

Effects of training multiple form classes on acquisition, generalization and maintenance of word retrieval in a single subject

MALCOLM R. MCNEIL†‡, PATRICK J. DOYLE†‡,
KRISTIE SPENCER‡, AMY JACKSON GODA‡,
DIANE FLORES†‡ and STEVEN L. SMALL†§

† Department of Communication Science & Disorders, University of Pittsburgh, PA, USA

‡ Aphasia Rehabilitation Research Laboratory and Clinic, VAMC, Pittsburgh, PA, USA

§ Department of Neurology, University of Maryland, MD, USA

Abstract

A study by McNeil *et al.* reported no generalization effects in two individuals with aphasia following application of a word finding treatment in which subjects were trained sequentially on lexical items arranged by form class. The present investigation examined, in one of the two subjects from the original study, whether training on lexical items from a variety of form classes concurrently would result in greater response generalization than was observed previously in this subject. Results replicated earlier findings with positive acquisition and maintenance effects and little evidence of generalization to untrained items within or across form classes.

Introduction

Most treatment approaches for dysnomia posit a failure at the lexical-semantic, phonological, or at both linguistic levels (Lesser 1989). Treatment tasks have frequently used picture naming (e.g. Hillis 1989, Thompson *et al.* 1986), word reading (e.g. Seron *et al.* 1979), or elicited, but internally generated words (e.g. McNeil *et al.* 1997). Treatment targets are typically selected on the basis of form class (Thompson and Kearns 1981), semantic features (Boyle and Coelho 1995, Thompson and Kearns 1981), or phonological factors (Thompson *et al.* 1991) such as phonological relatedness (Hillis 1989). Most stimuli are also controlled for such psycholinguistic variables as age of acquisition, operativity, frequency, familiarity, imageability, concreteness, phoneme length, and visual complexity of the stimulus pictures (Nickels and Howard 1995). Techniques used for facilitating the acquisition, generalization, and maintenance of lexical targets generally take the form of either multiple stimulus repetitions (Brookshire 1975), pre-stimulus cues (Howard *et al.* 1985, Lowell *et al.* 1995, Patterson *et al.* 1983) or post-stimulus cued

subject errors (e.g. Linebaugh 1997). Among all of these important variables, perhaps the most critical and illusory aspect of successful treatment for lexical retrieval deficits, and perhaps the most critical and illusory aspects of treatment for all speech and language disorders, is that of treatment generalization.

While it is possible to have the mere acquisition and maintenance of a finite corpus of words (or their use in selected environments) as the goal of a naming treatment, the ultimate measure of treatment effectiveness is the extent to which treatment effects extend to other stimuli, responses, or contexts. The alternative is to target and treat all lexical entries (or environmental or linguistic contexts) that form the desired corpus to be used by the patient. Attempts to develop a technology of generalization for aphasiology have been derived in large measure from the applied behavioural analysis literature and have been drawn primarily from the experimental work on the acquisition of new skills or behaviours from developmentally disabled individuals. This technology has been borrowed liberally and applied to the verbal behaviours of persons with aphasia (e.g. Kearns 1985, Thompson and Kearns 1981). According to this framework, the fundamental issue in promoting generalization is one of determining what represents a stimulus or response class. Several general principles for promoting this discovery and promoting treatment generalization have been proposed (Baer 1981). These include the following.

- (1) Train only behaviours that meet and interact with natural communities of reinforcement. Speech and language production are naturally occurring and naturally reinforced behaviours and, therefore, typically meet this strategic goal.
- (2) Train to a high level of 'fluency' with the task. In the context of attention and resource allocation theory, this translates to training until the task is performed with a great deal of automaticity and hence without a great commitment of processing resources. This usually requires extended practice beyond the typical criterion levels.
- (3) Choose the best examples to treat. This involves choosing the most basic examples or those that contact the greatest number of class members.
- (4) Train multiple exemplars of the response class at the same time. Like many of the other principles for promoting generalization, this principle promotes slower acquisition, but it is believed to promote better generalization.
- (5) Use delayed reinforcements.
- (6) Shift the evaluation of performance from the clinician to the patient.

Although there have been no programmatic research efforts to investigate the utility or the efficacy of these principles for the treatment of aphasia, their effectiveness is often assumed and forms the motivation for many treatments. It was the direct attempt to assess the effects of one of these principles that motivated the current study.

This investigation sought to replicate and extend the treatment findings for one of two patients with dysnomia previously reported by McNeil *et al.* (1997). Specifically, we examined whether training on lexical items from a variety of form classes concurrently would result in greater response generalization than was observed when training was conducted on items within a single form class treated sequentially.

