APHASIOLOGY, 1998, VOL. 12, NO. 7/8, 587–600

Rule-based treatment for acquired phonological dyslexia

DIANE L. KENDALL†, MALCOLM R. McNEIL†
and STEVEN L. SMALL‡

† Department of Communication Science and Disorders, University of Pittsburgh, PA, USA
‡ Department of Neurology, School of Medicine, University of Maryland, USA

Abstract

In the context of a multiple-baseline design, this study demonstrated the positive effects of behavioural treatment using grapheme to phoneme correspondence rules to treat a patient with phonological dyslexia 17 years after stroke onset. Treatment used repeated exposure to real and nonsense word stimuli embodying the regularities of two grapheme to phoneme correspondence rules (GPCR) with hierarchical cueing and knowledge of results. Results revealed a pattern of performance that increased beyond baseline variability and coincided in time with the institution of treatment. Generalization of these treatment effects occurred to words requiring knowledge of other GPCR and to an independent processing based reading measure.

Introduction

Phonological dyslexia, a reading impairment involving the decreased ability to read aloud non-words when compared with real words, was first described by Beauvois and Derouesne (1979). Their patient demonstrated near intact ability to read real words (94%) with poor performance pronouncing non-words (10%). Impaired reading of non-words when compared with real words is also part of deep dyslexia but the presence of semantic errors distinguishes it from phonological dyslexia. Other characteristics of deep and phonological dyslexia are an impaired ability to read function words inordinate to that of content words, along with the ability to read concrete words more accurately than abstract words (Coltheart 1996).

The disparity between real word and non-word reading in phonological and deep dyslexia, contrasted with good non-word reading and poor irregularly spelled word (i.e. tomb) reading in surface dyslexia, has been taken as evidence to support a dual-route model of reading (Coltheart 1985, Newcombe and Marshall 1985). This model assumes the existence of qualitatively distinct processing routes for reading words (via the lexical route) and non-words (via the non-lexical route). It is suggested that non-words cannot be processed via the direct route because non-words are not represented in the lexicon. As a result, non-words are processed via the indirect, non-lexical route using a series of grapheme–phoneme operations.

Address correspondence to: Diane L. Kendall, Communication Science and Disorders, University of Pittsburgh, 4033 Forbes Tower, Pittsburgh, Pennsylvania 15260, USA.

The phonological assembly of non-words within the dual-route model is often said to require three distinct operations executed serially (Coltheart 1996). The first operation is graphemic parsing where a letter string is converted to a grapheme string, so that, for example, the letter string CHOOTH is converted to the grapheme string [CH] + [OO] + [TH]. The second stage, phoneme assignment, uses grapheme–phoneme correspondence rules which convert the grapheme string to a phoneme string. This procedure depends upon knowledge of the language-specific rules specifying the correspondences between orthographic segments and their phonological counterparts (Coltheart 1985). The final stage is phoneme blending, which converts a phoneme string into a single phonological form.

Treatment for impaired grapheme–phoneme correspondence has been applied to deep dyslexia (de Partz 1986, Matthews 1991, Mitchum and Berndt 1991). De Partz re-trained grapheme–phoneme conversion rules in three stages in a French-speaking person with acquired deep dyslexia. First, single letter graphemes were matched to a word (e.g. the letter 'C' was associated with his wife's name 'Carole' and the letter 'B' with 'baby'). The second step involved association of letters with the first phonemes of the word codes. This was accomplished by confronting the patient with the letter and asking for the word. For example, if confronted with 'A', the patient would say 'Aaaello' and then was cued to produce orally only the first phoneme /a/. In the final stage of the treatment, the contextual criteria for the troublesome rule-based pronunciations were established. The patient was asked to pronounce separately each of the phonemes corresponding to the different letters of the non-words, then to try to combine them into a simple oral production. An example of one of the rules trained was 'the conversion rule of the letter S, which is pronounced most frequently /s/ but which changes to /z/ in an intervocalic position' (de Partz 1986, p. 162). In this stage, the three rules were explained to the patient explicitly and he was trained in their correct use in reading aloud. Parallel to the reading therapy, writing training was also undertaken. The writing training used the same relay strategy as the reading therapy (word code to transcode each phoneme into the corresponding grapheme). In the post-training assessment of reading, the patient read irregular words and low-imagery words correctly.

Matthews (1991) provided treatment to a deep dyslexic patient, motivated by Luria's notion of functional reorganization in which the general principle is to substitute indirect, conscious activity for a function that had formerly been unconscious or automatic, but now is impaired. In this study, the patient was impaired in the use of grapheme–phoneme rules. The goal of treatment was to re-establish these rules through the kinesthetic system, then to train the patient to 'sound out' those words that he could not otherwise read. Stimulus cards were designed to illustrate the phoneme–grapheme correspondences which included a picture of the articulatory posture for bilabial, alveolar, and velar stop consonants with its grapheme. Results of this study showed improved ability to read real words (from 5/15 to 13/15) and pseudohomophones (from 1/15 to 10/15).

