GENERALISED TREATMENT EFFECTS AFTER REHABILITATION IN PATIENTS WITH NEUROPSYCHOLOGICAL DEFICITS: THE ROLE OF COGNITIVE MODELS.

by

Lara Harris

A thesis submitted to the University of Birmingham for the degree of

DOCTOR OF PHILOSOPHY

Behavioural Brain Sciences Centre
School of Psychology
University of Birmingham
March 2011

Abstract

The current thesis explored diagnosis and rehabilitation of deficits in memory and language, using a multiple neuropsychological case study design. Broadly, the work evaluated the use of cognitive theory to diagnose patients' clinical presentations and inform rehabilitation methods, and explored how outcomes from these interventions can be used to test cognitive theory in turn. This bi-directional link was explored in two ways: Firstly, theoretically-motivated groupings of word stimuli (e.g. 'neighbourhoods') were used to evaluate patterns of post-therapy generalisation, testing hypothesised associations between types of word stimuli. Secondly, the work identified proposed links between cognitive functions, using rehabilitation to test the validity, and nature, of these associations. The thesis is therefore comprised of two parts: Part 1 explored 'neighbourhood' effects in language and how they might be used to direct generalised improvement following rehabilitation (Chapters 2, 3, 4 and 5) and Part 2 evaluated associations and dissociations of functions in the cognitive architecture, across therapeutic and experimental contexts (Chapters 6 and 7). The work demonstrated that using theoretically-driven stimuli sets in rehabilitation can maximise generalised improvements following language treatment, and detailed how rehabilitation can be harnessed to test the integrity of associations between cognitive functions in the context of multiple deficits.

Acknowledgements

I am very grateful to my supervisors, Glyn Humphreys and Andrew Olson, for their constant support and guidance throughout numerous discussions about the work. Glyn and Andrew have not only taught me most of what I know about research design and theory, but have ultimately led me to pursue a career in Neuropsychology. I would also like to thank the life-long friends I have made at Birmingham for their help, encouragement, understanding, and generally for making the PhD course such an enjoyable experience. Thanks Hayley for being such an amazing friend. Thank you to my wonderful mum and family for always being there to listen and to offer support. Many thanks are also to the School of Psychology for awarding me a school scholarship to fund my degree. Crucially, I am indebted to the inspirational patients who selflessly and enthusiastically gave up their time and energy to assist me in this work. Without them, the thesis would not have been possible.

Table of Contents

List of Figures

List of Tables

List of Appendices

Chapter 1: General Introduction

Cognitive neuropsychology is an important tool for exploring various patterns of disorders, and thorough assessment and diagnosis is essential for guiding the focus of rehabilitation (Caramazza, 1989). Many rehabilitation studies have used cognitive theory past the stage of diagnosis, applying cognitive models to treatment design (e.g. in surface dyslexia: Coltheart and Byng, 1989, sentence production: Schwartz, Saffran, Fink, Myers, and Martin, 1994, and memory and sentence comprehension: Francis, Clark and Humphreys, 2003). Outcomes from cognitive rehabilitation studies can, in turn, be an informative test of cognitive models, an idea that is currently under explored and forms a major part of the current thesis. Without careful consideration, this type of approach can render links between the lost function and the training designed to remedy it ambiguous as well as removing the possibility of exploring the nature of the deficit. Identifying and resolving problems faced by cognitive rehabilitation methodologies at the planning stage is a crucial step towards framing rehabilitation outcomes in terms of theoretical findings. This first chapter identifies and discusses some of the various issues facing cognitive neuropsychology and rehabilitation using examples from the current literature. The relationship between cognitive theory and rehabilitation is outlined in two directions, first discussing how cognitive theory can help guide effective cognitive rehabilitation, and then describing how treatment outcomes can inform cognitive theory.

1. The use of cognitive models in assessment and diagnosis

Coltheart and Byng (1989) reported a patient presenting with surface dyslexia (EE). Referring to dual route models of reading (e.g. Coltheart, 1985, Ellis and Young, 1988, Patterson and Shewell, 1987) the authors first identified the processes of word recognition, word comprehension and retrieval of phonological forms as possible causes of surface dyslexia. After assessing functional abilities at each level, the authors concluded that EE suffered impaired written word recognition, on account of his poor performance when comprehending homophones. The authors reasoned that if one word is taken to mean another with the same phonology (e.g. break, brake), the processes of letter identification, translation from orthography to phonology, pronunciation and comprehension are all preserved, and the damage must therefore lie at the level of recognising a written word he has already processed. This is a good example of how neuropsychological assessment can be used to identify the locus of a patient's impairment. The diagnosis led to a treatment producing a dramatic treatment gains on the trained items, with some generalised improvement to untreated items.

2. Directly treating the impaired function

There are a number of assertions that can be deduced from Coltheart and Byng's (1989) findings: i) remediating lost written word recognition leads to improved reading performance in this case of surface dyslexia, and ii) the effectiveness of the treatment reinforces the accuracy of the initial diagnosis. However, Patterson (1994) highlights the difficulty in making conclusions about the nature of the treatment effect in Coltheart and Byng's (1989) study, warning that the strategy may in fact promote use of preserved skills in word comprehension to aid written word recognition. Their rehabilitation procedure targeted a fairly high-level skill (matching written words to pictures), and so could potentially have trained word comprehension more generally, leading to a better performance on written word recognition tasks post therapy due only to strengthened semantic representations of written word items. This example describes a ‘reorganisation’ method of rehabilitation, whereby promoted use of preserved cognitive skills facilitates impaired performance on a task geared at assessing the impaired function, i.e. the impaired function is not targeted directly. In other words, there is some uncertainty about whether the targeted skill was actually treated and improved. Patterson (1994) reasons that a more theoretically informative method might adopt an ‘ABACADA’ design, where ‘A’ denotes baseline and ‘B/C/D’ denote treatment strategies. In this sense, the authors could have hierarchically applied three treatment strategies, each targeted at improving skills at each level specified by a dual–route model, with the baseline phases assessing the extent and nature of success in each case. This type of staged rehabilitation design has been applied to the treatment of various cognitive impairments in order to compare efficacy of each intervention (Domahs, Bartha and Delazer, 2003, Sage, Hesketh and Lambon Ralph, 2005).

However, there are many considerations and potential problems even within this empirically grounded approach. Firstly, a staged therapy must consider the ordering of processing stages in cognitive models. In surface dyslexia, when considered within dual route models of reading, it is intuitive to first train the lowest-level function, which in this case is letter identification, the ability to use letter information (orthographic input lexicon), then word comprehension (semantic system) and finally word recognition (integrity of the link between orthographic input and the semantic system). Treating the higher-level function that most closely resembles the end diagnosis may, as in this case, lead to successful a treatment outcome, but the possibility that one might be training a number of lower-level abilities renders assertions about the nature of the treatment effects tentative.

3. Ensuring clinical relevance

If a rehabilitation technique is to be clinically valid, it should produce the optimum and most generalised improvement achievable within the shortest possible amount of time. To this end, empirically-based design restrictions such as those outlined above may not be as important as creating a treatment most likely to improve the lost skill. Indeed, neuropsychological rehabilitation studies may not be as practicable within a clinical context, where one cannot be as generous with resources and time as in a research department.

However, there may be important practical implications of treating numerous levels contributing to a given cognitive process: If the produced improvement was due to support from a lower-level function than anticipated, then therapy could focus on remediating this specific behaviour, potentially cutting resources and time drastically. One of the themes of the current thesis is to report specifically-targeted and empirically driven rehabilitation studies in the hope that the most effective stage of therapy can be applied quickly and effectively in a clinical environment.

4. Reorganisation and restoration approaches

Rehabilitation strategies can adopt either a restoration or reorganisation approach, where restoration seeks to improve the impaired function directly and reorganisation recruits preserved regions and abilities to replace the lost skill. As discussed above, reorganisation cannot be ruled out in Coltheart and Byng's (1989) treatment geared at remediating written word recognition specifically (designed as a restoration method). In terms of producing theoretically revealing therapy outcomes, restoration may be the most fruitful approach. Thorough consideration of the patient's baseline abilities and the nature of the function being re-trained are instrumental in deciding on either a restorative or a reorganisation approach. Firstly, reorganisation approaches demand sufficient preserved function to form an alternative route by which the patient can arrive at the same outcome produced by the impaired function. Secondly, different skills may be more amenable to one approach over another. Arithmetic, for example, represents a special case where knowledge consists of many highly related facts that are inherently organised and limited, making arithmetic well suited to a reorganisation approach (Girelli and Seron, 2001).

When using rehabilitation methods to make inferences about cognitive theory, a restoration approach is perhaps the more sensible choice in most cases. Reorganisation techniques typically enlist other (preserved) functions to restore the impaired behaviour, negating any test of functional association between the improved skill and other abilities that may show generalised treatment gains.

5. Errorful and errorless learning paradigms

The rehabilitation literature has explored the use of errorless learning paradigms, where information is presented to the patient in a way that reduces the possibility of mistakes. This type of method has been applied successfully to patients with severe memory problems (Baddeley and Wilson, 1994; Clare, Wilson, Breen and Hodges, 1999; Evans et al., 2000; Glisky, 1995; Squires, Hunkin and Parkin, 1996; Wilson and Evans, 1996). Whether these approaches work by enlisting implicit or explicit memory mechanisms seems to depend on the integrity of these types of memory in patients. For example, patients with very poor explicit memory will depend on the more preserved mechanisms of implicit memory, but in cases where some explicit memory is spared, errorless learning may enlist both systems (Page, Wilson, Norris, Shiel, and Carter, 2001).

Errorful (or error-based) learning involves the repeated presentation of a given task, with regular feedback from a therapist as to the success of each trial. The effectiveness of this approach can be explained within a 'Hebbian learning' account. Hebbian theory explains any increase in synaptic efficacy during rehabilitation as resulting from the presynaptic cell's repeated stimulation of the postsynaptic cell (thus, synaptic plasticity). These neural changes reflect a modification in behaviour, i.e. the relearning of an impaired function.

Recent work has suggested that errorless learning may be as effective as errorful therapy in treating aphasic word finding difficulties (Fillingham, Hodgson, Sage and Lambon Ralph, 2003; Fillingham, Sage and Lambon Ralph 2005a, 2005b, 2006). Other studies have shown that errorful therapy may be more effective in some cases (Abel, Schultz, Radermacher, Willmes and Huber, 2005). One study by Conroy and colleagues has shown that a therapy using decreasing cues (beginning by providing the picture, presenting its name verbally and visually, and gradually decreasing cues) was as effective as a therapy using increasing cues (gradually working up to maximal cues) at improving naming speed and accuracy for nouns and verbs in aphasia (Conroy, Sage and Lambon Ralph, 2009).

In the current thesis, error-reducing approaches were always employed. This was the case where the locus of the treated impairment was likely to be memory-based (e.g. in the context of a graphemic buffer deficit in Chapter 3, and in the semantic and phonological STM-impaired patients in Chapter 6), and also when the therapies treated more online processes in production (of sentences involving past tense verbs in Chapter 2, and reading performance in letter-by-letter readers in Chapter 5).

6. Using rehabilitation to explore associations between cognitive functions

Many cognitive functions are thought to be functionally and neurologically linked. Memory in particular is thought to subserve a range of cognitive functions (e.g. text comprehension in Daneman and Carpenter, 1980 and following directions in Engle, Carullo and Collins, 1991). An association between memory and sentence comprehension has been posited by some researchers (e.g., capacity constraint theories: Just and Carpenter, 1992). Evidence from dual task methodologies has demonstrated that the ability to understand individual sentences decreases when individuals are required to simultaneously retain lists of digits for offline sentence comprehension (i.e. influencing accuracy judgements after the sentences have been read, Baddeley and Hitch, 1974, Baddeley 1986, Hitch and Baddeley, 1976, Gordon, Hendrick and Levine, 2002) and when retaining lists of words in an online capacity (i.e. affecting reading times, Federenko, Gibson and Rohde, 2006).

Most neuropsychological work on the association between sentence comprehension and memory has employed detailed assessment to identify a memory problem, and test for coinciding deficits in comprehension, with any significant match taken as evidence that the two processes are related (e.g., patient PV: Vallar and Baddeley, 1984 and 1987). There are several problems with this approach. Firstly, neurological insult often produces a range of clinical symptoms, rendering assertions about whether this reflects relatedness between cognitive behaviours difficult. Secondly, even where a language processing deficit in conditions with lesser demands on memory (e.g. single word comprehension) can be ruled out, the generality of such an approach fails to show how comprehension and memory are associated. Rehabilitation can therefore be a very informative measure of links between cognitive processes: Providing that the treated memory impairment *specifically* did improve, as is testable via baselines across behaviours, any generalised effect to sentence comprehension can provide convincing evidence as to whether or not the two functions are associated.

The topic of associations between working memory and sentence comprehension has been made more interesting by hypothesised distinctions between memory for phonological and semantic information. Phonological similarity and recency effects are thought to be products of the phonological component of STM, while lexicality and primacy effects are due to encoding of semantic features of words. Just as specialised memory tasks enlist memory for different types of information, dissociations between phonological and semantic STM have been noted in the neuropsychological literature (Martin and Romani, 1994). Distinctions between the retention of semantic and phonological information have been posited between different sentence types. In some conceptualisations of sentence comprehension, an ambiguous or 'garden path' sentence requires retention of phonological features, as the initial interpretation of the early parts of the sentence are discounted by the latter. For example, in the sentence 'Today the panel finally gave in after warping and sagging for months' the dominant interpretation of 'panel' is that of a committee required to make a decision, but this initial interpretation is invalidated by the latter part of the sentence, that confirms the definition of panel as a component of a fence.

Memory for lexical features is necessary for relative structures (e.g. 'I met the girl that the grandma drew') as the reactivation at the position of the gap is semantic and not phonological (the reactivation is of the word meaning rather than word form)

(Martin, 1990). More recent work on this topic has made distinctions based on the modality of testing rather than intrinsic differences in the sentences themselves. Sentence repetition is thought to assess phonological STM, while comprehension (tested using acceptability judgements) employs semantic STM (Hanten and Martin, 2000). Employing a staged ABACA design treating phonological and semantic STM separately, then testing for generalised improvement to comprehension on a range of sentence stimuli may show which sentences use each type of STM. This idea is elaborated on in the next section.