Methods

Subject

At the time of enrolment in the study, B.O. was 63 years old, right handed, male, a retired business owner, and professional musician with 14 years of formal education. He was 19 years post-onset of a single, thromboembolic, left middle cerebral artery stroke with moderate residual right hemiparesis, mild aphasia, and mild apraxia of speech. His performance on the *Western Aphasia Battery* (Kertesz 1982) classified him as having anomia with an AQ of 87.5. Clinician judgements of conversational speech classified him as having Broca's aphasia. Overall *Porch Index of Communicative Ability* (Porch 1981) performance placed him in the 86th percentile for left-hemisphere damaged aphasic adults. Performance on the *Test of Adolescent/Adult Word Finding* (German 1990) placed him within the 46th percentile for age-matched normal adults and his performance was characterized primarily by semantic paraphasias. Performance on the *Apraxia Battery for Adults* (Dabul 1986) revealed behaviour consistent with 'mild' apraxia of speech. Additionally, B.O. performed at the 43rd percentile on the *Revised Token Test* (McNeil and Prescott 1978), read 88 of 100 items correctly on the *Reading Comprehension Battery for Aphasia* (LaPointe and Horner 1979), scored 28/36 on the *Coloured Progressive Matrices* (Raven 1965), scored 67/80 on the *Discourse Comprehension Test* (Brookshire and Nicholas 1993), and scored 3 on the Beck depression inventory (Beck *et al.* 1961), indicating that he was not depressed.

At the initiation of this study, B.O. was 3.5 months post-completion of the previous treatment trial (McNeil *et al.* 1997) in which a double blind, placebo-controlled, multiple baseline design was used to compare placebo and drug treatments in the presence and absence of lexical-semantic activation/inhibition therapy (L-SAIT) (described by McNeil *et al.* 1995).¹ In the McNeil *et al.* (1997) study, baselines were established using lists of 10 words representing each of four form classes (i.e. adjectives, verbs, nouns, and prepositions), and equated for word frequency and syllable number. Each list was probed and trained *sequentially* by form class. The general result of that study was that there was a positive acquisition and maintenance effect, but no generalization was observed to untrained items within or between form classes.

Design

In the current study, a single-subject multiple baseline design across six word lists was used to examine the effects of L-SAIT on the subject's oral production of antonyms and synonyms for trained and untrained lexical items within each list. Each of the six, 10 word lists (60 total) included exemplars from adjective, verb, noun, and preposition form classes.² Repeated measures of the subject's oral production of antonyms and synonyms served as the dependent variables and were obtained during probe sessions conducted during baseline, treatment, and maintenance phases of the study.

¹ Among the purported advantages of L-SAIT over other treatments for dysnomia are the increased ecological validity of elicited, but internally generated naming without picture naming or reading, along with the applicability of the procedures to individuals whose primary disturbance is semantically or phonologically localized, dominated, or equivalently affected.

² A complete list of the words used in this investigation can be obtained from the senior author.

Baseline

Prior to the initiation of treatment, each of the 60 stimulus words was presented orally to the subject during 10 consecutive baseline sessions that spanned a period of 5 weeks. These 60 items were presented twice in each baseline session. During one presentation, the subject was required to provide synonyms for each of the 60 words, and during the other he was required to provide antonyms for those same words. Task order was randomized across sessions, as was the order of presentation of words within each list and the order of presentation of lists within each session. Responses were scored correct when they conformed to antonyms or synonyms provided in a dictionary (Cayne and Lechner 1991) and when the response was given within a 5 s interval following presentation of the stimulus word.

Treatment

Following the collection of baseline data, L-SAIT (McNeil *et al.* 1995) was administered sequentially across four lists for a total of 27 treatment sessions over an 8 week period. Each treatment session consisted of the presentation of the targeted list of 10 words, six times, for a total of 60 training trials. During 30 of these trials the subject's task was to produce antonyms for the words and during the other 30 presentations the subject's task was to produce synonyms for these stimuli. The order of presentation of words comprising the treatment list was randomized, as it was during the collection of baseline data. Each training trial consisted of the clinician orally presenting a lexical item from the treatment list and asking the subject to produce either an antonym or synonym for the word. Correct responses were verbally reinforced, incorrect responses or failure to respond within a 5 s interval were consequence using a cueing hierarchy (McNeil *et al.* 1995) until the desired response was evoked. Acquisition criterion was set at 80% accuracy during six consecutive probe sessions.