Mitchum and Berndt (1991) also treated a patient with a deficit limited to the grapheme–phoneme conversion system. The basis for their study was the intervention described by de Partz (1986). The primary goal of remediation was to establish a mechanism by which the patient could convert the individual graphemes of unfamiliar words into meaningful phonemic strings. Treatment consisted of two sequential steps: training in auditory analysis and then explicitly teaching grapheme–phoneme correspondence rules. The tasks involved in teaching rules consisted of sounding out specific graphemes then blending the graphemes into a particular word. The intervention targeting grapheme to sound correspondence showed increased speed and accuracy. However, the treatment for individual phoneme blending did not generalize to untrained stimuli.

In the present study the following research questions were addressed: Does rule-based treatment improve the use of grapheme to phoneme correspondences in a single patient with phonological dyslexia? Do any positive effects of treatment generalize to untreated behaviours (i.e. a different rule and to untreated words and non-words) or to a non-standardized reading version of a standardized aphasia test, a reading vocabulary test, or homophones? Are any positive effects of treatment and generalization maintained following treatment termination?

Method

Subject

W.T. was a 42-year-old, right-handed female who sustained a left hemisphere ischaemic stroke after being struck by lightning at the age of 25. The patient initially had right hemiplegia and global aphasia. She received intensive, though unspecified speech–language, occupational, and physical therapies for approximately 5 years. Her pre-morbid language skills are assumed to have been at a reasonably high level based on her education (16+ years) and occupation as a nurse anaesthetist. An MRI scan at the age of 39 showed an infarct in the left middle cerebral artery distribution with involvement of the left fronto-temporal cortex and left basal ganglia. Extensive speech, language, cognitive and reading evaluations were conducted at the time of this study. A summary of the most relevant of these findings is presented in table 1.

Notable among the language and communication measures administered, the *Revised Token Test* (RTT) (McNeil and Prescott 1978) showed a mild auditory comprehension deficit with a score of 14·20, placing her in the 92nd percentile for left-hemisphere damaged aphasic subjects. Performance on the *Discourse Production Test* (DPT) (Nicholas and Brookshire 1993) showed a mild production deficit with performance characterized by slowed rate and a decreased number of content information units (70% content information units, mean = 86% for normal, non-aphasic subjects). The *Apraxia Battery for Adults* (Dabul 1986) revealed vowel errors, visible searching, errors that increased as the number of sounds in the sequence increased, and fewer errors in automatic than volitional speech. Speech articulation characteristics showed intersyllabic intrusive schwa, abnormal prosodic features, and awareness of errors as evidenced by attempts to self-correct. These errors were also evident during spontaneous speech. Speech production during oral reading included errors in addition to those errors evidenced during spontaneous speech. When reading aloud, W.T. omitted entire words and added/deleted suffixes which she did not do during spontaneous speech.

The *Gates–MacGinitie Reading Test* (GMG) (Gates and MacGinitie 1978) is a test of adult reading comprehension where the subject is required silently to read multiple paragraphs and vocabulary words and then respond to multiple choice questions. W.T.'s performance on the GMG showed impaired reading comprehension with a grade equivalence of 12·8 for vocabulary words and a grade equivalence of 9·2 for paragraph comprehension. Further testing used non-

Table 1. Assessment measures for W.T.

Formal measure	Results
Western Aphasia Battery (Kertesz 1982)	
Aphasia Quotient	95·3
Cortical Quotient	88·12F
Revised Token Test (Overall) (McNeil and Prescott 1978)	92 %ile (aphasic)
Pyramids and Palm Trees Test (Correct) (Howard and Patterson 1992)	101/104
Test of Adolescent/Adult Word Finding (German 1990)	92/107
Boston Naming Test (Kaplan *et al.* 1983)	59/60
Discourse Production Test (% CIU) (Nicholas and Brookshire 1993)	70 (84 = normal)
Apraxia Battery for Adults (Dabul 1986)	Mild–moderate apraxia of speech
Gates–MacGinitie Reading Test (Grade Equivalence) (Gates and MacGinitie 1978)	
Vocabulary words	12·8
Paragraph comprehension	9·2
Non-standardized reading measures	
Oral reading of non-words (% accuracy)	40
Repetition of non-words (% accuracy)	87
Identification of real-word rhymes	10/10
Oral reading real word homophones	
One letter different from true spelling	13/15
More than one letter different from true spelling	31/45
Wholistic method of word recognition	13/15
Verbal letter production	28/28
Reading Comprehension Battery for Aphasia (LaPointe and Horner 1979)	99/100
Wechsler Adult Intelligence Scale—Revised (Wechsler 1987)	
Verbal memory	105
Visual memory	117
General memory	111
Attention/concentration	74
Delayed recall	128