7. The importance of categorising test stimuli

Most rehabilitation studies do not categorise untreated stimuli, instead pooling different types of items together (Sage et al. 2005, Francis et al. 2003). This is particularly problematic for the treatment of processes that may contain dissociations (e.g. the phonological and semantic components of comprehension and STM discussed above), but also for any rehabilitation attempt employing word stimuli. Effects of lexical variables have long been documented in the language literature (e.g. word frequency effects in production, Jescheniak, and Levelt**,** 1994). While any confounding effects introduced by most lexical variables are removed by using pre- and post- treatment baselines as the main statistical comparison, the level of relatedness to / groupings of these stimuli with, the trained set are important considerations as these are factors that may affect the post-therapy baseline. Coltheart and Byng's (1989) paradigm promoting the use of mnemonics in word recognition capitalises on the semantics of a given item. It is therefore plausible to predict that semantic representations of untreated items from the same semantic category as a word in the treated set may be particularly activated, boosting performance on these words. This is a consideration that applies to most rehabilitation studies using word stimuli.

Connectionist theories assert that activation of a word receives input from its phonological and orthographic neighbourhoods (e.g. McClelland, Rumelhart and Hinton, 1986), and that this effect may be mediated by frequency / familiarity of neighbouring items (Baus, Costa and Carreias, 2008). Failing to control / manipulate the neighbourhood variable in stimuli could lead to some untreated items benefiting from neighbouring items in the training set, facilitating performance on untreated items at post-, but not pre- treatment baselines.

Furthermore, some types of stimuli may benefit from a rehabilitation approach more so than other types. For example, in the letter-by-letter reading literature, words comprised of low confusability letters remove the length effect and produce effects of lexical variables, both findings that are indicative of parallel processing (length effect: Fiset, Arguin, Bub, Humphreys and Riddoch, 2005, lexical variables: Fiset, Arguin and McCabe, 2006). It might therefore be argued that low-confusability stimuli benefit more so by a therapy encouraging parallel processing, despite eliciting the same type of serial reading behaviour as high-confusability letter-words at baseline. Essentially, looking at changes in performance on different types of untreated stimuli tells us something about the *nature* of a given treatment effect, and may also shed light on how cognitive processes in the normal brain use sets of over-learned representations to support performance on other words, and in what learning contexts and with which types of associations (e.g. phonological, semantic) this support might be facilitative.

This approach could also uncover explanations for stark contrasts in the treatment effects produced by similar interventions with patients presenting with the similar types and extents of damage, a major problem facing neuropsychological rehabilitation research as a clinical tool. Specifically, can the degree of relatedness between untreated and treated stimuli predict the likelihood of generalised improvement? Manipulating stimulus groupings may be fundamental to understanding how rehabilitation takes place.

Current thesis

Howard and Patterson (1989) note that outcomes from rehabilitation studies with neuropsychological patients can go further than standardised tests in exploring cognitive systems. Throughout, the thesis will describe examples where rehabilitation studies test the direct relations between specific cognitive processes by improving one ability and testing for treatment gains in another, and how contrasting post-therapy performance between different groups of stimuli (e.g. orthographic/phonological neighbours) following treatment of a stimuli set can evaluate how information is represented and used.

We present data from patients presenting with deficits in language (DS, JF, MAH, DM) and short term memory (DS, AK, MM, JF), collated throughout different rehabilitation strategies. A key idea pursued by the current thesis is to use thorough diagnostic investigations and findings from rehabilitative interventions to test cognitive theories.

Chapter Two investigated the impact of learning stem and past-tense verb associations on untreated verbs that take the same (rhyming) versus different transformations as the treated set. This method tested whether single mechanism assumptions that verbs benefit from phonological support from these ‘verb clusters’ in a nonfluent aphasic patient.

Chapter Three tested several specific assumptions made by single mechanism models of past tense processing. One key aim was to evaluate the relative contributions of support from these verb clusters as well as from meaning knowledge and lexical frequency in a patient with impaired semantic knowledge.

Chapter Four explored whether graphemic buffer impairments use lexical information in contrasting post-therapy improvement on untreated neighbourhoods with shared versus different middle sections to explore whether the *N* effect observed in these patients is due to items receiving lexical input, or to a facilitative effect of learning middle of word combinations. The reasoning behind this study was that central letter positions are especially problematic in GB patients, and letter combinations in the central positions are usually held constant over orthographic neighbourhoods.

Chapter Five assessed the recent claims that letter confusability might mediate serial reading behaviour in letter-by-letter reading, and considers how this finding can be used to optimise rehabilitation outcomes.

Chapter Six evaluates the link between memory and sentence comprehension by improving phonological and semantic STM in a rehabilitation study of phonological and semantic STM patients, and testing comprehension of sentences that use phonological vs. semantic information at pre- and post- treatment baselines.

Chapter Seven investigates the claim that phonological and semantic STM patients suffer type-specific interference, and explores the extent to which their memory impairments are explained by susceptibility to types of PI.

Part 1: 'Neighbourhood' effects in language and their role in directing generalised improvement after rehabilitation.

Chapter 2: Rehabilitation of past tense verb production and non-linear sentence production in left inferior frontal non-fluent aphasia

Abstract

An impairment in generating regular relative to irregular past tense verbs has been noted in nonfluent aphasia. Dual mechanism accounts attribute this deficit to an impaired rule-based route responsible for stem affix (‘play’ + ‘ed’) operations, and spared retrieval of past tenses from lexical knowledge. In contrast, single mechanism accounts suggest that the dissociation reflects an impairment in the phonological representation of verb forms. A rehabilitation study was reported, targeted at improving the production of past tense verbs and nonlinear sentences in a nonfluent aphasic patient (DS). The study used a thematic mapping procedure requiring DS first to describe what was happening in a picture with two characters, and then describe the events again but starting with the other character. Two thematic questions followed and DS was asked to produce the linear and non-linear sentences again. Generalised improvements were seen in the production of non-linear sentences and in producing treated and untreated regular past tense verbs. Interestingly, the improvement on treated irregular verbs generalised only to untreated irregulars from the same neighbourhood groups as treated words. The findings support an account positing generalisation based on lexical neighbourhoods.

Introduction

Patients described in the literature as fitting a non-fluent aphasic profile have shown impairments in i) forced production of sentences with non-linear orders of thematic roles (Caplan and Hanna, 1998; Menn, 1999), ii) verb retrieval and verb inflection (Bird, Lambon Ralph, Seidenberg, McClelland and Patterson, 2003, Longworth, Tyler and Marslen-Wilson, 2003), iii) phonological and articulatory processing (Kurowsi, Hazen, and Blumstein, 2003), and iv) comprehension and production of semantically reversible sentences (e.g. 'the boy is kissing the girl'; see Berndt, Mitchum and Haendiges, 1996 for a meta-analysis). The failure of some patients with non-fluent aphasia to process semantically reversible sentences is thought to arise from a failure to associate nouns to their thematic roles.

Syntactic processing in aphasia

Most generative approaches (e.g. the Principles and Parameters Approach, Chomsky, 1981) hold that sentences are coded in a hierarchical manner, with words assigned to slots in the hierarchy on the basis of their syntactic role, and their fit with a given thematic role. It has been argued that non-fluent aphasic patients can follow rules of hierarchical syntactic organisation, but cannot correctly assign thematic roles in non-linear sentences due to difficulties in representing traces. Failure to assign roles in sentences leads to a superficial assignment of agent status to the first noun class in a sentence, and theme status to the second noun, and this difficulty is apparent in both production and comprehension (Caplan and Hildebrandt, 1988).

Garrett's sentence production model (1975, 1980) provides a valuable framework for rehabilitation research into non-fluent aphasic speech. The model includes two levels, the positional level, that codes thematic relationships between noun classes, and the functional level, responsible for the ordering of constituents to form the syntactical frame. In non-fluent aphasia, incorrect thematic role assignment may arise from failures in mapping between these two levels, prompting treatment procedures targeted at improving mapping between the positional and functional levels of sentence production (e.g. Schwartz, Saffran, Fink, Myers, and Martin, 1994).

Rehabilitation in aphasia

Sentence processing

Rehabilitation attempts levelled at improving the processing of sentences with noun phrases moved from their linear positions have focussed on the processing of object relatives (e.g. the ball is kicked by the man). Other types of sentences requiring the movement of an argument noun phrase are who and what; and when and where questions. Production of such structures is problematic in non-fluent aphasia (e.g. Friedmann 2002). Thompson, Shapiro, Tait, Jacobs, Schneider (1996) report treatment effects of seven non-fluent aphasic patients following training designed to remediate production of these forms by stressing lexical and syntactic features of the sentences and the *wh*- movement required. The treatment required the patients to listen to a sentence (e.g. the soldier is pushing a woman in the street) and they were then trained to produce who and what questions (prompted by showing patients videotaped examples of actors acting out the procedure) based on the sentence (in this example the target was ‘who is the soldier pushing in the street?’). The patients were required to identify the verb, the verb argument structure, and the thematic roles of the noun phrases. Production of *wh*-movement was treated by asking patients to move the sentence constituent and replace this with the *wh*-morpheme, and to produce the final targeted *wh*-question. Following therapy, sentence production abilities improved in all seven patients, across constrained sentence production and discourse tasks. Further, three patients produced correct argument movement on both treated and untreated items following training indicating some generalisation from training.

Many rehabilitation attempts targeting improvements in the sentence processing deficits of non-fluent aphasic patients have focussed on comprehension rather than production. Mapping therapies have used various procedures. Schwartz et al. (1994) use a sentence query approach, which required the patient to identify the verb in the sentence, and respond to questions about thematic roles ('who is kicking' etc.). Post-therapy, three patients still demonstrated impaired processing of sentence structure / content and two showed no sentence comprehension deficits after treatment. The authors state that the best post-therapy improvements were seen in patients with pure agrammatism, whereas patients with more complex pre-therapy patterns of deficits showed limited improvement (see Mitchum, Greenwald and Berndt, 2000 for a review of sentence comprehension treatments).

A key focus of the current study is on rehabilitation of sentence production. Rochon, Laird, Bose and Scofield (2005) investigated the production of reversible linear and non-linear sentences after mapping therapy was conducted with five non-fluent aphasic patients experiencing left hemisphere cerebral vascular accident (CVA). Rochon et al. (2005) used pictures depicting the reversible target sentences, with an indicator to emphasise the difference between the agent and object. As they were interested in the understanding of thematic relationships only, Rochon et al. (2005) did not investigate verb inflection, and therefore did not penalise inflection errors ('the patient calls by the lawyer was accepted for 'the patient is called by the lawyer'). The patients were given a series of detailed cues by the examiner in the treatment, including both the present and past tense verb inflections, and information regarding syntactical frame (e.g. "the one being chased is the tall teacher", "the one doing the chasing is the nurse"). Following the 2.5-month rehabilitation procedure, the patients showed cross-task generalisation of treated and some untreated sentence structures on tasks of constrained sentence production, although there was no generalised improvement in comprehension abilities. In this study, the aim was to include the production of past tense verbs in sentences in the analysis to see whether past tense production can be treated successfully within a mapping therapy procedure.

Past tense processing

One main focus of theoretical debate has concerned how regular and irregular past tense verbs are processed, and whether two distinct mechanisms are necessary to explain performance. Dissociations in performance between regular and irregular past tense forms in neurological patients have been reported; with nonfluent aphasic patients sometimes showing a profile in which irregular past tense forms are generated more accurately than regular past tense forms. In contrast, the reverse dissociation is reported in patients with semantic difficulties (Patterson, Lambon Ralph, Hodges and McClelland, 2001). There are, however, reports of the opposite profile of dissociations reported in German and Dutch agrammatic patients (a major impairment in producing irregular past tense forms and a relatively preserved ability to produce regular past-tense forms, Penke and Westerman, 2006), suggesting that this pattern may be specific to the English past tense. However, recent work has attributed this finding to greater phonological complexity of the irregular past tense in the German language (Seidenberg and Arnoldussen, 2003).

Dissociations between regular and irregular past tenses in neurologically damaged patients are accounted for by both dual-route and single-mechanism accounts of past tense production. Dual route theorists (e.g. Marslen-Wilson and Tyler, 2007; Tyler, Randall and Marslen-Wilson, 2002; Tyler, Stamatakis and Marslen-Wilson, 2005) assert that such dissociations indicate that there are two separate mechanisms at work during past tense processing: a lexical store responsible for irregular forms and a rule-based process enabling the stem + 'ed' regular affix to be produced. Their assertion is that, in nonfluent aphasic patients, the rule-based process is impaired whereas in fluent aphasic patients damage is to the lexical store. Conversely, single mechanism theorists (e.g. Patterson et al., 2001) explain the dissociation demonstrated by some nonfluent aphasic patients as due to difficulties in phonological processing, which disproportionately affects simple regular past tense forms, as they are often more phonologically complex than irregulars. Cases where irregular forms are worse than regular are explained as stemming from a reduced input from word meaning (e.g. in cases of degraded semantic information), which leads to difficulties in generating lower frequency irregular verbs.

Dual mechanism account

Nonfluent aphasic patients sustaining left inferior frontal damage have been shown to experience difficulties specifically with regular verb inflections (Tyler at al., 2002), while research suggests that irregular verb inflections are more problematic for fluent aphasic patients with posterior left cortical regions (Ullman et al., 1997). Evidence from fMRI studies indicates that a fronto-temporal network linking the anterior cingulate, left inferior frontal cortex, and bilateral superior temporal gyrus via the arcuate fasciculus mediates regular past tense processing, whereas irregulars do not activate the frontal cortex to the same extent (Tyler et al., 2005; Marslen-Wilson and Tyler, 2007, though see Desai, Conant, Waldron and Binder, 2006 for conflicting results).

The dual systems account (e.g. Marslen-Wilson and Tyler, 2007; Tyler et al., 2002; Tyler et al., 2005; Tyler et al., 2005) describes two separate neural systems underpinning language function: Regular verb inflections are possible by way of phonological parsing via a fronto-temporal network, which allows identification of stem and affix. Irregular verb inflections, comprising no internal morpho-phonological structure to allow predictable mappings, are instead recalled from a rote-learned knowledge store of whole forms through a temporal lobe route (Tyler, Randall and Marslen-Wilson, 2002).