Treatment and maintenance probes

Probe data for trained and untrained items were collected prior to the beginning of each treatment session. Procedures identical to those used during baseline determination were employed during probe sessions.

Results

Figure 1 displays the baseline, probe, and maintenance data for the trained and untrained antonym lists. Two judges, who were sophisticated in the analysis of multiple baseline data, independently judged whether or not treatment and maintenance effects were present for each of the four treated lists, and whether or not generalization to as yet untreated lists had occurred for both the antonym and synonym data. Only those positive and negative judgements in which both judges agreed are reported. Judges made their perceptual judgements with the instruction to perform at the 95% confidence level for each judgement. These judgements revealed rapid and marked acquisition effects on all lists that were treated (lists 1–4) for both antonyms and synonyms. Maintenance data for the first antonym and synonym lists showed a marked decrease in performance immediately following treatment termination before stabilizing at levels well above baselines (excluding the final three probes) (figure 1). Examination of performance on lists preceding

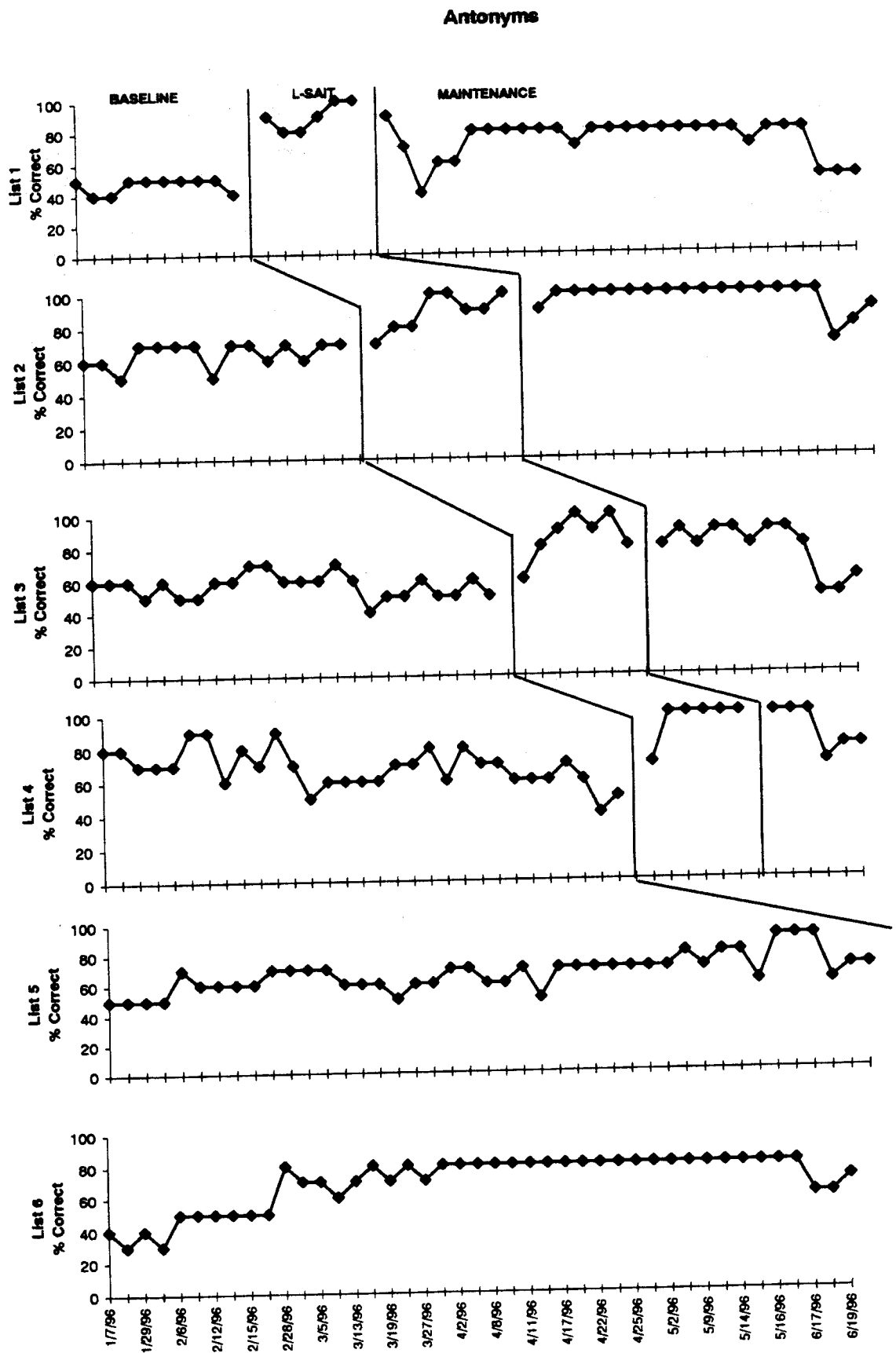


Figure 1. Percentage of correct baseline, probe, and maintenance items named across form classes for the trained and untrained antonym lists. Vertical lines across lists represent initiation and termination of L-SAIT treatment.