standardized tasks borrowed in construct, and some specific stimulus items from Saffran and Marin (1977). Paragraph reading aloud revealed omission of whole words, morphologic errors (adding and deleting suffixes) and errors producing low frequency words. Oral reading of non-words (i.e. birk) showed 40 % accuracy and oral repetition of non-words 87 % accuracy. Identification of real-word rhymes (forced choice, field of four) was 10/10, oral reading of real-word homophones with one letter different from true spelling (e.g. kup for cup) was 13/15, and oral reading of real-word homophones with more than one letter different from true spelling (e.g. minnit for minute) was 31/45. To investigate the possibility that words were not recognized as constellations of visual features, typographical perturbations such as mixing upper and lower case letters, vertical presentation, and vertical displacement of letters or placing ' + ' marks between letters of words were evaluated. W.T. was able correctly to identify 13/15 of these stimuli. To investigate W.T.'s knowledge of permissible letter strings, an evaluation of English orthography (e.g. EVAR vs. DVCE) was 10/10 accurate. Finally, verbal letter production was 28/28 accurate. Based on this evaluation, W.T. was judged to have met the criteria for the diagnosis of sublexical/phonological reading impairment.

Experimental design

A multiple-baseline across behaviours, with generalization probes, single-subject design was used in this study to evaluate the effects of an intervention programme. In this treatment (described below) stimuli appropriate for elicitation of the 'c-rule' were baselined for six sessions before the initiation of six treatment sessions. When all stimuli reached the criterion of 80% accuracy for three consecutive sessions for verbal and written responses, treatment was discontinued on the 'c-rule'. The maintenance of these effects was probed over a period of 11 consecutive and subsequent sessions. After three of these sessions, stimuli appropriate for the elicitation of the 'g-rule' were targeted for treatment. This three-session interval occurred secondary to scheduling conflicts where treatment could not be delivered although all treatment stimuli were probed. The 'g-rule' was baselined prior to treatment for 16 sessions during the treatment and maintenance of the 'c-rule'. The maintenance effects of treatment to the 'g-rule' were probed over three sessions following treatment termination. In addition, non-standardized homophones, the Gates–MacGinitie vocabulary subtest, and a reading version of the *Revised Token Test* were measured across the duration of the experiment and used as generalization probes.

Stimuli

The treatment stimuli were composed of real and nonsense words appropriate for eliciting the 'c-rule' and the 'g-rule' (rules described below), varied by 'simple' and 'difficult' levels of complexity and divided into four lists per rule (see Appendix 1). List 1 contained 10 words of one-syllable nouns of high concreteness, frequency, imagery, and meaningfulness (Paivio *et al.* 1968). List 2 contained 21-syllable nonsense words beginning with phonemes representing each rule. List 3 contained 10, 2–3 syllable nouns of low concreteness and low imagery (Paivio *et. al.* 1968) and list 4 contained 20 2–3 syllable non-words beginning with phonemes representing the rule. In summary, each rule contained a total of 60 words that were trained (10 simple real words, 20 simple non-words, 10 difficult real words, and 20 difficult non-words). All 60 words within each rule were randomly presented during treatment and were 'unrandomized' and presented in their respective lists when probed.

Treatment

A treatment paradigm focusing on improving the usage of grapheme to phoneme correspondence rules via systematic exposure to exemplars of each rule with instruction on pronunciation was administered for six treatment sessions for the first rule and five treatment sessions for the second rule over a 6 week period. The rules trained were two out of seven rules for grapheme–phoneme correspondence that occur consistently enough or cover enough words to warrant teaching children with impaired reading skills (May and Elliot 1973). The two rules trained were the 'c-rule' and the 'g-rule'. The 'c-rule' states that when *c* comes just before *a*, *o*, *u*, it is produced as /k/, otherwise as an /s/. The 'g-rule' states that when *g* comes at the end of words or just before *a*, *o*, or *u*, it is produced /g/, otherwise as /dz/.

The rules were not 'taught' explicitly; instead, the patient was exposed to multiple exemplars of non-words embodying the rule, with incorrect responses cued until correct using a cueing hierarchy. When an error occurred, W. T. was first given a phonetic cue (e.g. /s/ for the stimulus 'cylecaber'). If the response was still in error, she was told 'it begins like the word cycle'. If W. T. erred following the real word cue, she repeated the correctly produced non-word. The cueing hierarchy paired specific phonemes of non-words with phonemes in real words (dictated by the orthographic context) as in the preceding example.