An alternative view within the dual-route literature is the 'Words and Rules Account' (Marcus; Pinker; Ullman; Hollander; Rosen, and Xu, 1992; Pinker, 1999; Pinker and Ullman, 2002). As in the dual-mechanism model, regular stem-affix transformations are conducted online and whole irregulars are recalled from a lexical store. However, it is noted that a number of irregular past tense forms can follow a rule-based procedure that could potentially be applied to other similar past tense forms (i.e. keep – kept, sleep – slept). Therefore, as opposed to the dual-mechanism model, the Words and Rules account portrays the lexical store as capable of identifying common types of morphological form, and enabling generalisation to novel forms. For example, sleep – slept may be applied to nonwords with similar morpho-phonological structure (e.g. weep-wept). Some studies have also found irregular-style generalisations to non-words ('splung' was produced as a past tense for 'spling' in Prasada and Pinker 1993).

Noting the degree of morpho-phonological consistency across groups of irregular past tense forms, Pinker (1999) suggests that irregular verbs are processed by a lexical mechanism that is capable of detecting properties like frequency and similarity. Before producing a past tense verb, lexical memory is searched to find the past tense form. If no form is found, lexical memory produces a novel exception form for inputs similar to real irregular past tense verbs by default, resulting in generalisation of irregular past tenses to other irregulars, and 'irregularisations' on novel (non-) words. If no similar irregular past tense form is found, the root verb enters the morpho-phonological parsing process, giving the regular '+ed' affix. Ullman asserts that more frankly irregular verbs such as 'went' cannot be produced by applying the '+d affix', and are instead recalled from lexical memory.

Single mechanism account

McClelland and Patterson (2002a) note that while the words and rules account is correct to acknowledge verb clusters in English irregular verbs, it does not explain why so many exceptions share properties with regular past tenses. McClelland and Patterson assert that the vast majority of irregular verbs are 'quasi-regular', and cite only two irregulars where the past tense forms are completely different to their stems (the extremely frequent forms 'go-went', and 'am-was'). They note that it is not clear how the finding of mixed errors such as tear = 'tored' can be accommodated within the words and rules account as, under this account, successful generation of an irregular past tense form should block application of the regular affix.

The single mechanism view can accommodate the findings of both irregular-style transformations applied to nonwords and regulars, and instances of mixed errors containing an irregular whole past tense form with a regular affix. The single mechanism approach is based on Parallel Distributed Processing (PDP) networks (e.g. McClelland and Rumelhart, 1986; Joanisse and Seidenberg, 1999). Such connectionist models describe a single process whereby a single network of units and connections map from verb stems to their past tense forms. In these models, the production of past tense verbs is modulated by frequency: High frequency irregular verbs are accessed quickly through strong mapping processes and are not affected by competition from the regular past tense inflection, whereas lower frequency past tenses are accessed less quickly and may be subject to some interference from the regular inflection resulting in mixed errors such as 'tored'. A key notion is that all verbs receive additional input from semantic knowledge about the word, but that low frequency irregular verbs rely on this additional input more so than high frequency past tense irregular verbs and regulars, as low frequency irregular verbs are not as firmly captured by the direct phonological transformation process (Patterson, Lambon Ralph, Hodges, and McClelland, 2001).

A key assumption of the single mechanism approach (e.g. McClelland and Patterson, 2002a and 2002b, Lambon Ralph, Braber, McClelland and Patterson, 2005) is that the processing of both regular and irregular forms depends on one complex procedure (Patterson, Lambon Ralph, Hodges and McClelland, 2001). They assert that neurological dissociations in past tense verb processing arise from reduced input from word meaning (impacting on lower-frequency irregular verbs), and phonological deficits that make regular and novel forms problematic (Patterson et al., 2001; Lambon Ralph et al., 2005). The approach describes an integrated connectionist network that processes irregulars in the same way as regulars. That is, to copy the stem to the past tense form and apply the /d/, /t/, or /^d/ affix depending on the final consonant (McClelland and Patterson, 2002a). McClelland and Patterson (2002a) outline nine groups of irregular past tense verbs, all of which contain aspects of the regular /d/, /t/ or /^d/ affix. For example, many verbs involve the addition of /d/ or /t/ in their transformations (e.g. keep, sleep, creep, say, do, tell, sell hear etc.). Their analysis revealed that 59% of irregular verbs could be allocated into the 8 groups of 'quasi-regular' irregular past tense verbs, and assert that the vast majority of remaining irregulars can be classed as belonging to one group (9) involving a vowel change rather than a /d/ or /t/ affix.

Despite various rehabilitation reports aiming to improve non-linear sentence processing in non-fluent aphasic patients, there are few instances where rehabilitation studies have contributed to the past tense verb discussion, and no intervention studies to my knowledge that have directly tested the single-mechanism claim that sets of irregulars use the same adjustment of the regular stem-affix operation. The current study treats regular and irregular verbs using a treatment procedure targeting the production of non-linear sentences. My aim was not to make any assertions about the effect of treating syntax on verb inflection, or vice versa. Rather, the motivation was to produce optimal improvements in both abilities using the same treatment programme. The method will explore generalised improvement not only from treated regular forms to untreated regular forms, a finding that can be accommodated by both the single and the Marslen-Wilson and Tyler dual-mechanism account, but also to untreated irregular verbs from the same groups as treated items. This finding cannot be accommodated within the Marslen-Wilson and Tyler (e.g. 2007) dual-mechanism model as the lexical store responsible for irregulars is described as rote learned, and unable to generalise to untreated irregulars in this way.

Rehabilitation of past tense verb processing

Weinrich, Boser, and McCall (1999) report a rehabilitation procedure targeted at past tense verb production of a non-fluent aphasic patient. They trained their patient, EA, on seven regular and six irregular verbs and sentence structures using computerised visual communication (C-VIC). The display in the C-VIC programme they used shows a top row of icons depicting noun classes and actions, questions (when, who, did what, to what?), and the patient was trained to respond by dragging icons into the communication spaces to match the target phrase (e.g. past, woman, wash, plate). The treatment targeted the production of past tense verbs in sentences, and the authors suggest that previous success with the method, which does not involve any explicit training of syntax rules, reflects either stabilised functional representations, or strengthened connections between functional and positional representations. For the treated set, the sentences involved seven regular and six irregular verbs. To assess generalised improvement to other sentences, a set of pictures different from the pictures used in the therapy were used at baseline. These items included twelve irregular verbs, of which he had received training on six, and ten regular verbs, of which he had been trained on four. All verbs were highly imageable, and were matched in frequency for the past tense verbs assessed. Regulars and irregular verbs were matched for frequency range, though the mean frequency of irregular verbs was slightly higher for the irregular set (irregular: 45.8; regular: 39.5). Assessment of verbal and written past tense verb morphology, isolated from sentence construction, was possible using a sentence completion task: ‘Today the boy is riding the bicycle, tomorrow the boy _____’. The treatment produced significant on the untreated picture description task in the production of sentences, root verbs, and past tense forms. Past tense verb production in the untreated sentence production task showed treatment gains on the treated items, while performance on both regular and irregular untreated past tense forms remained poor. The sentence completion task targeting past-tense morphology (today the boy is riding the bicycle, yesterday the boy _____) produced improvements that generalised only to untreated items in the written modality for regular verbs. The finding that EA showed generalised improvement on regular, but not irregular past-tense verbs is interpreted by the authors as support for the dual route perspective of past tense verb processing, as regulars are possible using rule application that operates irrespective of lexical knowledge, while irregular past tense forms are stored separately as complete items in the lexical store. However, the authors did not categorise untreated irregular verbs in terms of those taking phonologically similar versus dissimilar transformations. It was anticipated that generalised improvement to untreated irregulars might be achieved through treating items with similar present to past verb phonology.

The rehabilitation procedure

The current study used a treatment programme targeted at syntactic processing, and tested for generalisation across different types of syntactic operations, and across the expression of regular and irregular verb types.

No rehabilitation studies to my knowledge have explored how irregular past tense verbs behave in re-learning programmes. The current study aimed to promote past-tense verb production in a non-fluent aphasic patient, DS, using a thematic query approach requiring the description of pictures. This procedure targeted the past tense inflection by requesting nonlinear sentence structures, but as in Weinrich et al. (1999) the method did not involve any explicit training of syntax rules. As a consequence, any treatment gains produced in past tense verb inflection reflect either stabilised functional representations or strengthened connections between functional and positional representations, and not any superficial trained application of a regular /d/, /t/ or /^d/ affix.

The procedure in the current study used pictures that were made available to DS throughout the training period. The pictures depicted different types of sentences (see materials section), and involved both regular and irregular verbs. In the picture-description tasks, DS was required to describe what was happening in the picture beginning with each noun class (e.g. boy [subject], and dog [object]). Throughout the rehabilitation sessions, this same procedure was used, except that DS also answered two thematic questions about the picture (e.g. 'what is the role of the boy [/dog]'). A timetable of the complete treatment procedure is provided in Table 1.1.

Five baseline tests were conducted, involving treated and untreated sentence production, past tense verb production, and PALPA comprehension assessments (see Table 1.1 for details). Untreated same group irregular words were taken from the same groups as treated irregular words (groups based on McClelland and Patterson, 2002a, e.g. sleep-slept, keep-kept), while the untreated different-group irregular words belonged to different groups from those irregulars included in the treatment. Both sets were tested at baselines in order to test for generalised improvement based either on the verb form (generalisation to same-group items) or more general recovery (generalisation to both the same and to different groups). Improvement was expected in the production of the treated past tense verbs. Existing research predicts generalised improvement to untreated regular verbs following treatment of a set of regular verbs (Weinrich, Boser, and McCall, 1999). Generalisation to the production of untreated sentences and to the comprehension of non-linear sentences was anticipated, given that improved ability to link noun phrases with their thematic roles should be applicable to both production and comprehension tasks.

Table 1.1. Timetable of the rehabilitation procedure

Type of session	**Stimuli**	**Assessment**	**Duration**
	Verb meaning task	Comprehension of verbs to be used in the treatment was tested.	1 one-hour session
Pre-therapy Baseline 1	Sentence production assessment-treated	'To be treated' sentences and verbs (regular and irregular)	2 one-hour weekly sessions
Pre-therapy Baseline 2	Sentence production assessment-untreated	Untreated sentences and verbs (regular and irregular)	2 one-hour weekly sessions
Pre-therapy Baseline 3	Sentence production assessment-nonwords	Untreated sentences and nonwords	2 one-hour weekly sessions
Pre-therapy Baseline 4	Past tense verb production	'To be treated' verbs (regular and irregular), untreated verbs (regular and irregular), and nonwords.	1 one-hour session
Pre-therapy Baseline 5	PALPA tests (Kay, Lesser and Coltheart, 1992)	PALPA, including comprehension assessments.	4 one-hour weekly sessions
Rehabilitation sessions	Procedure requiring DS to 1. Produce linear and nonlinear sentences to describe target pictures, 2. Answer two thematic questions about the picture and 3. Produce the linear and nonlinear sentences again.	Treated sentences and verbs (regular and irregular)	10 1-hour weekly sessions
Post-therapy Baseline 1	Sentence production assessment-treated	Treated sentences and verbs (regular and irregular)	2 one-hour weekly sessions
Post-therapy Baseline 2	Sentence production assessment-untreated	Untreated sentences and verbs (regular and irregular)	2 one-hour weekly sessions
Post-therapy Baseline 3	Sentence production assessment-nonwords	Untreated sentences and nonwords	2 one-hour weekly sessions
Post-therapy Baseline 4	Past tense verb production	Treated verbs (regular and irregular), untreated verbs (regular and irregular), and nonwords.	1 one-hour session
Post-therapy Baseline 5	PALPA tests (Kay, Lesser and Coltheart, 1992)	PALPA, including comprehension assessments.	4 one-hour weekly sessions
6-month follow-up (Baseline 1)	Sentence production assessment-treated	Treated sentences and verbs (regular and irregular)	2 one-hour weekly sessions

Method

Case Description

DS was a 73-year-old patient presenting with non-fluent aphasia subsequent to a stroke affecting the left inferior frontal cortex. Prior to his stroke DS worked as a bus inspector and his education history included a non-university diploma. A structural MRI scan was taken, and the lesion was created in SPM and added as an overlay onto a standard multi-slice template in MRIcron. The SPM analysis was a one sample t-test with the covariates healthy (140 brains aged 40+) vs. patient, age and gender. Any region showing a significant difference at $p < .01$ between DS's image and the control overlay was judged as a lesion. The analysis revealed a lesion in the left inferior frontal gyrus, involving the pars opercularis, pars triangularis and pars orbitalis, left rolandic operculum, left insula, left middle frontal gyrus, left precentral gyrus, left postcentral gyrus, left caudate and putamen (Figure 1.1).

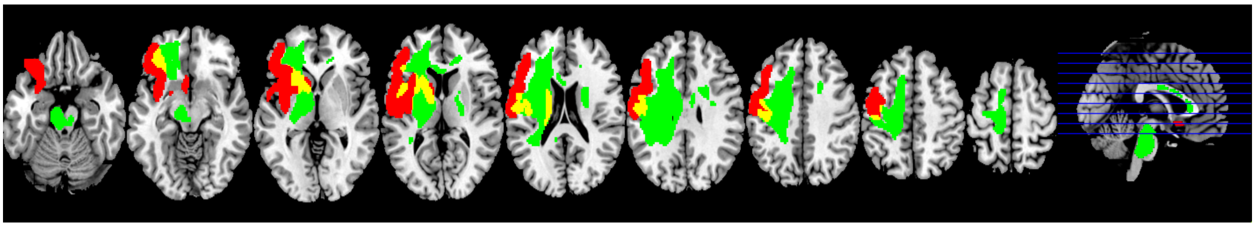


Figure 1.1. Brain images of patient DS: Corrected FWE p =0.01 Extent threshold including only significant blobs containing ≥100 voxels. N.B. Grey matter lesion appears in red and white matter lesion in green. The lesion was created in SPM and added as an overlay onto a standard multi-slice template in MRIcron. The SPM analysis was a one sample t- test with the covariates healthy (140 brains aged 40+) vs. patient, age and gender.

DS's free speech was hesitant and effortful, and featured frequent omissions, most notably of inflection and function words. DS's baseline score on the PALPA 55 Auditory Sentence-Picture Matching (Kay, Coltheart and Lesser, 1992) was impaired at 37/60, and his performance showed an effect of sentence reversibility, as his most frequent error was in selecting the distracter item portraying the reverse thematic role relationship (9/11 errors). This trend towards poor judgements with reversible sentences was echoed in the written version (PALPA 56 Written Sentence-Picture Matching) where DS's score was impaired at 24/60 and most errors were reverse errors (12/14 errors). This profile is in the direction predicted by Grodzinsky's Trace-Deletion Hypothesis account of the agrammatic comprehension impairment (TDH; Grodzinsky 1986, 1995). Performance on the PALPA 12 sentence repetition task was impaired, at 27/36. He showed a profile in the production of root verb - past tense verb transformations in which regular forms were worse than irregular transformations: Regular 17/57, irregular 27/54 (in non-phonologically matched materials, X^2 = 26.917, p =.001, see Table 1.2 for stimuli details). In this task, the experimenter presented DS with root verbs, and required him to express each verb 'as if it happened yesterday': DS demonstrated good phonological skills, scoring 24/24 on a test of word repetition with stimuli comprised of up to 3-syllables (PALPA 7, Kay et al., 1992), though repetition may not be a sensitive test of phonological deficits (Crisp and Lambon Ralph, 2006).