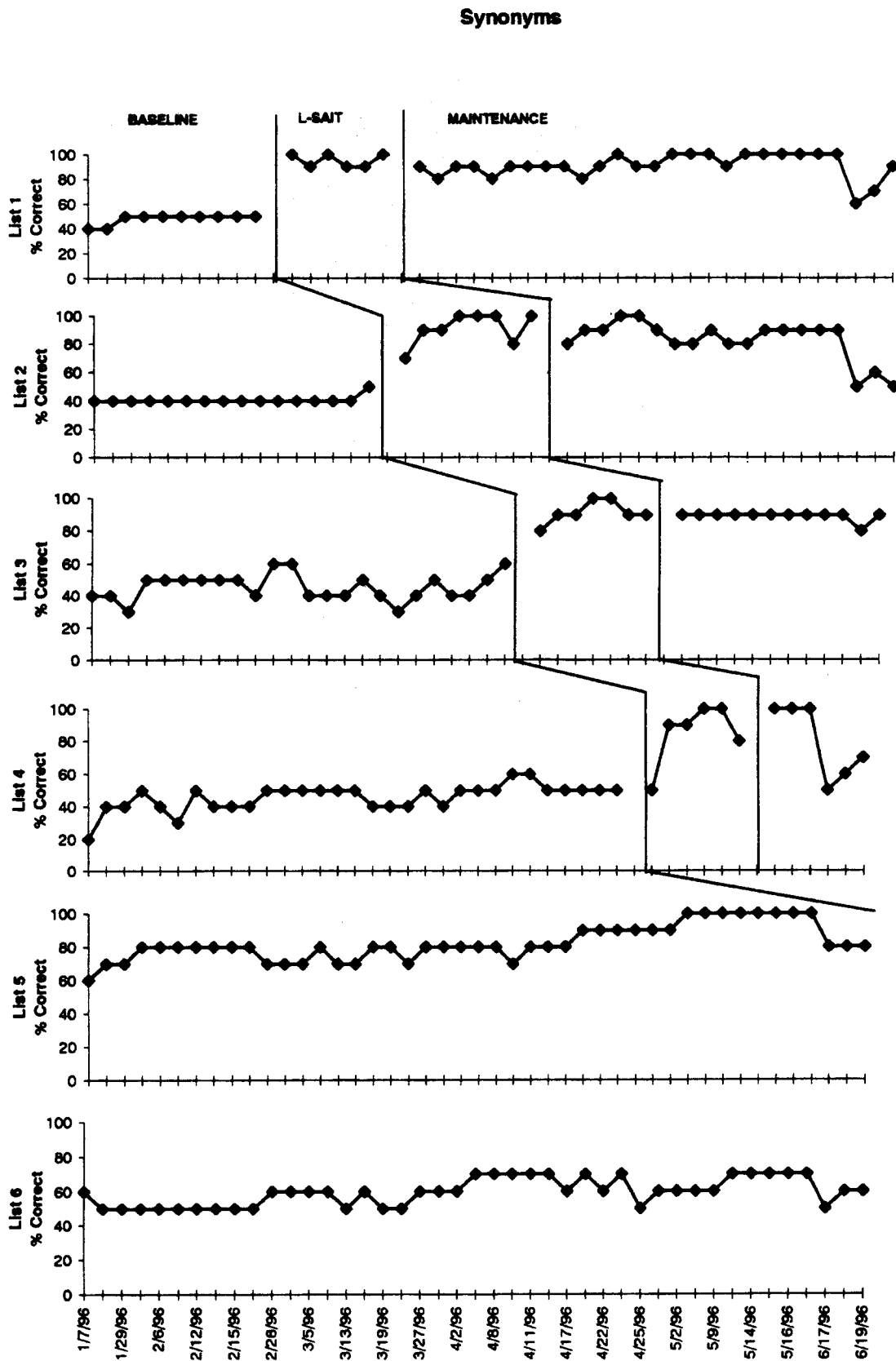


Figure 2. Percentage of correct baseline, probe, and maintenance items named across form classes for the trained and untrained synonym lists. Vertical lines across lists represent initiation and termination of L-SAIT treatment.