It was believed that if the rule was trained explicitly, the subject would necessarily develop a strategy that would then be employed when decoding unfamiliar or novel words where the rule would apply. By definition, the strategy would use limited available processing resources and could further limit reading effectiveness. By contrast, it was believed that repetitive exposure to stimuli in which the rule was employed would promote an internalization of the rule and a more automatic, less resource-demanding reading process. The choice of non-words as treatment targets was motivated by a desire to avoid the possibility of producing the word via the semantic route while maintaining the necessity of applying the rule to the construction of orthographically and phonetically permissible sequences.

Initially, baselines were established for the vocabulary words subtest of the GMG, a reading version of the RTT (Odell 1983), a 30-real-word homophone list (Appendix 2) and eight word lists described above (Appendix 1). Each of the 11 treatment sessions lasted approximately 2 hours with the first hour dedicated to the collection of baseline and probe data and the second hour dedicated to the delivery of treatment. All real and non-words for the 'c-rule' were randomized and treated until criterion level was achieved (80% accuracy for both written and verbal responses for three consecutive sessions). This was followed by treatment of all randomized real and non-words for the 'g-rule'. Verbal responses were scored as either correct or incorrect prior to, and independent of, the application of the cueing hierarchy. There were verbal responses to 180 stimuli per session (60 words × 3).

During each treatment session, each word was presented once, visually, with printed black lettering on a white 3 × 5 inch (80 × 125 mm) card and W. T. was instructed to provide first a verbal ('say this word') and then, after the card was removed, a written ('write this word') response. The written response was required to reinforce further proper graphemic segmentation and correct spelling. As stated above, when an incorrect verbal response occurred, the cueing hierarchy was used to evoke the target response. Response time was measured for each verbal response and time was calculated from the onset of visual presentation of the word to onset of verbal production.

Data analysis

Three individuals, who were knowledgeable about and have published data from single-subject experimental design studies, served as judges for this study. Judges were asked to evaluate, by visual inspection, all treated and maintenance stimuli and probe measures to determine whether treatment, generalization, or maintenance effects had or had not occurred for both accuracy and for response time. Judges were asked to make their binary judgements at a 95% confidence level considering ranges of baseline scores as well as magnitude and slope of changes relative to the timing of conditions. In order to conclude that a treatment or no-treatment effect had occurred, all three judges were required to be in agreement.

Reliability

To determine inter-rater reliability, one judge (uninvolved in this study) evaluated W. T.'s previously recorded verbal productions as either correct or incorrect on 41 % of the total corpus of treatment stimuli with 94 % accuracy. In addition, intra-rater reliability was also performed by the first author with 99 % accuracy.

Results

Orientation

The multiple-baseline across behaviours generalization probes design used in this study is illustrated in figures 1 and 2 respectively. Figure 1 illustrates the subject's performance on the stimuli used for the treatment of the 'c-rule' and the 'g-rule' separated by stimulus type into eight 'lists' or categories of stimuli. The top four panels display the baseline, treatment, and maintenance data for the 'c-rule', those for the 'g-rule' are represented in the bottom four panels. The first full vertical line indicates initiation of treatment and the second full vertical line indicates treatment termination. The baseline, treatment, and maintenance phases are labelled across the top of the figures. Each number at the bottom of the figures represents one data collection session spanning 6 weeks, seven baseline and six treatment sessions for the 'c-rule', three probe and five treatment sessions for the 'g-rule', and three maintenance sessions. For figure 1, the vertical axis on the left represents the percentage accuracy for the verbal response (filled square) and the right vertical axis shows response time in seconds (filled triangle). The 'c-rule' was treated from sessions 8 to 13 and the 'g-rule' was treated from sessions 17 to 21. Following the baseline phase, all words were continuously probed in order to reveal a general performance change (in which case performance change could *not* be attributed to the behavioural treatment) or a systematic, temporally coherent treatment/generalization of treatment to untreated stimuli (in which case the effect could be attributed to the intervention and not to any number of other possible explanations for improved performance).

Figure 2 summarizes average scores of the three reading measures that were periodically probed throughout the course of the treatment regimen. Panel 1 illustrates verbal and written responses to a non-standardized homophone word list, panel 2 plots the vocabulary subtest portion of the GMG and panel 3 plots the overall scores of the reading version of the *Revised Token Test*.