Set	**N**		**Celex frequency**	**Log frequency**	**Familiarity**	**Letter length**	**Imageability**
Regular	57	Present	176.42	1.96	486.36	4.24	121.86
		Past	149.25	2.09	39.82	6.63	49.3
Irregular	54	Present	196.93	2.06	474.84	4.28	134.63
		Past	159.43	1.96	324.52	4.28	53.86

Table 1.2. Mean scores on lexical variables for the verbs in the sentence completion task

Materials

Treated set

The materials consisted of 38 black and white simple hand-drawn pictures that portrayed sentences involving characters, animals, and objects. All verbs were presented as pictures, and all pictures portrayed reversible sentences. Below each picture, the target verb in its root form was written (Figure 1.2). All sentences targeted a transitive, accusative sentence structure (sentence group 1 and 5 in Appendix A); though there were occasions where DS produced sentence structures that deviated from this sentence frame, particularly in the nonword condition. Guidelines for scoring syntax across various possible sentence structures are provided in Appendix A.

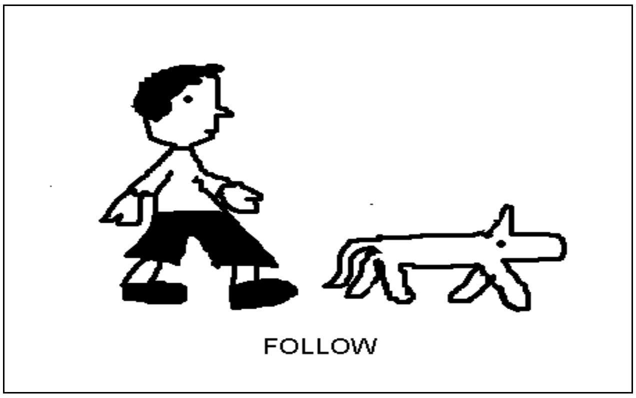


Figure 1.2: Example of test stimuli for the target item 'The boy is following the dog'/ 'The dog is being followed by the boy'

12 pictures portrayed sentences involving irregular verbs, and 26 contained regular verbs. Fewer irregulars than regulars were used due to the limited number of irregular 'neighbourhoods' and the restricted size of each group. The regular and irregular verbs were matched as far as possible on frequency, familiarity, letter length, and imageability (Table 1.3).

Set	**N**		**Celex frequency**	**Log frequency**	**Familiarity**	**Letter length**	**Imageability**
Regular	26	Present	69.31	1.84	506.27	4.5	199.27
		Past	48.10	1.53	-1	6.38	-1
Irregular	12	Present	69.99	1.67	492.91	4.25	221.66
		Past	81.24	1.70	245.16	4.58	166.5

Table 1.3. Mean scores on lexical variables for the verbs in the treated set

Untreated sets

Untreated sets of regular and irregular verbs (with same and different affixes to those in the treated groups) were used (see Table 1.4 for lexical score details). Untreated same group irregulars shared the same phonological transformations as treated irregulars (i.e., they rhymed, e.g. sleep-slept, keep-kept, groups based on McClelland and Patterson, 2002a), and untreated different-group irregulars belonged to different groups from those irregulars included in the treatment (e.g. drink-drunk – build-built). Both sets were tested at each baseline phase. See Appendix B for the sets of irregulars with treated and untreated affixes.

Set	**N**		**Celex frequency**	**Log frequency**	**Familiarity**	**Letter length**	**Imageability**
Regular	57	Present	312.02	1.45	379.62	5.36	110.22
		Past	381.14	1.21	296.67	6.30	85.41
Irregular (untreated affix)	27	Present	329.02	1.83	393.59	4.15	126.22
		Past	372.40	1.69	288.67	4.30	87.41
Irregular (treated affix)	23	Present	301.83	1.87	443.91	4.26	118.91
		Past	291.81	1.75	366.52	4.43	74.17

Table 1.4. Mean scores on lexical variables for the verbs used in the untreated sentence completion task (Baseline 4).

Nonwords were based on the regular and irregular sets, and were generated by making a letter change at random and equally distributed positions over the words. Therefore, the nonwords that appeared underneath pictures in place of the real verbs in the untreated picture description task were matched for length and CV structure.

Procedure

Throughout the rehabilitation sessions, each picture was presented with the written corresponding verb in its root form appearing below (see Figure 1.1). For each treated item, DS was asked to describe what was happening in the picture starting with the boy (subject) and the dog (object). Two thematic questions followed: 'Tell me about the role of the boy [subject]' and 'Tell me about the role of the dog [object]'. Finally, DS was asked again to describe what was happening in the picture starting with the boy (subject) and the dog (object). Requested subject / object order (canonical / non-canonical target structures) was alternated. The nonlinear form of one treated item ('the girl hides the box'; 'the box is hidden by the girl') targeted the past participle verb (hidden) rather than the simple past tense form (hid). For this item (along with any untreated structures that included a past participle verb), the syntax in the nonlinear response was analysed as usual, but in addition, DS was provided with the first part of the sentence 'Yesterday' and asked to produce another sentence starting with the subject (e.g. boy), the target sentence being 'Yesterday, the boy hid the box'. DS's response to that question was assessed for verb production.

If DS still made errors in sentence production after responding to the thematic questions, the following feedback was given:

For errors in producing the linear form twice, e.g. 'the boy is following by the dog'; 'the dog is following by the boy', the examiner advised 'so the boy is following and the dog is following? In the picture only one of the characters is followed. Would you like to try again?'.

In terms of inflection errors, it was decided that where over-regularisation errors (e.g. catched), mixed errors (e.g. tored), or *irregularisation* errors (e.g. patch-paught, similar to catch-caught) occurred, the correct past tense form would be provided aurally, and DS would be asked to produce the sentence again. Errors in word choices would be corrected in the same way. Other syntax errors (besides linear - nonlinear syntactical frame substitutions) were not corrected.

The assessment (baseline) sessions included a verb meaning test, where the examiner read out each verb to be used in the treatment, and DS was instructed to describe what the word meant. DS performed well on the task and correctly paraphrased an approximate meaning for every verb in the set. The sentence production assessments for the treated and untreated pictures / verbs, and the (untreated) nonwords (Baselines 1, 2 and 3) were conducted in the same way as is detailed in the rehabilitation procedure, only without feedback from the examiner (see above). The nonword assessment used different pictures from both the treated and untreated materials, with nonwords that differed from the verb set by one consonant. The past tense verb production assessment required DS to say each verb stem provided verbally by the examiner 'as if it happened yesterday'.

See Table 1.1 for a full timetable of the procedure.

Syntax analyses

To score responses a point-scoring system was used, loosely based on Caplan and Hanna (1998, see Appendix A for the full guidelines). Broadly, correct application of a linear syntactical frame was awarded for the production of 'the' and 'is' in 'The boy is following the dog'. For correct nonlinear structures, the production of 'is being' and 'by the' was required (e.g. 'The dog is being followed by the boy'. A correct score was only given when correct different syntactical frames were applied across canonical and non-canonical conditions for each item. The order of canonical vs. non-canonical requests was randomised across sessions, but they were always posed in pairs.

Results and Discussion

The study aimed to assess the effectiveness of the mapping therapy in improving DS's production of nonlinear sentences and past tense verbs, and to test whether there was any generalisation of improvement. The data are discussed with reference to existing rehabilitation studies and models of past tense processing. Full pre- vs. post- therapy raw and percentage scores are reported in Table 1.5.

Table 1.5: Scores across tasks, pre- and post-treatment

Baseline	**Description**	**Pre-therapy scores**	**Post-therapy scores**
Pre-therapy Baseline 1	Sentence production assessment-treated	Syntax analyses 5/38 (13.16%) Regular past tense production 7/26 (26.9%) Irregular past tense production 3/12 (25%)	Syntax analyses 35/38 (97.22) Regular past tense production 26/26 (100%) Irregular past tense production 11/12 (91.66%)
Pre-therapy Baseline 2	Sentence production assessment-untreated	Syntax analyses 9/160 (5.63%) Regular past tense production 44/82 (53.65%) Irregular past tense production 39/78 (50%)	Syntax analyses 150/160 (93.75%) Regular past tense production 66/82 (80.48%) Irregular past tense production 60/78 (76.92%)
Pre-therapy Baseline 3	Sentence production assessment-nonwords	Syntax analyses 0/40 (0%) Nonword – regular inflection score 4/40 (10%)	Syntax analyses 35/40 (87.5%) Nonword - regular inflection score 13/40 (32.5%)
Pre-therapy Baseline 4	Untreated past tense verb production	Regular past tense production 17/57 (29.82%) Irregular past tense production 25/50 (50%) Irregular untreated predictable affix 16/27 (59.25%) Irregular treated predictable affix 9/23 (39.13%)	Regular past tense production 33/57 (57.89%) Irregular past tense production 33/50 (66.66%) Irregular untreated predictable affix 15/27 (55.55%) Irregular treated predictable affix 18/23 (78.26%)
Pre-therapy Baseline 5	PALPA tests (Kay, Lesser and Coltheart, 1992)	Auditory sentence-picture matching 36/60 (60%) Written sentence-picture matching 24/60 (40%) Sentence repetition 27/36 (75%)	Auditory sentence-picture matching 44/60 (73.33%) Written sentence-picture matching 30/60 (50%) Sentence repetition 30/36 (83.33%)

Sentence production

Analyses of syntax revealed a significant improvement in applying the correct syntax in both linear ('the boy **is** following the dog') and non-linear sentences ('the dog is **being** followed **by the** boy'). Syntax scores at Baseline 1 were 5/38 pre-therapy, and 35/38 post-therapy (McNemar $X^2 = 28.03$, $p < .001$). This effect was enduring, withstanding a 6-month period without therapy (5/38 pre-therapy vs. 36/38 at 6-month follow-up, McNemar $X^2 = 29.032$, $p < .001$). The improvement in the production of syntactical frames seen on the treated items in Baseline 1 was generalised to untreated sentences at Baseline 2 (9/160 vs. 150/160, McNemar $X^2 = 139.01$ $p < .001$, Figure 1.3).

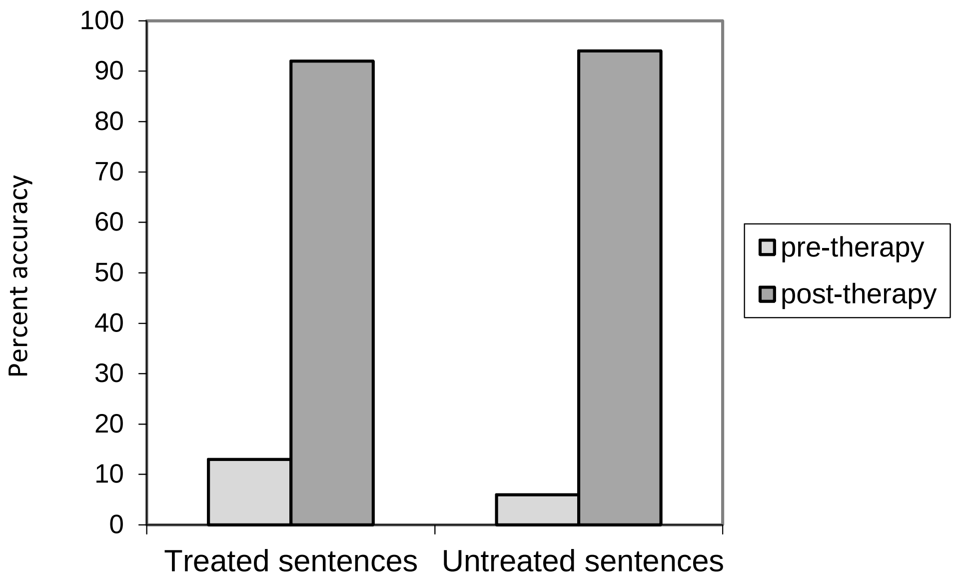


Figure 1.3: Data showing improvement in the application of distinct syntactical frame for linear and nonlinear sentences, across treated (Baseline 1) and untreated (Baseline 2) sets

The improvement seen on treated items demonstrates positive effects from (i) posing questions about the thematic roles in sentences, and (ii) describing that it was incorrect to use the same verb and syntactical frame to report an action sentence describing an agent – 'experiencer' relationship. This occurred even though the correct syntactical frame was never presented. That this improvement generalised to untreated sentences indicates that improvement was not specific to the treated sentences, suggesting too that the training might be beneficial for DS's communication in everyday life. While this assumption was not tested directly (e.g. by analysing free speech data pre and post- rehabilitation), DS self-reported gains in speech production.

Sentence comprehension

More modest gains were produced in comprehension. DS's performance on the PALPA auditory sentence-picture matching task improved from 36/60 pre-therapy to 44/60 post-therapy (McNemar $p < .005$). The suggestion that the treatment improved DS's thematic role understanding is supported by the significant reduction in the number of reverse errors made on reversible sentences in the PALPA set post-therapy (81.8% vs. 7/16 43.7%, Friedman's $X^2(1) = 5.0$ $p < .05$; Wilcoxon post hoc showing a decline in reverse errors: $T = -2.236$, $p = .025$). The results from prior reports of sentence production therapies differ in their findings of generalisation across production and comprehension (no improvement in comprehension abilities despite gains in production were reported in Rochon and Reichman, 2003 and Rochon et al, 2005; in contrast Jacobs and Thompson (2000) report generalisation in production and comprehension following comprehension-only treatment but not production-only treatment). The improvement produced here might be explicable by the focus of the mapping therapy approach on thematic role understanding, which might bring about improvement in the comprehension of reversible sentences.