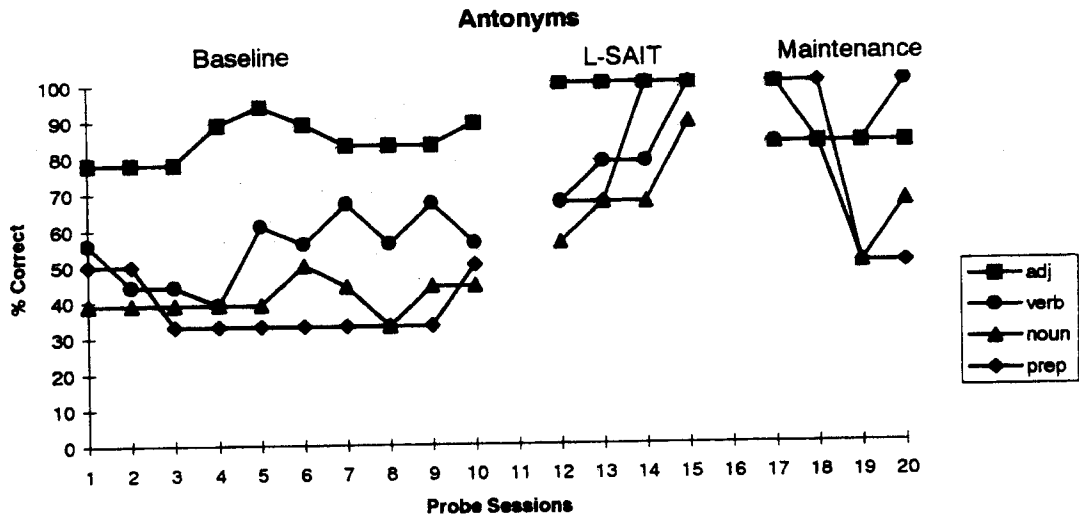


Figure 3. Treated and untreated performance on the antonym lists by form class (adjectives, verbs, nouns, and prepositions) across the first 10 baseline probes, the first four treatment probes, and the first four maintenance probes.

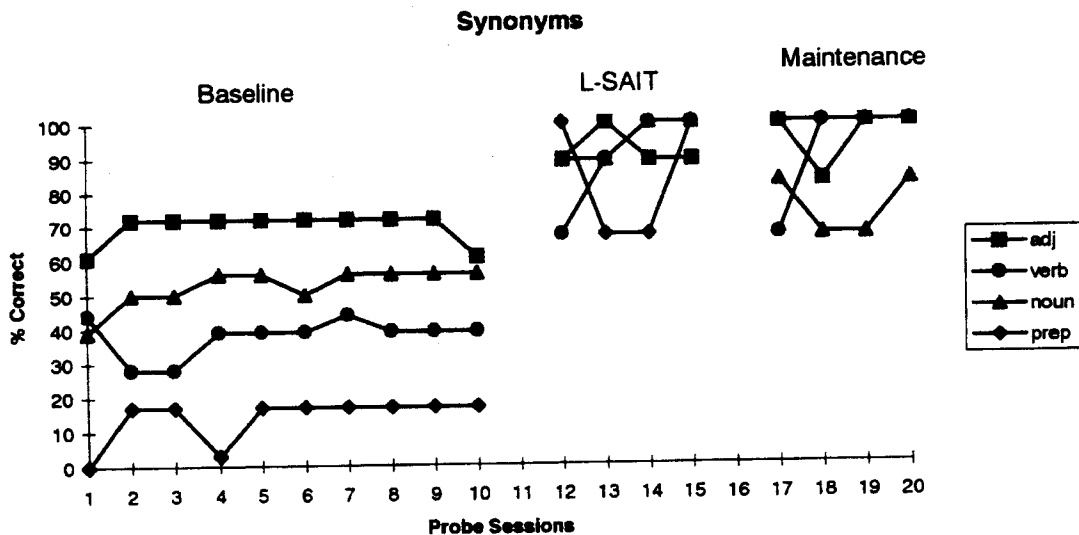


Figure 4. Treated and untreated performance on the synonym lists by form class (adjectives, verbs, nouns, and prepositions) across the first 10 baseline probes, the first four treatment probes, and the first four maintenance probes.

treatment and on untreated lists 5 and 6 evidenced no generalization for either antonyms or synonyms with the single exception of antonym list 6. In this case, generalization was judged to have occurred with the initiation of treatment on list 1.