Results

A pattern of performance was observed which increased beyond baseline levels and variability in magnitude and/or slope, subsequent to the institution of treatment for the 'c-rule' on all four stimulus types (panels 1–4). Performance on the untreated 'g-rule' during the treatment of the 'c-rule' demonstrated generalization

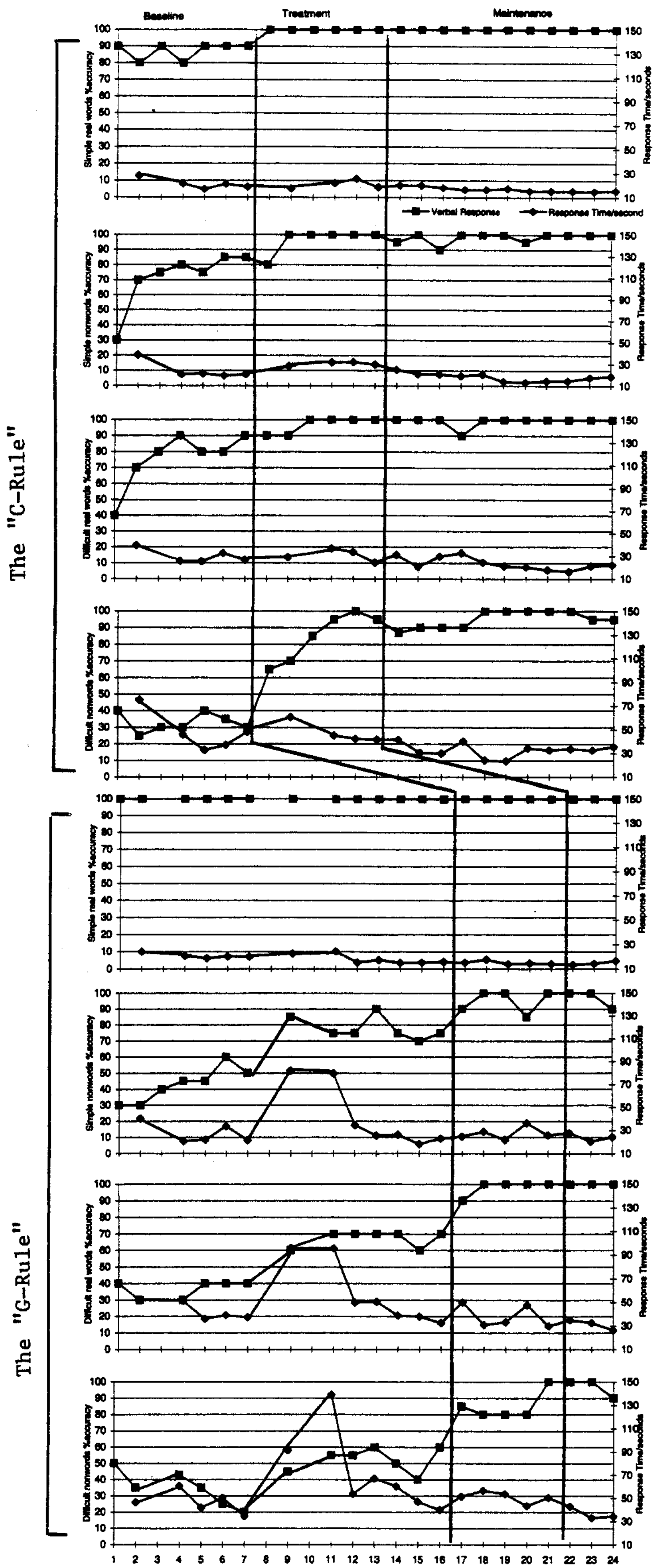

Figure 1. For legend see opposite.