Regular and irregular past-tense verbs

Following therapy, the production of the treated regular and irregular past-tense verb forms in sentences (using the picture description procedure at Baseline 1) improved significantly (pre-therapy 10/38 vs. post-therapy 38/38, McNemar $X^2 = 26.03$, $p < .001$). The most frequent pre-therapy error in verb production during the picture-description task was due to a failure to attempt the non-linear form (e.g. ‘the dog is following the boy’ was produced for the target ‘the dog is followed by the boy’). Therefore, failures in verb production at Baselines 1, 2, and 3 may not reflect an inability to generate the past tense form. The assessment at Baseline 4, requiring the conversion of root verbs to their past-tense forms (change the word as if it happened yesterday) allowed past-tense verb production to be assessed independently of failures to produce non-linear sentences. Before therapy, DS often failed to inflect regular and irregular verbs, though some over-regularisation errors (e.g. catched) were made on irregular forms. Data describing the production of regular and irregular past tense verbs across phases of the treatment are discussed separately (see Appendix C for full details of McNemar tests across sessions).

Past-tense verb production

Performance on regular past tense verbs in nonlinear sentences at Baseline 1 improved from 7/26 pre-therapy to 26/26 post-therapy (McNemar $p < .001$). This improvement was maintained at a six-month follow-up (regular verbs 7/26 pre-therapy vs. 26/26 at six-month follow-up, McNemar $p < .001$, Figure 1.4).

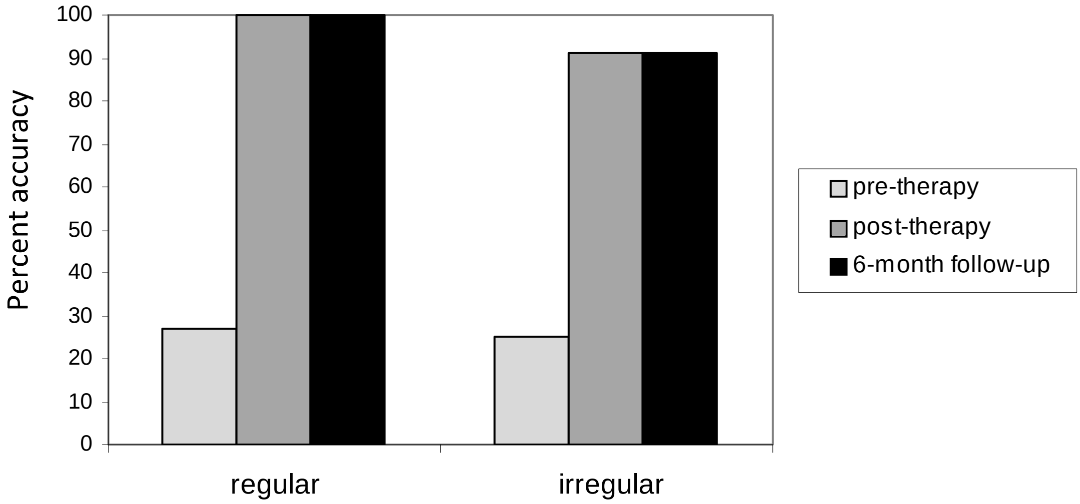


Figure 1.4: Accuracy of the production of treated past-tense verbs in sentences, pre- and post-therapy, and at a 6-month follow-up (Baseline 1).

A generalised improvement was also found in generating the regular past-tense inflection in untreated sentences comparing untreated regular verbs at Baseline 2 (44/82 vs. 66/82, McNemar, $p < .001$). The production of untreated regular past tense forms in a task requiring the conversion of root verbs (Baseline 4) was also significantly improved (17/57 versus 33/57 McNemar $p = .001$, Figure 1.4). This result suggests that improvements in past tense verb production were not specific to the picture-description task used in the therapy.

Performance on irregular past tense verbs in nonlinear sentences improved from 3/12 pre-therapy to 11/12 post-therapy (Baseline 1, McNemar $p < .01$) and 3/12 vs. 11/12 for irregular verbs at a six-month follow up (McNemar $p < .01$, Figure 1.4). Following improvements in the production of non-linear sentences, treatment gains were also produced in untreated irregulars at Baseline 2 (39/78 versus 60/78, McNemar $p < .01$, Figure 1.5).

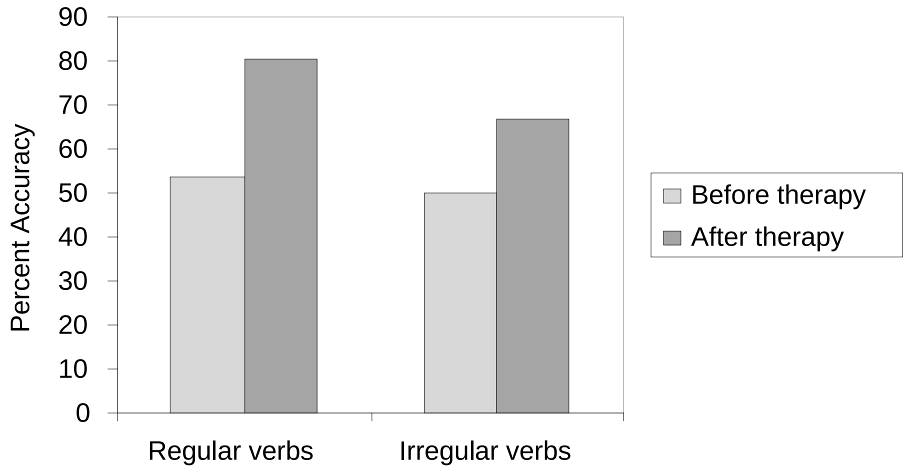


Figure 1.5: Generalised improvement in untreated regular and irregular forms (% accuracy)

Nonword data

The past-tense conversion of non-words (where only regularly-inflected non-words were scored as correct, as '–ed' should be the default affixation for novel forms), revealed significant improvements at Baseline 3 (4/38 vs. 13/38, McNemar $p < .005$, see the section on 'error analyses' for details on 'irregularisations' produced in the nonword tasks).

Irregular verbs

Importantly, an improvement was seen in the production of untreated irregular items at Baseline 4 (requiring expression of root verbs 'as if they happened yesterday'). This result shows that improvement was not explicable in terms of an increased ability to generate nonlinear sentences, as an improvement was seen on generating past tense verbs on a simple elicitation task (25/50 pre-therapy versus 33/50 post-therapy, McNemar $p = .001$).

A generalised improvement from treated to untreated irregular past tense forms is not consistent with the dual-mechanism approach (e.g. Marslen-Wilson and Tyler, 2007). The dual-mechanism account proposes that irregular past tense forms are possible through a passive, rote-learned lexical store and so predicts that treatment of regular and irregular past tense forms should produce generalisation to regular items only. Moreover, the finding of generalised improvement in regular past tense verb production seen in DS following a treatment that did not explicitly train syntax is not consistent with their assertion that a morpho-phonological parsing mechanism is damaged in nonfluent aphasics.

The current intervention also showed that inflection scores for irregular verbs with the same affixes as those in the treated irregular group were significantly better post- relative to pre-therapy, whereas performance on items from affix groups left untreated did not change (performance on irregulars with affixes left untreated was 16/27 pre-therapy vs. 15/27 post-therapy, McNemar $p > .05$; where irregulars with treated affixes was 9/23 pre-therapy vs. 18/23 post-therapy, McNemar $p < .05$, Figure 1.6).

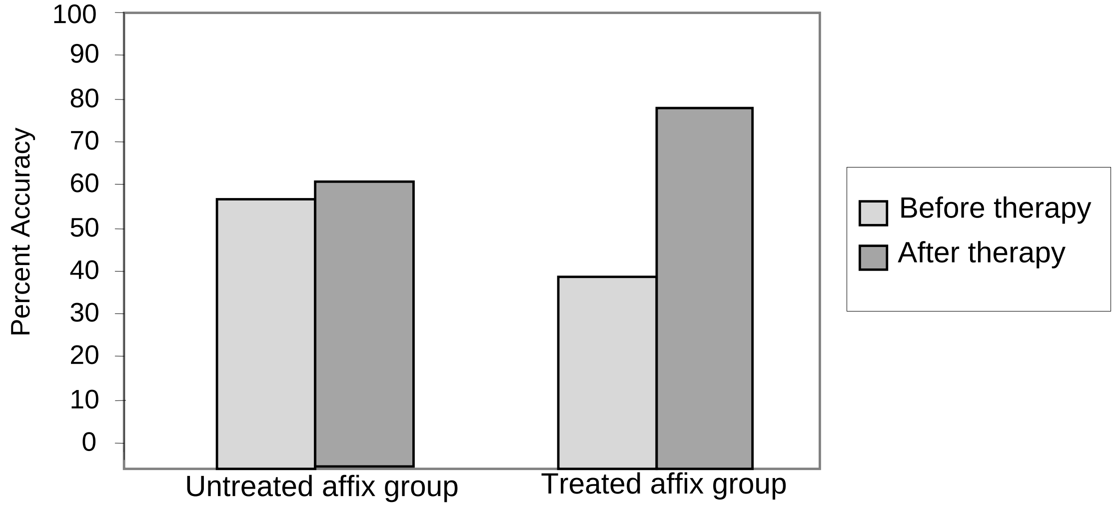


Figure 1.6: Generalised improvement in untreated irregular forms with the same predictable affix as irregular forms in the treated set, % accuracy (Baseline 4).

For the untreated set, the pre-therapy baseline for untreated irregular affixes was higher than those irregulars whose affixes were subsequently treated. This is because it was not possible to match on the range of lexical variables (Table 1.4) as well as keep the irregular groups intact within the two sets. Therefore, it was decided that the set where DS gave a better performance should be left untreated, so that any positive effect observed as a function of the therapy could not be attributed to a better ability to inflect these irregulars naturally, before therapy.

The current Chapter described the sentence and past tense verb production performance of DS, a nonfluent aphasic patient presenting with poorer production of regular than irregular past tense verbs. The treatment produced improvement in sentence and past tense verb production. Treatment gains were evident for treated regular and irregular past tense forms, and they generalised to both untreated regular and irregular verbs. There were also generalised treatment gains in sentence comprehension. The finding of improvement in sentence production and comprehension supports the use of thematic mapping therapy in treating sentence processing deficits in nonfluent aphasic patients. Perhaps the most important finding was that past tense verb production treatment generalised to untreated items, and this pattern appeared to be modulated via irregular verb ‘neighbourhoods’.

Generalised improvement was produced on untreated irregular verbs from the same groups as treated irregular items (sleep-slept; weep-wept), but not on the different-group untreated irregular set. Broadly, there was generalisation within subgroups of irregular verbs but not across all irregular items, which could indicate non-specific improvement. This evidence is consistent with the single-mechanism view that by strengthening connections between treated root verbs and their past tense forms, improvement can generalise to other irregulars from the same groups as treated items. However, given that many groups of irregulars are similar in their morpho-phonological structure, the pattern of improvement could also be explained in terms of the Words and Rules framework. The Words and Rules theory might account for improvement to only same-group untreated irregulars as gains on trained irregulars generalised to untreated items with a similar morpho-phonological structure, because the lexicon is described as partly associative, and able to generalise to other items in this way. However, the findings are inconsistent with a model positing isolated lexical items (Marslen-Wilson and Tyler, 2007; Tyler et al., 2002, 2005a and 2005b).

The improvements on untreated irregulars seen in the rehabilitation study here were not produced by another study with an expressive aphasic patient treating the production of verbs in sentences (EA, Weinrich et al, 1999). Like DS, EA presented as a nonfluent aphasic patient. However, EA’s aphasia appeared to be more severe than DS’s: where EA’s speech was limited to some stereotyped phrases, DS’s spontaneous speech was more varied, and he rarely used writing to respond to questions, unlike EA. Weinrich et al.’s study produced generalisation to written regular but not irregular verbs, evidence they interpret as supporting a dual mechanism account. Their paper does not include details about the treated and untreated irregulars, nor does their analysis separate untreated irregulars unto those with treated and untreated affix transformations; therefore it is difficult to make comparisons between the current data set and the data in Weinrich et al. (1999). Differences between the study findings might be explained, for instance, by the sets of irregular verbs Weinrich et al. (1999) used in the treated and untreated irregular sets: If irregulars in the untreated irregular set were from different groups to the treated items, then no generalisation would be expected, even within the single mechanism account.

Error types

Post therapy, DS made frequent over-regularisations on irregular verbs. This again suggests that treating a set of regular past tense verbs generalises to other forms. The result also implies that there may be cases where the regular affix is more competitive than certain irregular forms. The explanation for this could be in the weighting of the treatment items, in that there were more regular verbs than irregular verbs in the treated set (12 irregular and 26 regular – due to DS's more pronounced difficulty with regular verbs) here.

Interestingly, these were also instances of 'mixed' errors observed in the production of irregular past tense forms throughout testing (e.g. 'losted'). Mixed errors were first noted and discussed by Patterson, Lambon Ralph, Hodges and McClelland (2001). Patterson et al. note that such errors are not predicted by the dual-mechanism account, which proposes that verbs are subject either to a regular root verb + '*ed*' operation, or retrieval of an irregular whole past tense form. The Words and

Rules theory (Pinker 1999, Pinker and Ullman 2002) accounts for instances where the 'ed' affix is applied to irregular root verbs (e.g. 'catched'), in that if a whole irregular past tense form is not found, an affix is then applied. It is also possible that dual mechanism accounts (e.g. Baayen et al., 1997) could explain the current findings in terms of increased productivity of applying the treated suffixes via training.

However, it is not clear how these accounts can explain the finding of mixed past tense irregular forms with 'ed' affixes (e.g. 'tored'), as the theory proposes that successful retrieval of a whole irregular prevents the application of the regular +'ed' rule. Perhaps there are some cases where an irregular past tense verb is retrieved from the lexicon, and that past tense form is entered into the morpho-phonological parsing process, but this explanation has not been made explicit in their model.

In the single-mechanism PDP framework, the production of past tense verbs is modulated by frequency: High frequency irregular verbs are accessed quickly and are not affected by competition from the regular past tense inflection. In contrast, lower frequency past tenses are accessed less quickly, and may be subject to some interference from the regular inflection resulting in mixed errors such as '*tored*'. It follows that a single mechanism model would also predict instances where high-frequency irregular past transformations influence low frequency regulars and nonwords. Performance on a nonword task, which required DS to change nonwords from root to past (Baseline 4) contained responses similar to irregular past-tense forms (e.g. 'surnt' was produced for the nonword 'surn', which is similar to burn-burnt). There were also instances where DS did not change the nonword in its root form when the past-tense was required (e.g. 'chid'). These no-change responses occurred more frequently when the nonword had a terminal consonant of 'd' or 't', as with real no-change irregulars. These findings support both (i) the single-mechanism notion that irregular past tense transformations are processed in much the same way as regulars, producing 'irregularisations' on similar nonwords, and (ii) the Words and Rules account that learned irregulars generalise not only to untreated morpho-phonologically alike irregulars, but also to nonwords with similar morpho-phonological structure.

