Y. (1997). Tense and agreement in agrammatic production: Pruning the syntactic tree. *Brain and Language*, *56*, 397–425.

Garrett, M. F. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.), *Language production. Vol. 1: Speech and talk* (pp. 177–220). New York: Academic Press.

Garrett, M. F. (1988). Processes in language production. In F. J. Newmeyer (Ed.), *Linguistics: The Cambridge Survey: III Language: Psychological and Biological Aspects* (pp. 69–96). Cambridge: Cambridge University Press.

Goodglass, H. (1976). Agrammatism. In H. Whitaker & H. A. Whitaker (Eds.), *Studies in neurolinguistics, Vol. 1* (pp. 237–260). New York: Academic Press.

Goodglass, H., & Kaplan, E. (1983). *The assessment of aphasia and related disorders* (2nd edition). Philadelphia: Lea & Febiger.

Grodzinsky, Y. (1986). Language deficits and syntactic theory. *Brain and Language*, *27*, 135–159.

Grodzinsky, Y. (2000). The neurology of syntax: Language use without Broca's area. *Behavioral and Brain Sciences*, *23*, 1–71.

Grodzinsky, Y., & Finkel, L. (1998). The neurology of empty categories: Aphasics' failure to detect ungrammaticality. *Journal of Cognitive Neuroscience*, *10*, 281–291.

Haendiges, A. N., Berndt, R. S., & Mitchum, C. C. (1996). Assessing the elements contributing to a "mapping" deficit: A targeted treatment study. *Brain and Language*, *52*, 276–302.

Hare, M. L., & Goldberg, A. E. (1999). Syntactic priming: Purely syntactic? *Proceedings of the Cognitive Science Society.*

Hartsuiker, R. J., & Kolk, H. H. J. (1998). Syntactic facilitation in agrammatic sentence production. *Brain and Language*, *62*, 221–254.

Hartsuiker, R. J., Kolk, H. H. J., & Huiskamp, P. (1999). Priming word order in sentence production. *Quarterly Journal of Experimental Psychology, 52A*, 129–147.

Inglis, A. L. (2003). Taking expectations to task in aphasic sentence comprehension: Investigations of off-line performance. *Aphasiology*, *17*, 265–289.

Jacobs, B. J., & Thompson, C. K. (2000). Cross-modal generalisation effects of training noncanonical sentence comprehension and production in agrammatic aphasia. *Journal of Speech, Language, and Hearing Research*, *43*, 5–20.

Jones, E. V. (1986). Building the foundations for sentence production in a non-fluent aphasic. *British Journal of Disorders of Communication*, *21*, 63–82.

Le Dorze, G., Jacob, A., & Coderre, L. (1991). Aphasia rehabilitation with a case of agrammatism: A partial replication. *Aphasiology*, *5*, 63–85.

Linebarger, M. C., Schwartz, M. F., & Saffran, E. M. (1983). Sensitivity to grammatical structure in so-called agrammatic aphasics. *Cognition*, *13*, 361–394.

Marshall, J., Pring, T., & Chiat, S. (1993). Sentence processing therapy: Working at the level of the event. *Aphasiology*, *7*, 177–199.

Marshall, J., Pring, T., & Chiat, S. (1998). Verb retrieval and sentence production in aphasia. *Brain and Language*, *63*, 159–183.

Miceli, G., & Mazzucchi, A. (1990). The nature of speech production deficits in so-called agrammatic aphasia: Evidence from two Italian patients. In L. Menn & L. K. Obler (Eds.), *Agrammatic aphasia: Cross-language narrative sourcebook*. Baltimore, MD: John Benjamins.

Miceli, G., Silveri, M. C., Romani, C., & Caramazza, A. (1989). Variation in the pattern of omissions and substitutions of grammatical morphemes in the spontaneous speech of so-called agrammatic patients. *Brain and Language*, *36*, 447–492.

Mitchum, C. C., & Berndt, R. S. (1994). Verb retrieval and sentence construction: Effects of targeted intervention. In G. Humphreys & J. Riddoch (Eds.), *Cognitive neuropsychology and cognitive rehabilitation* (pp. 317–348). Hove, UK: Lawrence Erlbaum Associates Ltd.

Mitchum, C. C., & Berndt, R. S. (2001). Cognitive neuropsychological approaches to diagnosing and treating language disorders: Production and comprehension of sentences. In R. Chapey (Ed.), *Language intervention strategies in aphasia and related neurogenic communication disorders* (4th edition, pp. 551–571). Philadelphia: Lippincott, Williams, & Wilkins.