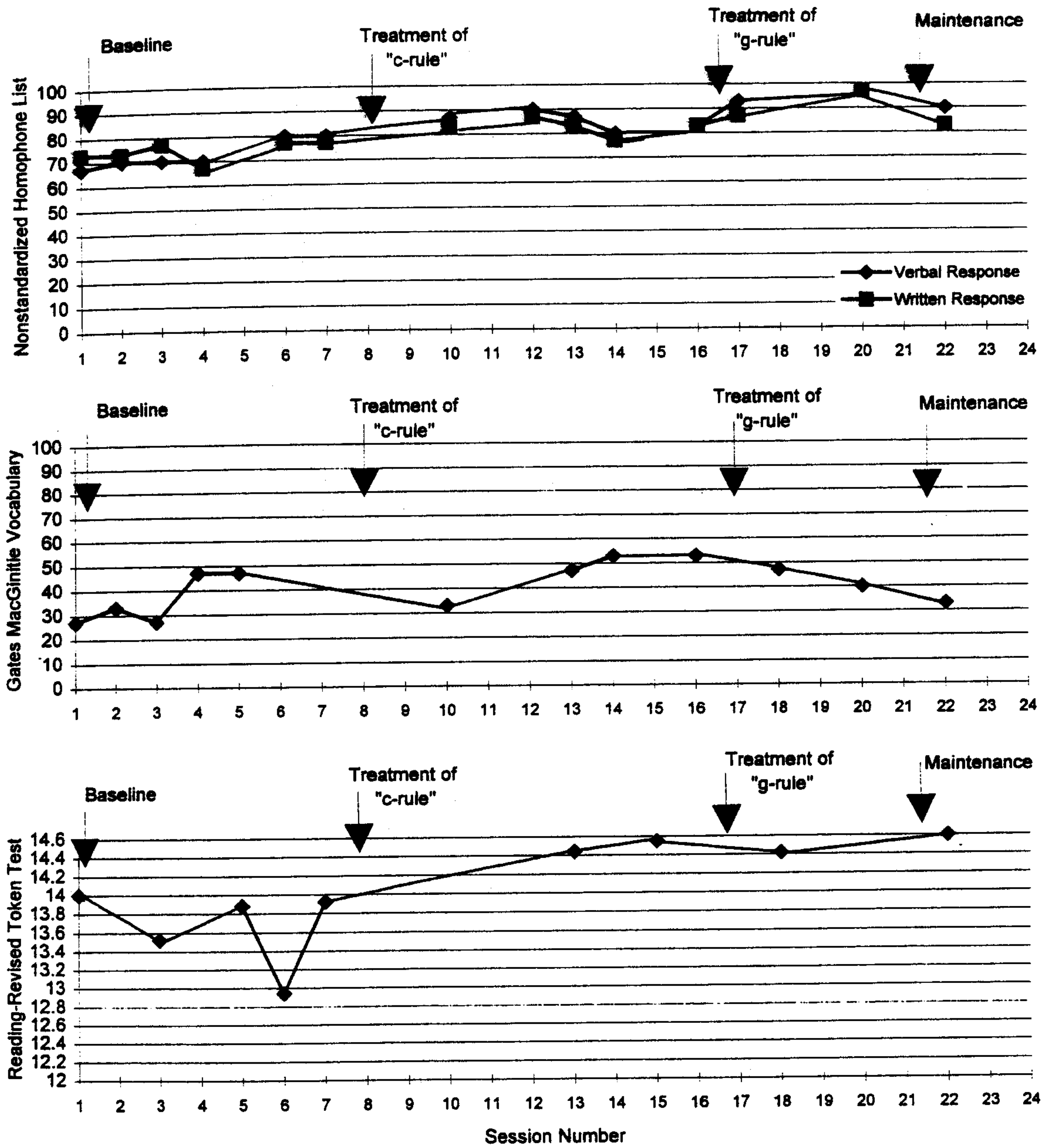

Figure 2. Accuracy for the generalization probes.

to all stimulus types (panels 6–8) except those represented in panel 5, whose levels were at ceiling. Maintenance of acquisition effects for the 'c-rule' was apparent for all stimulus types. No treatment and, hence, no maintenance effects were observed for response times.

Treatment of the 'g-rule' was judged to exhibit an immediate acquisition effect on simple non-words (panel 6), difficult real words (panel 7), and difficult non-words (panel 8). Maintenance of acquisition effects for the 'g-rule' was judged to be present for all lists in which an acquisition effect was present. The 'g-rule' simple real words (panel 5) were judged to show no acquisition effects because baseline performance levels were at ceiling prior to treatment.

Figure 1. Baseline, treated, and untreated probe and maintenance data for accuracy and response time for the 'c-rule' and the 'g-rule'.

Generalization to the non-standardized homophone list was judged to exceed baseline levels with the initiation of treatment of the 'c-rule' and again with treatment of the 'g-rule' for both verbal and written responses, but no maintenance effects were observed. No effects were observed on the vocabulary subtest portion of the GMG with scores remaining within the range of baseline during the targeted treatment of both rules and during the maintenance phases of each. An increase in performance exceeding baseline levels during treatment of the 'c-rule' and maintenance was observed on the overall scores of the reading version of the RTT.

Discussion

The purpose of this study was to investigate the effects of a treatment designed to improve the use of grapheme to phoneme correspondence rules in a person with aphasia and phonological dyslexia. Results of this study revealed a pattern of performance that increased to levels beyond baseline variability in magnitude and/or slope. These increases coincided in time with the institution of treatment for both rules. These treatment effects generalized to untreated word lists and to untreated reading measures.

The first research question asked whether treatment using exemplars derived from specific GPC rules improved the use of grapheme to phoneme correspondence in reading (the acquisition effect). This question can be answered positively. The presence of a treatment effect was defined and judged to be dependent on the presence of generalization to untreated stimuli (research question number 2). This question can also be answered positively. The third question, if treatment and generalization effects were maintained following treatment termination, can also be answered positively. The judges agreed that the treated and generalization behaviours were maintained for both rules and for the reading version of the RTT.