Suggestions for future research

The current study suggests that both regular and irregular past tenses are generated by a single mechanism involving connections between root and past tense verbs. The likelihood of correct generation of past tense forms is modulated by frequency, neighbourly support and semantics. However, the single-mechanism view that the production of irregular past tenses is modulated by frequency was not tested directly here. Patterson et al. (2001) assessed the performance of patients with semantic dementia on irregular past tense verbs and found performance on irregulars to be highly predicted by performance on a synonym judgement task. It would be interesting to test the single-mechanism claim that processing past tense irregular verbs is modulated by word frequency in patients with semantic deficits and problems with past tense irregular verbs. A single-mechanism perspective predicts that more errors, and specifically more 'mixed' errors (tear-'tored') will occur in lower-frequency irregulars less assisted by neighbourly support. Their assumption is that these forms are likely to be subject to interference from the regular –ed affix. An investigation into the factors mediating past tense verb production in a patient with degraded semantic information is detailed in Chapter 3.

Conclusions

The rehabilitation method for treating non-linear sentence production used in the current study produced a significant improvement for applying correct syntactical frames for nonlinear sentences: Post therapy, DS was able to use language to specify the reverse thematic relationship (e.g. using 'being' and 'by the', rather than 'is'), which generalised to untreated sentences. Significant improvement was also seen in the production of treated and untreated past tense regular verbs, and in a task requiring the inflection of nonwords. Improvement on the treated irregular verbs was found, challenging the dual-mechanism claim that irregular past tenses need to be encoded individually into a lexical store, but this improvement generalised only to irregulars with the same affix as treated irregulars (e.g. sleep-slept, keep-kept). Broadly the main findings can be accommodated by both the single mechanism account and the Words and Rules theory: A single-mechanism account might interpret the current findings as evidence that irregular past tense verbs are processed in much the same way as regulars, and a Words and Rules account would explain that learned irregulars generalise to same-group irregulars as they have similar morpho-phonological structure. However, DS made a number of 'mixed' errors in verb production (e.g. 'losted'), a result that poses some difficulty for the Words and Rules Theory (e.g. Pinker and Ullman, 2002).

The data suggest that irregular verb rules are coded in a generalised form (see Albright and Hayes, 2001). Taken together, the current findings support the use of a mapping therapy targeting gains in sentence production, and the generation of past tense forms by promoting understanding of thematic roles in patients presenting with non-fluent aphasia. The results on past tense verb production are interesting from a theoretical standpoint, and promising from a rehabilitation perspective, given the pattern of generalised improvement.

Chapter 3: Exploring the impaired production of irregular past tense verbs in a patient with a semantic deficit

Abstract

A key notion within a single mechanism account of verb processing is that the production of past tense verb forms is mediated by frequency, neighbourly support and semantic knowledge (Lambon Ralph, Braber, McClelland, and Patterson 2005; McClelland and Patterson 2002). The current study assessed the relative contributions of frequency, neighbourly support and semantic knowledge to the past tense verb production of JF, a primary progressive aphasic patient with impaired semantic knowledge. When tested on a series of production tasks, JF's past tense verb production was predicted by the degree of regularity, frequency, irregular phonological 'neighbourhood' size, and integrity of semantic knowledge about the to-be-produced verbs. The data suggest that the production of the irregular past tense is mediated by frequency, neighbourhood support and semantics in this case, as is predicted by a single mechanism account of verb production.

Introduction

A main focus in the literature has concerned whether two distinct processes are necessary to explain our production of regular and irregular past-tense verbs. Irregular past tense verbs (e.g. catch-caught) do not follow a common transformation / rule whereas regular past tense forms (e.g. follow-followed) always take a /d/, /t/ or /ed/ inflection, depending the stem-final consonant. Neuropsychological studies have played an important role in this debate based on dissociations in regular and irregular verb production in different patients. Some non-fluent aphasic patients present with better performance with irregular compared with regular past-tense verbs (Marslen-Wilson and Tyler 1997, Ullman, Corkin, Coppola, Hickok 1997), while some patients with semantic difficulties show the reverse dissociation (Patterson, Lambon Ralph, Hodges and McClelland, 2001).

These dissociations have been interpreted differently by single and dual-mechanism approaches to verb production. Dual mechanism theorists (Marslen-Wilson and Tyler 2007; Tyler, Randall, and Marslen-Wilson 2002; Tyler, Stamatakis, and Marslen-Wilson 2005a) argue that the deficit for irregular over regular past tense verb production dissociation reflects, on the one hand, impaired retrieval of irregular verbs and, on the other, preserved rule-based parsing process that connects the verb stem to the inflection for regular past tense verbs. The dual mechanism account proposes that regular and irregular verb production is supported by distinct functional and neural mechanisms.

In contrast to this, a single-mechanism approach holds that all verbs are produced by a single retrieval process that maps phonological representations of present tense forms to phonological representations of past tense forms. This process is influenced by a variety of factors, including lexical frequency, knowledge of word meaning, the strength of the phonemic representations of verbs (determining the strength of the present tense form), the number of words sharing the same mapping and the complexity of the phonological representations themselves. According to this account, poorer production of irregular relative to regular past tenses will arise when semantic knowledge is impaired, when retrieval may reflect the relative strength of present-past tense mappings more than lexical input. In contrast, worse production of regular than irregular forms can arise in patients sensitive to the phonological complexity of the stimuli (regular forms being more complex, Burzio, 2002, Desai et al., 2006). This account also holds that the degree to which semantic input mediates verb retrieval will depend on factors such as lexical frequency of the verb form – with high frequency forms less dependent on semantic knowledge than low frequency items (Patterson, Lambon Ralph, Hodges, and McClelland, 2001). There are other clear lines of evidence for broader impact of semantic in atpical mappings including past tense (for instance, in finding that applying Transcranial Magnetic Stimulation to the Anterior Temporal Lobe had a selective impact on normal participants' ability to generate the past tense of irregular verbs: Holland and Lambon Ralph, 2010).

Quasi-regularity in the English irregular past tense

Many irregular verbs contain aspects of the regular /d/, /t/ or /^d/ affix (e.g. keep=kept, sleep=slept, creep=crept). Patterson et al. (2001) argue that irregulars involving a terminal consonant change, or a combination of a vowel and end-consonant change (e.g. 'tell-told') may represent an 'intermediate case' between regulars and vowel change irregulars, where past tense conversions are less uniform. In contrast, a dual mechanism account holds that both quasi-regular and truly irregular verb forms are retrieved through a lexical production process. These contrasting positions make it interesting to study quasi-regular forms in patients with problems in retrieving irregular verbs.

Neighbourhood support in sets of irregular past tense verbs

English irregular verbs fall into ‘neighbourhoods’ of phonologically similar forms (e.g. dive-dove, drive-drove). A number of studies have found that neighbourhood size predicts accuracy of producing irregular past tense verbs for both children and adults (Bybee and Slobin, 1982, Bybee and Moder, 1983, and Stemberger and MacWhinney, 1986). McClelland and Patterson (2002a) outline nine groups of irregular past tense verbs, each comprised of verbs that share the same phonological relationship between present and past tenses. Each group differs in terms of neighbourhood size and the degree of within-group phonological similarity. McClelland and Patterson propose that around 59% of irregular verbs can be allocated into eight groups of ‘quasi-regular’ irregular past tense verbs, and assert that the vast majority of remaining irregulars can be classed as belonging to one group (9) that take a vowel change rather than the addition of a /d/ or /t/ terminal consonant. Here the production of verbs with high versus low numbers of phonologically similar present-past tense pairings was assessed (e.g. sleep-slept, keep-kept). By considering the degree of regularity and the size of these phonologically similar groups separately, the aim was to isolate the contributions of quasi-regularity from neighbourly support on past tense production in patient JF.

Stemberger (2004a) tested normal participants’ production of irregular past tense verbs in a primed sentence production experiment. The sentences comprised of ‘The – Noun – Verb’ structures, where the preceding noun primed the phonology of either the present tense verb vowel (the ball fall), the past tense verb vowel (the bell fall), or an unrelated vowel (the bin fall). Stemberger found facilitatory effects of priming the present tense verb, where maintenance of the vowel caused an increase in over-regularisation errors compared with an unrelated prime condition, but an apparent inhibitory effect with past-tense verb primes (priming 'the bell fall' made the production of 'fell' less likely compared with an unrelated noun condition). However, the normal participants in Stemberger's experiment made very few errors overall. To test the effects of lexical neighbourhoods on production in JF, this same paradigm was adopted here. In a patient prone to over-regularise past tense endings, is performance affected by priming a lexical neighbourhood? Note that effects of priming would fit with a single mechanism account, whereas a dual route account holds that over-regularisations reflect a rule-based non-lexical process which may be isolated from lexical priming.

The effect of semantic impairments on past tense verb production

In their study of patients with semantic dementia, Patterson et al. (2001) found that production of past tense irregular verbs was highly associated with the degree of comprehension deficit, as measured by a verb synonym judgement task. The specificity of the impact of verb comprehension was evaluated here, testing whether understanding of a verb's meaning is a significant predictor of past-tense production accuracy for that verb at the level of individual verbs in a single case. Are any associations between meaning knowledge and production accuracy stronger for low frequency and low neighbourhood irregulars compared to other irregular verb sets?

Both single and dual-mechanism theories can accommodate the idea that the production of irregular past tense verbs is impeded by a semantic deficit, the former attributing this to reduced input from semantics and the latter accounting for a selective difficulty for irregular past tense verbs in terms of damage to the lexical mechanism of verb retrieval. However, the finding of any frequency effect in past tense verb production is not predicted by dual mechanism accounts that hold that a semantic deficit does not necessarily produce an impairment for the irregular past tense (Tyler et al., 2004). On the other hand, single mechanism models predict that the degree of frequency influences the likelihood of correct past tense verb production.

Similarly, a single-mechanism perspective may be better able to accommodate some differences between contrasting types of irregular verbs. Effects of quasi-regularity can be attributed to the degree of overlap between the stem and the past tense form, in that the stem serves as a more effective cue for the past tense form for some clusters of irregulars compared with others (e.g. build-built vs. fly-flew).

In addition, the theories predict different outcomes regarding the influence of semantic knowledge on verb retrieval. A dual mechanism view holds that retrieval will be mediated by the strength of lexical representations (i.e. poor performance where a verb is 'missing' from the semantic system) whereas a single mechanism view predicts a graded, comprehension-modulated deficit for low frequency irregulars less assisted by phonological 'neighbourly' support (Patterson et al. 2001). In order to examine the contributions of lexical and semantic knowledge independently, the current study uses one task levelled at assessing comprehension of verb stems, and another lexical decision task to measure whether lexical representations are intact or degraded. Within a single mechanism view, low frequency irregulars with little neighbourly support rely on semantic input, and so on these irregular sets, semantic knowledge about a verb should modulate production accuracy. A dual mechanism view suggests that effects of meaning may be related to whether the part of the lexical store containing a given verb is damaged or preserved.

In the current study, a patient with a semantic deficit (JF) is tested throughout several studies designed to assess past tense verb production. Study 1 contrasted JF's production of regular and different types of irregular past tense verbs (those that take a terminal consonant change, a vowel and terminal consonant change, and a vowel change). Study 2 contrasted his production of regular and irregular verbs of high and low lexical frequency, and in irregulars from large and small phonological 'neighbourhoods'. Study 3 compared his production on irregular verbs for which meaning knowledge was impaired versus preserved. Study 4 assessed JF's production of irregular verbs of high and low lexical frequency under priming conditions that primed the present or past tense vowel sound. The outcomes from the studies will test assertions made by single and dual mechanism approaches in the following ways: If performance is abnormally graded by regularity in Study 1, then this is consistent with an influence from phonological transformation frequency that may be more pronounced in a case of impoverished semantics, an outcome that can be explained by both approaches. Similarly, a frequency effect (Study 2) can be explained in terms of input from frequency under a single mechanism connectionist view, and frequency modulating the probability of lexical access from a dual mechanism perspective. However, finding an effect of irregular phonological 'neighbourhood' in this study would be at odds with any dual mechanism model incorporating a lexical store with limited interactivity between lexical items (e.g. Tyler et al. 2004).

Case Description

JF was a 64-year-old right-handed male, who formerly worked as a self-employed mechanical engineer, and also managed his own consultancy firm. JF was referred to the Behaviour and Brain Sciences Centre at Birmingham University with a provisional diagnosis of primary progressive aphasia, having presented with a three to four year history of problems in spoken and written language, with no memory impairment or behavioural symptoms. A structural MRI scan revealed bilateral posterior atrophy, particularly around the region of the left intra-parietal sulcus, and minimal changes in the medial temporal regions (see Figure 2.1.1). He began to show increased difficulties with certain tasks throughout testing sessions conducted between 2007 and 2008, exhibiting signs of anomia. A more recent scan conducted in 2009 revealed some small areas of change in superior and inferior parietal gyri and a small lesion in the rolandic operculum when compared with the earlier scan. Figure 2.1.1 shows the uncorrected slices from the recent scan. A coronal image from JF's standard scan is provided in Figure 2.1.2, which indicates bi-parietal atrophy.

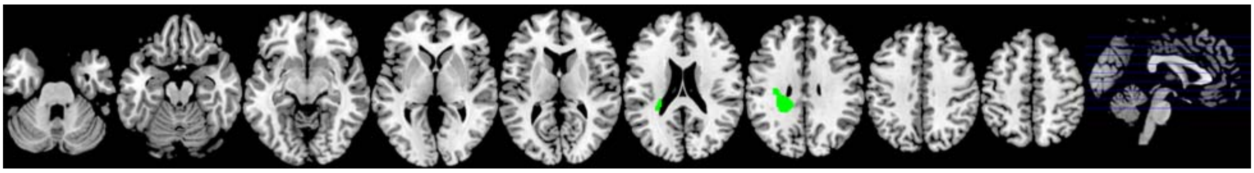


Figure 2.1.1. JF brain images from the recent (2009) scan, uncorrected. N.B. Grey matter lesion appears in red and white matter lesion in green. The lesion was created in SPM and added as an overlay onto a standard multi-slice template in MRIcron. The SPM analysis was a one sample t- test with the covariates healthy (140 brains aged 40+) vs. patient, age and gender.