Results for the synonym lists (figure 2) parallel those of the antonyms except that no generalization was judged to have occurred for any list. Maintenance for all treated lists was generally high and did not drop below treatment termination levels until the final three probe sessions before the experiment was terminated. At this point, performance on all behaviours dropped substantively and to levels consistent with pre-treatment baselines.

In order to evaluate differential effects of the treatment across form classes, acquisition and maintenance probes were combined across lists and plotted according to form class for the antonyms (figure 3) and synonyms (figure 4). Results for the antonyms (figure 3) generally revealed a hierarchy of correct

production by form class, with adjectives being retrieved most effectively, followed by verbs, followed by nouns, which were followed by prepositions. While adjectives were at ceiling throughout treatment, the pattern of acquisition across the other three form classes was quite similar. Maintenance was, however, poorer for the nouns and prepositions than for the adjectives and verbs.

There was a somewhat different form class hierarchy for the synonyms from that for the antonyms, with better performance on the adjectives, nouns, verbs, and prepositions respectively (figure 4). There was no substantive difference across form class for the acquisition or maintenance except that prepositions were generally performed more poorly in both study phases. Pre-treatment performance on the synonym task was generally lower than that of the antonym task. However, acquisition and maintenance effects were not different between the two lexical-semantic activation/inhibition tasks.

Discussion

This study replicated the positive effects of L-SAIT for the acquisition and maintenance of word retrieval for subject B.O. These effects were evidenced for all treated antonym and synonym lists. However, as with the results from the previous study in which form classes were treated sequentially for subject B.O., generalization was *not* evidenced within or across form classes, when treatment was administered to exemplars of all form classes concurrently. The single exception to this conclusion was the presence of generalization on list 6 of the antonyms. While the timing of the change on this list corresponded to the initiation of treatment on list 1, it is difficult to have much confidence that it was in fact the generalization of treatment that caused this change, given that words across the four form classes were randomly assigned to the six lists within each antonym and synonym task. Although the actual level of difficulty of the lists was not equivalent for this subject, a logical explanation for a selective generalization to this, the most difficult list, is not apparent.

It should also be noted that there are suggestions of a positive treatment effect in addition to that described above. While these effects did not reach the judges' 95% confidence level and did not correspond with the initiation of treatment (an expectation of all single-subject, multiple baseline across behaviours design), the design used in this study may have precluded this effect. That is, all lists were similar in composition. Therefore, there may be no *a priori* reason to expect that generalization to a subsequent list would correspond to the termination or initiation of treatment on another list. In fact, if generalization to untreated items was to occur, it might be most logical that it would be produced after the cumulative experience with many stimuli across the lists which were composed of randomly assigned lexical items. If this were the case, the effects would be seen at points in time other than those observed with the initiation or termination of treatment for a particular list, such as those observed on synonym lists 5 and 6. Even with this possible evidence for treatment generalization, it is most conservative and most justified to conclude that there was no clear generalization effects demonstrated in this study.

As predicted, the rate of acquisition to criterion level was achieved more slowly in this subject for both the antonyms and synonyms when L-SAIT was administered to lists containing words from all form classes compared to those

results of the previous study with this subject. Six to eight treatment sessions were required in the current study to reach criterion level, although it took only three to four treatment sessions to reach criterion level when treatment was administered to lists segregated by form class (McNeil *et al.* 1997). Because maintenance was not adequately assessed in the previous studies, a comparison with that achieved in the current study cannot be made. However, in general, maintenance was high and clearly above baseline levels for all treated lists. It will be recalled, however, that in the present study, the final three probes (beginning at the 41st data point) during the maintenance phase decreased precipitously from the previous levels for all lists. There was an entire month between the 40th and 41st probes which corresponded to a period when an illness prevented B.O. from continuing the frequent post-treatment probes. There were no obvious signs of additional neurological or psychological insult at this time and no other account of this systematic and dramatic decrease in performance could be found.

An attempt to evaluate the acquisition, maintenance, and generalization effects separately by form class revealed a hierarchy that was generally consistent, with antonyms being generally easier to retrieve than synonyms. In addition, adjectives were retrieved more accurately than nouns, verbs, and prepositions respectively. Whether this hierarchy represents a true psycholinguistic processing difference for this subject is not clear; adjectives had received the greatest allocation of treatment in the previous study, with nouns followed by verbs receiving less treatment, and prepositions receiving no treatment. Although there were substantial baseline differences across the form classes and although the adjectives were at or near ceiling levels when treatment was initiated, the pattern of acquisition across form classes for both synonyms and antonyms did not differ substantively. Whether or not the poorer maintenance for the nouns in both treatment tasks and for the prepositions in the antonym task reflect a differential vulnerability of this form class in this subject is a matter of interest and a matter for further study.