Generalization was present for the 'g-rule' (accuracy) during treatment of the 'c-rule' and for two of the three reading measures that were periodically probed throughout the course of the regimen (the non-standardized homophones list and the reading version of the RTT). A goal of the treatment was to avoid the training of strategies in order to avoid processing demands that could impede efficient reading. The lack of changes in response times during the acquisition and generalization phases of the treatment do not provide support for its achievement. In addition, examination of recorded verbal responses during treatment reveal that W. T. in fact was utilizing a strategy (i.e. 'let me think of a real word that begins with *gi*, I can use *giant*, now maybe this word will be *girandole*'). Generalization to the 'g-rule' might best be explained by the fact that W. T. was applying a strategy learned from the 'c-rule'. Generalization to the non-standardized homophone list also suggests that W. T. was able to apply the new strategy ('the real word ____ starts like this word') to the initial consonants of the homophones and then produce the word correctly. W. T. was also observed to use an index card to segment visually the homophones and apply the new strategy to each sound segment.

One result not consistent with the strategy interpretations is that of the reading RTT. The written commands in the reading version of the RTT did not contain the treated rules, hence those strategies believed to have been learned could not have been employed. Generalization to this measure could indicate that W.T. became more efficient in processing written material, but there is no independent evidence to support this speculation.

Generalization was not judged to have occurred to the Gates–MacGinitie vocabulary subtest which consisted of 20 vocabulary words, each with a forced choice of five real words. The patient was required to read the target word silently and circle the correct meaning. Improvement was not anticipated on this measure, as treatment was not targeted to improve vocabulary knowledge. This was, however, baselined and probed throughout the course of treatment in order to provide a stable measure against which targeted behaviour change could be judged, within the context of a single-subject multiple-baseline design.

This study was based on the hypothesis that W.T.'s decreased ability to read non-words and unfamiliar words was due to an inefficient use of the sublexical route. Treatment targeted the underlying deficit in the form of impaired grapheme to phoneme rule usage and used a cueing hierarchy for incorrect responses in the form of a real word that begins with the same letter in which the particular rule was to be applied. It was the premise of this treatment that it was reasonable to treat the underlying deficit using these rules as opposed to focusing treatment on the end-product such as oral reading at a paragraph level. The limited results that are available from previous research suggest that when treatment is aimed exclusively at the level of production, without regard to the integrity of processes that might precede phonetic and articulatory implementation, there is a paucity of generalization effects to other contexts (Mitchum *et al.* 1993). If the underlying deficit was targeted and treated successfully, then performance should generalize to other untreated stimuli and tasks. In this study, generalization was evident to homophone reading, the reading version of the RTT as well as to the untreated 'g-rule'. If the underlying deficit was not treated, generalization would not have occurred.

What remains to be delineated is whether the underlying deficit is regarded as the grapheme to phoneme conversion rule *per se*, or whether the underlying deficit can be regarded as the inefficient implementation of the rule with the locus of the deficit better conceived as a processing resource allocation deficit. The variability in performance on treatment and generalization tasks supports the conclusion that W.T. had the rules in her repertoire but was unable to implement them on a consistent and efficient basis. Other than support by default, evidence is not available to argue persuasively for the processing resources account.

Further research in this area should include replication of these findings as well as treatment to other stages in the sublexical route such as graphemic parsing followed by treatment of grapheme to phoneme correspondence rules. It was discovered that this patient did demonstrate difficulty with syllabic parsing and may have benefited from treatment to improve parsing skills prior to institution of grapheme to phoneme correspondence rules. While this treatment was, to a substantive degree, successful in this single patient with acquired phonological dyslexia, further treatment studies with other subjects and ultimately within a clinical trial should be conducted.

Acknowledgements

The support of NIH under award R01-DC3378 to the final author (S.L.S.) is gratefully appreciated.

References

Beauvois, M. F. and Derouesne, J. 1979, Phonological alexia: three dissociations. *Journal of Neurology, Neurosurgery and Psychiatry*, **42**, 1115–1124.

Coltheart, M. 1985, Cognitive neuropsychology and the study of reading. In M. I. Posner and O. S. M. Marin (Eds) *Attention and Performance XI* (Hillsdale, NJ: Erlbaum).

Coltheart, M. 1996, Phonological dyslexia: past and future issues. *Cognitive Neuropsychology*, **13**, 749–762.

Dabul, B. 1986, *The Apraxia Battery for Adults* (Austin, TX: Pro-Ed).

de Partz, M. P. 1986, Re-education of a deep dyslexia patient: rationale of the methods and results. *Cognitive Neuropsychology*, **3**, 149–177.

Gates, A. and MacGinitie, W. 1978, *Gates–MacGinitie Reading Tests* (New York: Teachers College Press).

German, D. J. 1990, *Test of Adolescent/Adult Word Finding* (Allen, TX: D&M).

Howard, D. and Patterson, K. 1992, The Pyramids and Palm Trees Test (Suffolk, UK: Thames Valley Test Company).

Kaplan, E., Goodglass, H. and Weintraub, S. 1983, *The Boston Naming Test* (Philadeplphia: Lea & Febiger).