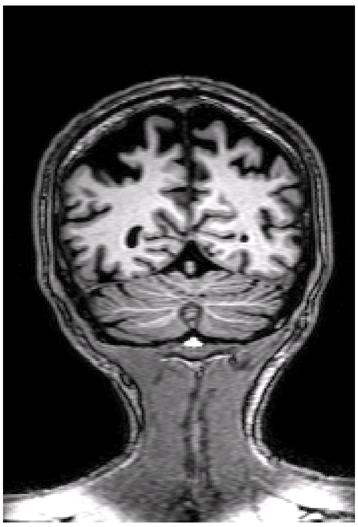


Figure 2.1.2. Coronal image from JF's standard clinical scan

Tests of semantic knowledge

Although JF showed no signs of a semantic deficit when he was first referred to the University, he began to show signs of difficulty in accessing semantic information throughout sessions conducted between 2007 and 2008. On standardised tests of semantic knowledge, JF frequently performed outside the control range. He scored 1.92 standard deviations below the control mean for high imageability words on a test of semantic association which required selection of the word that was closest in meaning to the target word amongst 1 semantically related and 3 unrelated distracter items (PALPA 51, Kay, Lesser and Coltheart 1992). For low imageability words JF scored 2.33 standard deviations below controls. In both cases the most frequent error was in choosing semantically related distracter items.

His performance fell just below the cut-off for control performance on the Pyramids and Palm Trees Tests (Howard and Patterson, 1992), scoring 49/52 in the three pictures condition, and 46/52 with one written word and two pictures (tested in June 2009). JF also made some errors on an auditory synonym judgement task (PALPA 49; Kay et al., 1992), scoring 28/30 with high imageability, and 26/30 with low imageability words. In the written modality, JF scored 28/30 on high imageability items and 27/30 on low imageability items (PALPA 50; Kay et al. 1992). No control norms are available for performance on the synonym judgement tasks, but one would expect fewer errors on this straightforward task from a control population. On a novel meaning-match task using low concrete words, JF scored 21/30, a score that was 3.74 standard deviations below the mean score from a group of five age-matched male controls.

Auditory processing

JF performed well on two tests of phonological segmentation (PALPA 16 Phonological Segmentation of Initial sounds: words 30/30, nonwords 15/15; PALPA 17 Phonological Segmentation of Final sounds: words 30/30 nonwords 14/15). He scored perfectly on a task of minimal pair discrimination of words (72/72; PALPA 2 Kay et al. 1992) and nonwords (72/72; PALPA 1 Kay et al. 1992).

He performed letter discrimination tasks taken from the PALPA battery (tests 18-21, Kay et al., 1992) with very few errors. JF successfully performed mirror-reversed letter discrimination (scoring 34/36 relative to a control mean of 35.44, PALPA 18), and cross-case matching of single letters scoring 26/26 in both upper to lower (PALPA 19) and lower to upper case (PALPA 20) modalities.

Speech processing

JF scored 2.29 standard deviations below the controls on spoken picture naming (PALPA 53; Kay et al., 1992), although he only made one visual-semantic error (producing ‘finger’ for a picture of a thumb). He made some other visual-semantic errors throughout testing (producing ‘strap’ for belt and ‘comb’ for brush), which were automatically self-corrected.

We assessed JF’s free speech using a picture description task taken from the Comprehensive Aphasia Test battery (CAT 19 Spoken Picture Description, Swinborn, Porter and Howard, 2004) that was recorded and later transcribed and analysed. JF was able to describe most of the important features in the picture, although he neglected to detail one aspect of the picture that portrayed books falling off a shelf. His speech was hesitant, containing frequent signs of word finding problems:

'It's in a house and err I see the the boy he the father I would think and its err err erm they seem to be relaxing there looking at err […] there's a err an err a cat up there that's trying to err chase out err fish and err the c there's a cup on the table and they seem to be in the living room.'

JF scored perfectly on a test of sentence reading (PALPA 37, Kay et al. 1992).

Spelling

JF's written production of letters was generally good, producing well-formed letters showing no additional impairment of graphomotor skills. He scored well (104/108, 96.2%) when copying words, but there were more errors when copying the same items after a delay (33/72, 45.8%) which may be attributable to difficulties in holding the orthographic representations in memory. Errors included producing mirror image letters, case errors, letter deletion and substitution.

JF's spelling performance is consistent with a graphemic buffer (GB) impairment. He showed signs of a length effect in spelling, and also had the same error pattern as that demonstrated by GB patients (see Rapp and Kong, 2002, Sage and Ellis, 2006). JF's spelling impairment is not the primary focus of this investigation, and is described in more detail in the next chapter (Chapter 4: Rehabilitation of spelling in a patient with a graphemic buffer impairment: The role of orthographic neighbourhood in remediating the serial position effect).

Study 1: Production accuracy across regular and different types of irregular past tense verbs

A specific impairment for irregular vs. regular past tense verbs has been reported in neuropsychological patients (Patterson, Lambon Ralph, Hodges and McClelland, 2001). Study 1a assessed JF's production of regular and irregular verbs to test the assumption that he should be significantly worse at generating irregular past tense verbs relative to regulars. Patterson et al. (2001) found an interesting pattern of results in production across different types of irregular verb transformations: Their semantic dementia patients scored highest with irregular verbs that take just a terminal consonant change (mean proportion correct 0.76), less well on verbs that take both a vowel and terminal consonant change (mean proportion correct 0.62) and least well on vowel change irregulars (mean proportion correct 0.55). However, these results were produced on limited numbers of stimuli, as this was not a principle focus of their study. The aim was to explore directly the role of quasi-regularity in irregular verb production. Study 1b assessed performance on three types of irregular inflection: vowel change irregulars (bleed-bled), irregulars that take a terminal consonant change (build-built), and those that require both a vowel change and the addition of a /t/ or /d/ terminal consonant (kneel-knelt).

Method

Participants

Patient JF and a group of five age- and education-matched male control participants were used. Control subjects received payment in exchange for their participation.

Materials

Study 1a used thirty-four regular (e.g. play) and thirty-four irregular (e.g. catch) frequency and imageability-matched verbs. In Study 1b sixteen terminal consonant irregulars, sixteen vowel change irregulars, and sixteen vowel and terminal consonant irregular verbs were used. Verbs in each set were matched as far as possible on a range of lexical variables, though close matching was not always possible due to restricted stimuli sets. See Tables 2.1 and 2.2.

Set	N		Celex frequency	Log frequency	Familiarity	Letter length	Imageability
Regular	34	Present	197.47	2.107	512.63	4.36	134.63
		Past	161.25	2.06	49.32	6.22	42.3
Irregular	34	Present	169.29	1.96	465.913	4.17	125.71
		Past	169.65	1.99	311.844	4.45	56.72

Table 2.1. Scores on lexical variables for Experiment 1a

Set	N		Celex frequency	Log frequency	Familiarity	Letter length	Imageability
Terminal consonant change irregulars	16	Present	45.72	1.50	444.28	4.57	178.5
		Past	51.445	1.39	222.21	4.78	86.28
Vowel-change irregulars	16	Present	116.56	1.52	374.43	4.75	123.96
		Past	85.68	1.36	269.72	4.62	102.26
Vowel and terminal consonant change	16	Present	274.60	1.84	478.74	4.20	157.02
		Past	251.87	1.76	398.53	4.71	127.46

Table 2.2. Scores on lexical variables for Experiment 1b

Procedure

The examiner read out sentences such as ‘Today I play, yesterday I _____’. JF was required to complete each sentence verbally, producing the past tense of the verb stem provided. This same procedure was used in studies 1a and 1b, and for all production tasks presented here.

Results and Discussion

Study 1a

Study 1a showed that JF had a significant deficit for irregular verbs compared with regular verb production. He scored 15/32 with irregulars, but scored perfectly on the 32 regulars (Pearson’s Chi Square test: X^2 (1) = 23.15, $p < .001$). Most errors were due to over-regularisations (e.g. ‘catched’), although JF also made four mixed errors (e.g. ‘tored’). He was not impaired compared to the controls on production of the regular verbs, but his score was significantly different to controls on the irregular set (relative to the control mean, X^2 (1) = 93.257, $p < .001$, Figure 2.1.3).

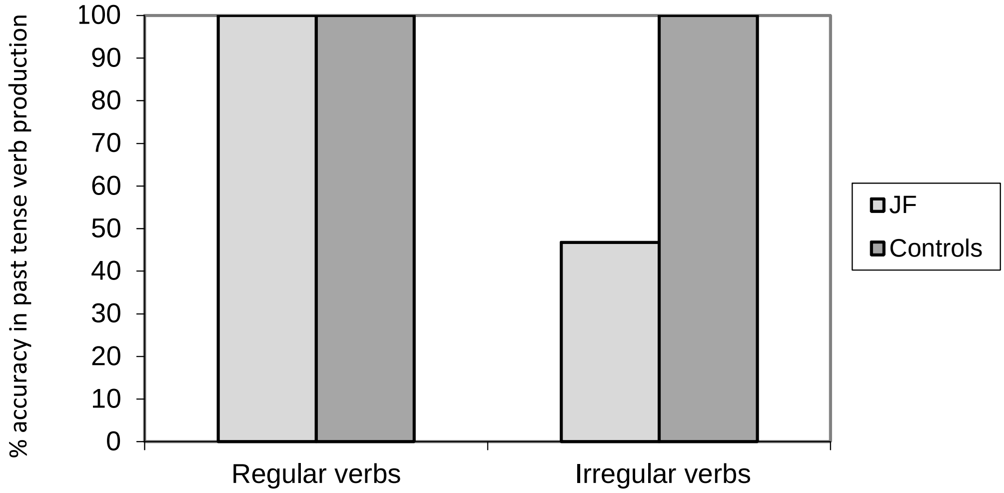


Figure 2.1.3: JF's performance on irregular and regular verbs, Study 1a.

Study 1b

Study 1b found that JF's performance on vowel change irregulars was significantly poorer (59.37%) compared to his performance on both the terminal consonant change only irregulars (87.5%, X^2 (1) = 4.451, p = .035), and the terminal consonant change and vowel change irregulars (84.37%, X^2 (1) =6.098, p = .014). There were no significant differences between terminal consonant and vowel change and terminal consonant irregulars (p = .772, Figure 2.12). JF scored 10.22 standard deviations below the mean for the terminal consonant and vowel change irregulars, and 10.88 standard deviations below the mean for the vowel change irregulars.

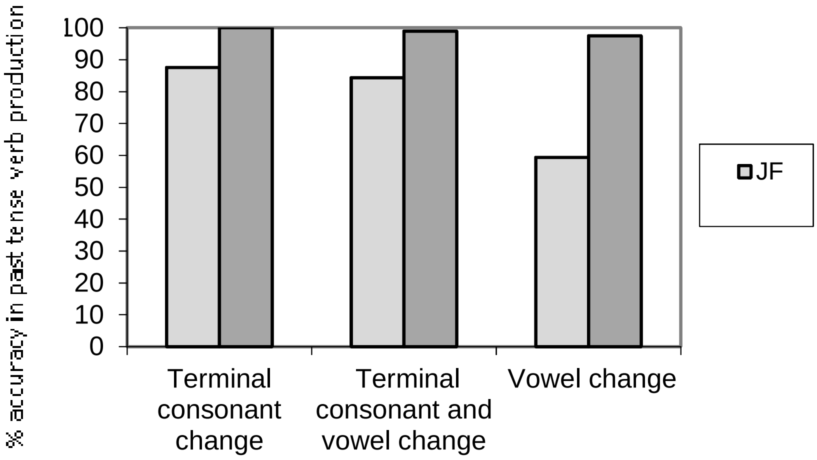


Figure 2.1.4: Percent production accuracy across the three types of irregular verb for JF and the controls (mean percentage), Study 1b.

The vowel change irregulars produced the lowest scores for JF, as is consistent with Patterson et al.'s (2001) findings with semantic dementia patients. However, accuracy did not seem to be graded in relation to the degree of regularity of the verbs. Patterson et al. found that accuracy on vowel and terminal consonant irregulars lay between vowel change and terminal consonant types, a pattern of performance that was not found with that of JF (Figure 2.1.4). However, the sets were not matched on the number of phonologically similar irregular present and past tense pairings shared by each item (e.g. sleep-slept; weep-wept): The vowel change irregulars contained the largest number of items with few phonologically similar pairings, whereas the other two terminal consonant change sets were comprised of items with many phonologically similar 'neighbours'. Study 2 assessed the impact of this neighbourhood support (in terms of phonological similarity between present and past pairings) on irregular past tense verb production. As other semantically impaired neuropsychological patients in the literature show a pronounced frequency effect in irregular verb production, whether items were of high or low frequency was also factored into the analysis.

Study 2: Production of regular and irregular past tense verbs across high and low frequency sets, and across irregulars with high and low neighbourhood support.

Study 2 aimed to test whether JF's impairment for irregular verb forms was modulated by frequency, with higher frequency verbs faring better than lower frequency items. A single mechanism view expects a frequency verb advantage for irregular verbs. Although JF performed at ceiling with the regular verbs used in Study 1a, it is still possible that he would make errors on low frequency regular items. Patterson et al. (2001) note that neighbourhood support amongst irregular verbs exerts a more specific impact than in the regular past tense due to the high degree of phonological similarity within irregular neighbourhood clusters. Although the regular inflection is the most pervasive, if the single mechanism account is correct then there should be effects of phonological similarity and neighbourhood with some regular verbs (i.e. there should be cases where the tendency to produce the –ed suffix is offset where regular verbs are less assisted by frequency, leading to '*irregularisation*' errors). Evidence for 'irregularisation' errors on low frequency regular items would support the single mechanism account, that all verbs are produced in one network and that performance on both irregular and regular verbs is mediated by frequency, neighbourly support and input from semantics. In contrast, a failure to find any errors on low frequency regular verbs would be more consistent with a dual mechanism view, which holds that regular and irregular verbs are produced by two separate and neurologically differentiated systems.

The nine groups of irregular verbs outlined by McClelland and Patterson (2002a) produce unequal and in some cases limited stimulus sets; making contrasts between the groups led to low statistical power, particularly after matching the items on lexical variables. Further, the McClelland and Patterson (2002a) categorisations are based on type of change from the present to past tense, rather than on phonologically similar pairings. For example, hear-heard, do-did and sleep-slept all appear in one group. For these reasons the analyses were based on phonological neighbourhood size (i.e. the number of phonologically similar rhyming pairings that exist for each irregular present and past tense pair). Classifying the stimuli in this way ensured that past tense information was specified at the stem verb level. For example, matching by phonologically similar (rhyming) pairings resulted in sleep-slept, weep-wept, keep-kept and creep-crept appearing in one 'neighbourhood'.