Mitchum, C. C., Greenwald, M. L., & Berndt, R. S. (1997). Production-specific thematic mapping impairment: A treatment study. *Brain and Language*, *60*, 121–123.

Mitchum, C. C., Greenwald, M. L., & Berndt, R. S. (2000). Cognitive treatments of sentence processing disorders: What have we learned? *Neuropsychological Rehabilitation*, *10*, 311–336.

Mitchum, C. C., Haendiges, A. N., & Berndt, R. S. (1993). Model-guided treatment to improve written sentence production: A case study. *Aphasiology*, *7*, 71–109.

Mitchum, C. C., Haendiges, A. N., & Berndt, R. S. (1995). Treatment of thematic mapping in sentence comprehension: Implications for normal processing. *Cognitive Neuropsychology*, *12*, 503–547.

Miyake, A., Carpenter, P. A., & Just, M. A. (1994). A capacity approach to syntactic comprehension disorders: Making normal adults perform like aphasic individuals. *Cognitive Neuropsychology*, *11*, 671–717.

Nespoulous, J. L., Dordain, M., Perron, C., Ska, B., Bub, D., Caplan, D., Mehler, J., & Lecours, A. R. (1988). Agrammatism in sentence production without comprehension deficits: Reduced availability of syntactic structures and/or grammatical morphemes? A case study. *Brain and Language*, *33*, 273–295.

Nickels, L., Byng, S., & Black, M. (1991). Sentence processing deficits: A replication of therapy. *British Journal of Disorders of Communication*, *26*, 175–199.

Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language*, *39*, 633–651.

Rochon, E., Fink, R. B., Schwartz, M. F., Myers, J. L., Sobel, P. R., Bluestone, R., et al. (1994). Stability of narrative production across narrative elicitation tasks. *Brain and Language*, *47*, 420–423.

Rochon, E., & Reichman, S. (2003). A modular treatment for sentence processing impairments in aphasia: 1. Sentence production. *Journal of Speech-Language Pathology and Audiology*, *27*, 203–210.

Rochon, E., Saffran, E. M., Berndt, R. S., & Schwartz, M. F. (2000). Quantitative analysis of aphasic sentence production: Further development and new data. *Brain and Language*, *72*, 193–218.

Saffran, E. M., Berndt, R. S., & Schwartz, M. F. (1989). The quantitative analysis of agrammatic production: Procedure and data. *Brain and Language*, *37*, 440–479.

Saffran, E. M., & Martin, N. (1997). Effects of structural priming on sentence production in aphasics. *Language and Cognitive Processes*, *12*, 877–882.

Saffran, E. M., & Schwartz, M. F. (1988). Agrammatic comprehension it's not: Alternatives and implications, *Aphasiology*, *2*, 389–394.

Saffran, E. M., Schwartz, M. F., Linebarger, M., Martin, N., & Bochetto, P. (1988). *The Philadelphia Comprehension Battery for Aphasia.* Unpublished manuscript.

Saffran, E. M., Schwartz, M. F., & Marin, O. S. M. (1980a). Evidence for aphasia: Isolating the components of a production model. In B. Butterworth (Ed.), *Language production, Vol. 1* (pp. 221–241). New York: Academic Press.

Saffran, E. M., Schwartz, M. F., & Marin, O. S. M. (1980b). The word order problem in agrammatism: 2. Production. *Brain and Language*, *10*, 263–280.

Schwartz, M. F., Fink, R. B., Rochon, E., & Myers, J. L. (1992). *Mapping therapy sentence bank*, Unpublished stimuli set. Philadelphia.

Schwartz, M. F., Fink, R. B., & Saffran, E. M. (1995). The modular treatment of agrammatism. *Neuropsychological Rehabilitation*, *5*, 93–127.

Schwartz, M. F., Linebarger, M. C., & Saffran, E. M. (1985). The status of the syntactic deficit theory of agrammatism. In M. L. Kean (Ed.), *Agrammatism* (pp. 83–104). New York: Academic Press.

Schwartz, M. F., Linebarger, M. C., Saffran, E. M., & Pate, D. S. (1987). Syntactic transparency and sentence interpretation in aphasia. *Language and Cognitive Processes*, *2*, 85–113.

Schwartz, M. F., Saffran, E. M., Fink, R. B., Myers, J. L., & Martin, N. (1994). Mapping therapy: A treatment programme for agrammatism. *Aphasiology*, *8*, 19–54.

Siegel, S., & Castellan, N. J. (1988). Nonparametric statistics for the behavioral sciences. New York: McGraw-Hill.

Thompson, C. K. (1989). Generalisation research in aphasia: A review of the literature. In T. Prescott (Ed.). *Clinical Aphasiology, Vol. 18* (pp. 195–222). Boston: College-Hill.