Kertesz, A. 1982, *The Western Aphasia Battery* (New York: Grune & Stratton).

LaPointe, L. L. and Horner, J. 1979, *Reading Comprehension Battery for Aphasia* (Tigard, OR: C.C. Publications).

Matthews, C. 1991, Serial processing and the 'phonetic route:' lessons learned in the functional reorganization of deep dyslexia. *Journal of Communication Disorders*, **24**, 21–39.

May, F. and Elliot, S. 1973, *To Help Children Read: Mastery Performance Modules for Teachers in Training* (Columbus: Charles C. Merrill).

McNeil, M. and Prescott, T. 1978, *The Revised Token Test* (Austin, TX: Pro-Ed).

Mitchum, C. C. and Berndt, R. S. 1991, Diagnosis and treatment of phonological assembly in acquired dyslexia. *Journal of Neurolinguistics*, **6**, 103–137.

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

Newcombe, F. and Marshall, J. C. 1985, Reading and writing by letter sounds. In K. E. Patterson, J. C. Marshall and M. Coltheart (Eds) *Surface Dyslexia: Neuropsychological and Cognitive Studies of Phonological Reading* (Hillsdale, NJ: Erlbaum), pp. 35–49.

Nicholas and Brookshire 1993, *Discourse Production Test* (Tucson: Communication Skill Builders).

Odell, K. H. 1983, Comparison between auditory and reading comprehension in aphasic adults. Unpublished master's thesis, University of Wisconsin.

Paivio, A., Yuille, J. and Madigan, S. T. 1968, Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology Monograph Supplement*, **76**, 1–25.

Saffran, E. and Marin, O. 1977, Reading without phonology: evidence from aphasia. *Quarterly Journal of Experimental Psychology*, **29**, 515–525.

Wechsler, D. 1987, *The Wechsler Memory Scale—Revised* (New York: The Psychological Corporation, Harcourt Brace Jovanovich,).

Appendix 1: Words used in treatment

The 'c-rule'

Simple real words	Simple non-words	Difficult real words	Difficult non-words
1. coach	1. calip	1. conciliator	1. carelam
2. coal	2. cuthy	2. conservationist	2. copected
3. cage	3. carld	3. corduroy	3. cipivity
4. cane	4. corpt	4. covenant	4. cobussy
5. carrot	5. cosp	5. corpulence	5. catamercy
6. ceiling	6. cery	6. cellophane	6. cendich
7. cymbal	7. cylim	7. centimetre	7. cuspectiment
8. centre	8. ceppi	8. certificate	8. covabaric
9. cell	9. cydem	9. ceaselessly	9. cellissify
10. cement	10. cerd	10. certitude	10. carrovercy
	11. corz		11. cavernate
	12. cawp		12. copencydum
	13. cuc		13. cuzzily
	14. cofa		14. caddilate
	15. cuvem		15. cyburcy
	16. cyra		16. ceficorb
	17. ceva		17. cidumacor
	18. cew		18. cebudrate
	19. cybrem		19. cyptezzi
	20. cek		20. cirfidek

The 'g-rule'

Simple real words	Simple non-words	Difficult real words	Difficult non-words
1. garden	1. gubi	1. galapagos	1. gabrellic
2. geese	2. gaws	2. gargantuan	2. gompostic
3. guard	3. gommer	3. goniometer	3. guprosect
4. garlic	4. gude	4. gubernatorial	4. gottify
5. guess	5. gahlf	5. gurgitation	5. gagalla
6. gem	6. giam	6. gelatification	6. gersify
7. gel	7. gerrum	7. genuflection	7. geffish
8. giant	8. gyder	8. genteelism	8. gimbered
9. george	9. giaf	9. girandole	9. givitilaceous
10. ginger	10. guse	10. gimbals	10. gerfiletti
	11. gort		11. golorgoric
	12. gurd		12. gaprogacy
	13. gu		13. gulligistic
	14. gop		14. gattifritude
	15. gald		15. gaffage
	16. gep		16. gechrose
	17. gif		17. gelligest
	18. gera		18. gipity
	19. gempy		19. gibbage
	20. gidd		20. gembeted

Appendix 2

Real word homophones that were periodically probed during baseline, treatment of the 'c-rule', treatment of the 'g-rule', and maintenance phases

1. phyl	11. phrend	21. kuvver
2. minnit	12. tayp	22. flud
3. pensul	13. weale	23. rayne
4. grayp	14. phell	24. shayk
5. bote	15. joon	25. phlore
6. hoam	16. bocks	26. rokit
7. phurst	17. pheeld	27. karpit
8. jayle	18. skarph	28. throte
9. kote	19. nekliss	29. shelph
10. nyfe	20. boks	30. kween