The demonstration of a strong acquisition and maintenance effect in the absence of a clear generalization effect puts this study in company with the majority of aphasia treatment studies, particularly those on the treatment of naming disorders. Whether this represents a negative finding is of clinical-philosophical interest. Although a treatment cannot be judged effective without strong acquisition and maintenance effects, it is not clear that the generalization of the treatment to untreated stimuli, linguistic and communicative contexts would also be required for a treatment to be judged efficacious. Indeed, Howard *et al.* (1985) targeted the maintenance of acquired naming abilities in addition to generalization. It is conceivable that a goal of a naming regimen might be that of providing the patient with access to a limited number of lexical items, thus targeting acquisition and maintenance, not generalization. None the less, for a treatment to be truly effective in addition to being efficacious, generalization should be expected and targeted in the design of the treatment.

There are any number of variables that could have affected generalization in this study. For example, an insufficient number of exemplars from each form class may have been presented. Likewise, the psycholinguistic make-up of these stimuli (e.g. age of acquisition, operativity, familiarity, imageability, or concreteness) were left largely uncontrolled. Perhaps the most important principle for promoting generalization, and one that is compatible with attention and resource allocation theory as applied to aphasia (McNeil *et al.* 1991), would be the length of time or

number of practice trials that the patient received *after* criterion had been met (or alternatively conceived, changing the criterion to include extended practice once a reasonable level of acquisition had been demonstrated). Although the number of treatment sessions at criterion levels was extended from three in the previous study to six in the current study, its effect on generalization was not evident. Extended practice on a newly acquired task is required before it can be expected to reach even a minimal level of automaticity (Schneider and Shiffrin 1977, Schneider *et al.* 1984, Shiffrin and Schneider 1977). This is true even after achieving a high level of accuracy on the task. Making the task automatic, freeing processing resources for the myriad mental activities (including but not limited to other linguistic computations) that compete with those required for lexical-semantic processing and naming is a goal consistent with an *access* theory of the deficits in naming disorders (Howard and Hatfield 1987). While little is known about the amount or nature of the extended practice necessary for making a naming, or any other linguistic task, 'automatic' for a person with aphasia, it is an area of research that is essential for understanding the resource allocation mechanisms that may underlie naming disorders. It is also an area of research that may be critical for understanding the routine failure of aphasia treatments to promote generalization to untrained stimuli, contexts, and environments.

In spite of what is unknown about the many treatment variables for naming disorders, L-SAIT, as conceived and implemented in this and previous studies, has been shown to produce rapid acquisition of naming ability for a limited repertoire of words. It also evidences maintenance of these effects when treatment is terminated. Generalization of these positive treatment effects has not been demonstrated convincingly to untreated stimuli, linguistic or communicative contexts. There is a great deal of careful experimental treatment research to be accomplished before this treatment (and most other published treatments) should be used for the treatment of naming disorders.

References

- BAER, D. 1981, *How to Plan for Generalization* (Lawrence, Kansas: H & H Enterprises).
- BECK, A. T., WARD, C. H., MENDELSON, M., MOCK, J. and ERBAUGH, J. 1961, An inventory for measuring depression. *Archives of General Psychiatry*, **4**, 561-571.
- BOYLE, M. and COELHO, C. A. 1995, Application of semantic features analysis as a treatment for aphasic dysnomia. *American Journal of Speech-Language Pathology*, **4**, 94-98.
- BROOKSHIRE, R. H. 1975, Effects of prompting on spontaneous naming of pictures by aphasic subjects. *Human Communication Journal*, **3**, 63-71.
- BROOKSHIRE, R. H. and NICHOLAS, L. E. 1993, *Discourse Comprehension Test* (Tucson, AZ: Communication Skill Builders).
- CAYNE, B. S. and LECHNER, D. E. (Eds) 1991, *The New Lexicon Webster's Dictionary of the English Language* (New York: Lexicon Publications).
- DABUL, B. 1986, *The Apraxia Battery for Adults* (Austin, TX: Pro-Ed).
- GERMAN, D. 1990, *The Test of Adolescent/Adult Word Finding* (Allen, TX: DLM).
- HILLIS, A. E. 1989, Efficacy and generalization of treatment for aphasic naming errors. *Archives of Physical Medicine and Rehabilitation*, **70**, 632-636.
- HOWARD, D. and HATFIELD, F. M. 1987, *Aphasia Therapy: Historical and Contemporary Issues* (London: Erlbaum).
- HOWARD, D., PATTERSON, K., FRANKLIN, S., ORCHARD-LISLE, V. and MORTON, J. 1985, Treatment of word retrieval deficits in aphasia. *Brain*, **108**, 817-829.
- KEARNS, K. P. 1985, Response elaboration training for patient initiated utterances. *Clinical Aphasiology*, **15**, 196-204.
- KERTESZ, A. 1982, *Western Aphasia Battery* (New York: Grune & Stratton).