Thompson, C. K., & Shapiro, L. P. (1994). A linguistic-specific approach to treatment of sentence production deficits in aphasia. In P. Lemme (Ed.), *Clinical Aphasiology, Vol. 21* (pp. 307–323). Boston: College-Hill.

Thompson, C. K., Shapiro, L. P., Ballard, K. J., Jacobs, B. J., Schneider, S. S., & Tait, M. E. (1997). Training and generalisation of wh- and NP-movement structures in agrammatic aphasia. *Journal of Speech, Language, and Hearing Research*, *40*, 228–244.

Thompson, C. K., Shapiro, L. P., & Roberts, M. M. (1993). Treatment of sentence production deficits in aphasia: A linguistic-specific approach to wh-interrogative training and generalisation. *Aphasiology*, *7*, 111–133.

Tissot, R. J., Mounin, G., & Lhermite, F. (1973). *L'Agrammatisme.* Brussels: Dessart.

Wambaugh, J. L., & Thompson, C. K. (1989). Training and generalisation of agrammatic aphasic adults: Wh-interrogative productions. *Journal of Speech and Hearing Disorders*, *54*, 509–525.

Weinrich, M., Boser, K. I., & McCall, D. (1999). Representation of linguistic rules in the brain: Evidence from training and aphasic patient to produce past tense verb morphology. *Brain and Language*, *70*, 144–158.

Weinrich, M., Boser, K. I., McCall, D., & Bishop, V. (2001). Training agrammatic subjects on passive sentences: Implications for syntactic deficit theories. *Brain and Language*, *76*, 45–61.

Weinrich, M., Shelton, J. R., Cox, D. M., & McCall, D. (1997). Remediating production of tense morphology improves verb morphology in aphasia. *Brain and Language*, *58*, 23–45.

APPENDIX
EXAMPLE OF THE TREATMENT PROTOCOL

A. Level 1 (Agent cue)

Example: A photograph of a nurse chasing a tall teacher is presented
Target sentence: "The nurse chases the tall teacher".
Examiner says:

- This is a picture about chasing.
- The verb in the sentence is "chases".
- In this picture the one being chased is the tall teacher (Theme).
- The one doing the chasing is the nurse (Agent)*
- Please make a sentence starting with the nurse.

Sentences treated: actives, subject clefts.

B. Level 2 (Theme cue)

Example: A photograph of a soldier hugging a farmer is presented.
Target sentence: "The farmer was hugged by the soldier".
Examiner says:

- This is a picture about hugging.
- The verb in the sentence is "hugged".
- In this picture the one being hugged is the farmer.
- The one doing the hugging is the soldier.*
- Please make a sentence starting with the farmer.

Sentences treated: passives, object clefts.

---

* Order of presentation/identification of agent and theme was balanced and randomised across all sentences in Levels 1 and 2.

C. Level 3 (Agent *and* Theme Cue)

Example: A photograph of a judge calling a baker is presented.
Target sentence: "The baker was called by the judge".
Examiner says:

- This is a picture about calling.
- The verb in the sentence is "called".
- In this picture the one doing the calling is the judge.
- The one being called is the baker.**
- Please make a sentence starting with the baker.

Sentences treated: actives, subject clefts, passives, object clefts.

D. Level 4 (Agent *or* Theme Cue)

Example: A photograph of an author shooting a farmer is presented.
Target sentence: "It is the author that is shooting the farmer".
Examiner says:

- This is a picture about shooting.
- The verb in the sentence is "shooting".
- In this picture the one doing the shooting is the author.
- The other person is the farmer.
- Please make a sentence starting with the author.
- Please start your sentence with "It is".***

Sentences treated: actives, subject clefts, passives, object clefts.

The following procedures were the same for all levels

Examiner:

- Placed icons above the photo as roles were identified to alert the participants to the thematic roles.
- Placed written role names in front of the participant.

If Correct:

- Examiner reinforced the response, e.g., for Level 1, "That's right. In the sentence: 'The nurse chases the tall teacher, the NURSE is the doer, the one doing the action'."

If Incorrect:

- Examiner provided correct model, e.g., "No. Actually in this sentence, 'the nurse' (icon is placed over agent) 'chases the tall teacher' (icon is placed over the theme), the NURSE is the doer, the one doing the action." So the sentence would be "The nurse chases the tall teacher".
- Icons were left on the photo as examiner repeated full sentence and pointed to each element in the picture.

** Order of presentation of agent and theme was fixed but random role order was used in request to begin sentence.
*** At all levels, when the target was a subject cleft or object cleft sentence, the examiner added: "Please begin the sentence starting with 'it is'."