- LAPOINTE, L. L. and HORNER, J. 1979, *Reading Comprehension Battery for Aphasia* (Tigard, OR: C. C. Publications).
- LESSER, R. 1989, Some issues in the neuropsychological rehabilitation of anomia. In X. Seron and G. Deloche (Eds) *Cognitive Approaches in Neuropsychological Rehabilitation* (NJ: Erlbaum), pp. 65-104.
- LINEBAUGH, C. 1997, Lexical retrieval problems: anomia. In L. L. LaPointe (Ed.) *Aphasia and Related Neurogenic Language Disorders* (2nd edn) (New York; Thieme), pp. 112-133.
- LOWELL, S., BEESON, P. M. and HOLLAND, A. L. 1995, The efficacy of a semantic cueing procedure on naming performance of adults with aphasia. *American Journal of speech-Language Pathology*, **4**, 109-114.
- MCNEIL, M. R. and PRESCOTT, T. 1978, *The Revised Token Test* (Austin, TX: Pro-Ed).
- MCNEIL, M. R., DOYLE, P. J., SPENCER, K., GODA, A. J., FLORES, D. K. and SMALL, S. L. 1997, A double-blind, placebo controlled, pharmacological and behavioral treatment of lexical-semantic deficits in aphasia. *Aphasiology*, **11**, 385-400.
- MCNEIL, M. R., ODELL, K. H. and TSENG, C.-H. 1991, Toward the integration of resource allocation into a general theory of aphasia. *Clinical Aphasiology*, **20**, 21-39.
- MCNEIL, M. R., SMALL, S. L., MASTERTON, R. J. and FOSSETT, T. R. D. 1995, Behavioral and pharmacological treatment of lexical-semantic deficits in a single patient with primary progressive aphasia. *American Journal of Speech Language Pathology*, **4**, 76-87.
- NICKELS, L. and HOWARD, D. 1995, Aphasic naming: what matters? *Neuropsychologia*, **33**, 1281-1303.
- PATTERSON, K. E., PURELL, C. and MORTON, J. 1983, Facilitation of word retrieval in aphasia. In C. Code and D. J. Muller (Eds) *Aphasia Therapy* (London: Edward Arnold), pp. 76-87.
- PORCH, B. E. 1981, *Porch Index of Communicative Ability* (Palo Alto, CA: Consulting Psychologists Press).
- RAVEN, J. C. 1965, *Coloured Progressive Matrices* (New York: The Psychological Corporation).
- SCHNEIDER, W. and SHIFFRIN, R. M. 1977, Controlled and automatic human information processing: I. Detection, search and attention. *Psychological Review*, **84**, 1-66.
- SCHNEIDER, W., DUMAIS, S. T. and SHIFFRIN, R. M. 1984, Automatic and control processing and attention. In R. Parasuraman and D. R. Davies (Eds) *Varieties of Attention* (Academic Press, New York), pp. 1-27.
- SERON, X., DELOCHE, G., BASTARD, V., CHASSIN, G. and HERMAND, N. 1979, Word finding difficulties and learning transfer in aphasic patients. *Cortex*, **25**, 149-155.
- SHIFFRIN, R. M. and SCHNEIDER, W. 1977, Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, **84**, 127-190.
- THOMPSON, C. K. and KEARNS, K. P. 1981, An experimental analysis of acquisition, generalization, and maintenance of naming behavior in a patient with anomia. *Clinical Aphasiology*, **11**, 35-42.
- THOMPSON, C. K., HALL, H. R. and SISON, C. E. 1986, Effects of hypnosis and imagery training on naming behavior in aphasia. *Brain and Language*, **28**, 141-153.
- THOMPSON, C. K., REYMER, A. and LE GRAND, H. 1991, Effects of phonologically based treatment on aphasic naming deficits. A model-driven approach. *Clinical Aphasiology*, **20**, 239-261.