Method

Participants

The study used the same participants described for Study 1.

Materials

There were 120 irregular verbs: 22 low frequency irregulars with few (3 or fewer) phonologically similar neighbours (LFLN), 20 low frequency irregulars with many (4 or more) phonologically similar neighbours (LFHN), 51 high frequency irregulars with few phonologically similar neighbours (HFLN); and 27 high frequency irregulars with many phonologically similar neighbours (HFHN). The sets were closely matched on a range of lexical variables (Table 2.3.1). Seventy-three regular verbs (43 high frequency and 30 low frequency) were also employed.

Table 2.3.1. Mean scores on a range of lexical variables for irregular verb stimuli across conditions

Set	N		Celex frequency	Log frequency	Familiarity	Letter length	Imageability	Phonological present-past pairing neighbours	Phonological pairing neighbourhood size
LFLN	22	Present	7.43	0.83	305.36	4.68	294.86	0.91	1.91
		Past	19.61	0.85	217.77	4.45	189.14		
LFHN	20	Present	7.55	0.86	302.96	4.69	275.88	3.42	4.42
		Past	18.86	0.90	209.58	4.55	167.42		
HFLN	51	Present	153.91	2.01	497.78	4.33	378.78	0.78	1.78
		Past	133.51	1.88	321.25	4.59	218.69		
HFHN	27	Present	151.02	1.91	478.48	4.37	398.93	4.74	5.74
		Past	105.75	1.70	295.41	4.30	220.85		

Procedure

The procedure was the same as used in Study 1.

Results and Discussion

Irregular verbs

JF's scores were 3.26 standard deviations below the control mean for the LFLN irregulars (11/22 vs. 98/110, *SD* = 2.70), 1.79 standard deviations below controls for the LFHN irregulars (17/20 vs. 93/100, *SD* = 1.34), 6.77 standard deviations below controls for HFLN items (44/51 vs. 246/255, *SD* = 1.09) and were within one standard deviation of the control mean for HFHN irregulars (25/27 vs. 130/135, *SD* = 1). Most

(16) of JF's errors were pure regularisations (e.g. 'catched'). Six errors were due to the production of phonologically similar words that often formed past tense verb substitutions (fill – filtered, flee – flew), and JF also made one 'mixed' error ('frozed').

JF's performance on the LFLN irregulars was significantly poorer than on the LFHN irregulars ($X^2(1) = 5.775$, $p = .016$), HFLN irregulars ($X^2(1) = 10.886$, $p = .001$), and HFHN irregulars ($X^2(1) = 5.547$, $p = .019$). No significant differences existed between LFHN, HFLN and HFHN sets (all $p > .05$), indicating that increases in accuracy can be brought about when either the influence of frequency or neighbourhood is strong. There was not an additional increase when both factors were strong (see Figure 2.2.1).

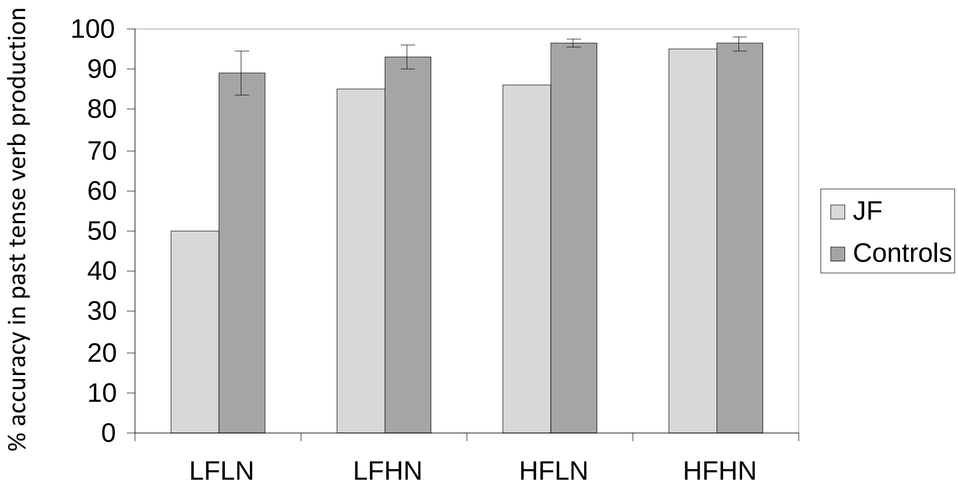


Figure 2.2.1: Percent production accuracy across frequency and neighbourhood size irregular sets for JF and controls (mean percentage)

Regular verbs

A significant association occurred between frequency and production accuracy on the regular verbs ($X^2(1) = 4.818$, $p = .028$). Only one error was made on the 43 high frequency regulars (walk – wowked), while five errors were made on the thirty low frequency regulars. Three errors involved an irregular verb style vowel change, two of which resulted in nonwords (e.g. fold-feld) and one in a real word (greet-grate/great). The two remaining errors were a regularised nonword (scream-scremmed) and a regular verb substitution (wipe-whipped). The controls performed perfectly in both conditions of the production task. JF score was significantly lower than controls for both irregular ($X^2(1) = 24.100$, $p < .001$), and regular verb production ($X^2(1) = 4.796$, $p = .029$) (see Figure 2.2.2).

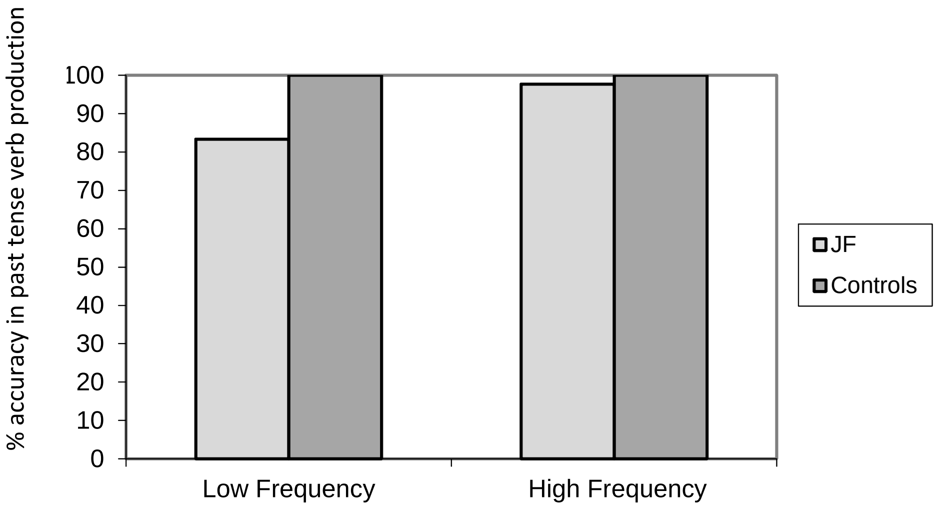


Figure 2.2.2: Percent production accuracy across low and high frequency regular verbs for JF and controls (mean percentage).

The findings in Study 2 emphasise the contributions of neighbourhood support and frequency to JF's generation of English past tense verbs. That a frequency effect was also found with regular verbs is consistent with a single-mechanism view, and suggests that frequency may facilitate performance on both types, albeit to a lesser extent with regular verbs due to the pervasiveness of the -ed inflection. Some of the errors produced on regular verbs comprised what may be deemed to be irregular-style transformations, suggesting that poorer performance on low frequency regular verbs may be partially attributable to interference from irregular verbs.

Study 3: To what extent does knowledge of meaning mediate irregular past tense production?

The single-mechanism account asserts that low-frequency irregular verbs that receive little benefit from neighbourly support rely on input from semantics more than other items. If this is the case, then low frequency irregular verbs with low phonological similarity should be more error-prone in JF, given that support from semantics is reduced in this case. A previous study assessed the association between past tense verb production and semantics by comparing overall scores on past tense verb production with those on synonym judgements made on the same verbs in a group of semantic dementia patients (Patterson et al., 2001). Here past tense verb production was predicted by performance on the verb synonym judgement task and the relationship between semantic knowledge and verb production was tested by assessing whether knowledge about an item's meaning predicts success in past tense production of that verb (particularly for low frequency, low neighbourhood irregular verbs which should depend most heavily on semantic support).

A dual mechanism account could accommodate the finding that production performance is predicted by meaning knowledge if there is damage to the lexical system. For example, poor performance on both production and meaning judgements about a given verb could indicate that the lexical representation for that verb is 'missing' from the lexical store. To test for this, a lexical decision task was also conducted on irregular verbs to dissociate possible impairments in verb meaning knowledge from degraded lexical representations.

To assess semantic knowledge of verbs a verb and definition matching task was employed, where a verb was presented and had to be matched against either of two meanings (one correct and one incorrect). A main focus was on low frequency low neighbourhood size irregular verbs, but performance on other irregular verbs where JF made incorrect responses on the meaning task was also tested to investigate which factors modulated production in this case (e.g. neighbourhood size, phonological similarity, quasi-regularity). Comparisons of the verb – meaning match task results with production accuracy evaluated the extent to which knowledge of irregular verb meanings predicts the accuracy of past tense verb production.

Method

Participants

JF was tested on all aspects of Study 3. Due to the near-ceiling performance of controls across all conditions of the irregular production task, it was not useful to make meaning and production comparisons on these data. Nevertheless controls were assessed on the meaning-match and lexical decision tasks to determine whether JF was impaired in his knowledge of irregular verb meanings and lexicality in comparison. Again, the controls received payment for their participation.

Materials

There were 22 low frequency irregular verbs with few phonologically similar present and past tense pairings (low neighbourhood size), 24 low frequency high neighbourhood size irregulars, 24 high frequency low neighbourhood size irregulars, and 24 high frequency high neighbourhood size irregulars. For the lexical decision task, the same 120 irregular verb stems as detailed in Study 2 were used, along with 120 nonwords that were created by making a single letter change either at the start or end of the irregular verb set. The production task used the same irregular verbs as in both the lexical decision and meaning knowledge tasks. Mean ratings on lexical variables for the stimuli are provided in Table 2.3.2.

Table 2.3.2. Mean scores on a range of lexical variables for the irregular verb stimuli across conditions

Set	**N**		**Celex frequency**	**Log frequency**	**Familiarity**	**Letter length**	**Imageability**	**Phonological present-past pairing neighbours**	**Phonological pairing neighbourhood size**
LFLN	22	Present	7.43	0.83	305.36	4.68	294.86	0.91	1.91
		Past	19.61	0.85	217.77	4.45	189.14		
LFHN	24	Present	7.47	0.85	297.56	4.59	271.55	3.70	4.70
		Past	17.82	0.89	212.12	4.46	172.32		
HFLN	24	Present	125.33	2.00	530.04	4.21	422.00	0.63	1.63
		Past	110.36	1.86	258.08	4.46	178.88		
HFHN	24	Present	116.42	1.95	401.13	4.21	323.29	5.21	6.21
		Past	105.07	1.78	290.33	4.17	228.13		

Procedure

For each trial, JF was presented with an irregular verb in its root form, and two choice definitions, one of which was the target and defined the verb, and the other was a foil meaning that defined another, unrelated, verb. The verbs and definitions were presented in written form, and were also read out by the examiner. The examiner asked JF to choose the definition that was closest in meaning to that of the current verb. JF was advised that he could re-read, or the examiner could repeat the verbs and meanings at any time, in order to prevent error due to difficulties in memorising the verbs and meanings. JF pointed to or read out the definition that he believed matched the verb, and the examiner recorded the accuracy of each response. The same task was given on six separate occasions. Verbs that JF responded to correctly on five or six trials were classed as 'meaning known' and verbs that produced correct responses on four or fewer trials were termed 'meaning not known'. No feedback was given to JF regarding his accuracy on the task. For the auditory lexical decision task, the examiner read out verbs and nonwords and advised JF to indicate whether each word was real or made-up. Performance on both tasks was compared with production scores on those same irregular verbs using the same sentence completion exercise described in Study 1.

Results and Discussion

A loglinear analysis was performed using production accuracy, meaning score, frequency (low or high), and neighbourhood (low or high) as factors. The analyses returned significant three-way interactions between production accuracy, meaning score and neighbourhood size (X^2 (1) = 5.823, p = .016) and meaning task score, frequency and neighbourhood size (X^2 (1) = 7.070, p = .008). This latter interaction reflects a difference in the number of words that JF knew (meaning known) in the different frequency and neighbourhood size groups. JF knew more low frequency and low *N* words (13/22), and more high frequency and high *N* words (12/24) than low frequency, high N and high frequency, low N words (8/24 and 8/24 respectively, Figure 2.3.1).

Of prime interest here, though, was JF's production accuracy in relation to his knowledge of the word meanings. To assess this, the interaction between production accuracy, meaning score and neighbourhood was tested using loglinear analyses (see Figure 2.3.1). On words from low neighbourhoods, JF's production was better for known than unknown words (known 13/21, 62% and unknown 10/25, 40%) and he performed better with high than low frequency words (LNHF 15/22, 68% vs. LNLF 8/22, 36%, $X^2(1) = 4.464$, $p = .035$). These two factors combined additively, and a highly significant three-way interaction was found between production accuracy, meaning score and frequency for the low neighbourhood verbs ($X^2(7) = 20.667$, $p = .004$).

For words with high neighbourhood densities, performance was unaffected by frequency when the word was known (low frequency 4/8, high frequency 5/12, $X^2(1) = .135$, $p = .714$) and not known (low frequency 8/16 vs. high frequency 9/12, $X^2(1) = 1.797$, $p = .180$). No significant interactions were found when the high N data were entered into a loglinear analysis (low frequency 12/24, 50% vs. high frequency 14/24, 58.3%, and meaning known 9/20, 45% vs. meaning not known 17/28, 61%, all p's > .2). One result that appears puzzling is the reversed effect of meaning knowledge for HFHN irregulars (with production being poorer on those irregulars for which meaning was known relative to not known, Figure 2.3.1). However, this result did not come through in the analysis, due, probably, to the unsystematic nature of the result, which was not replicated in either of the other two manipulations that were matched on frequency or neighbourhood size (LFHN, or HFLN). Note that meaning knowledge was not supportive where both frequency and neighbourhood support was high.

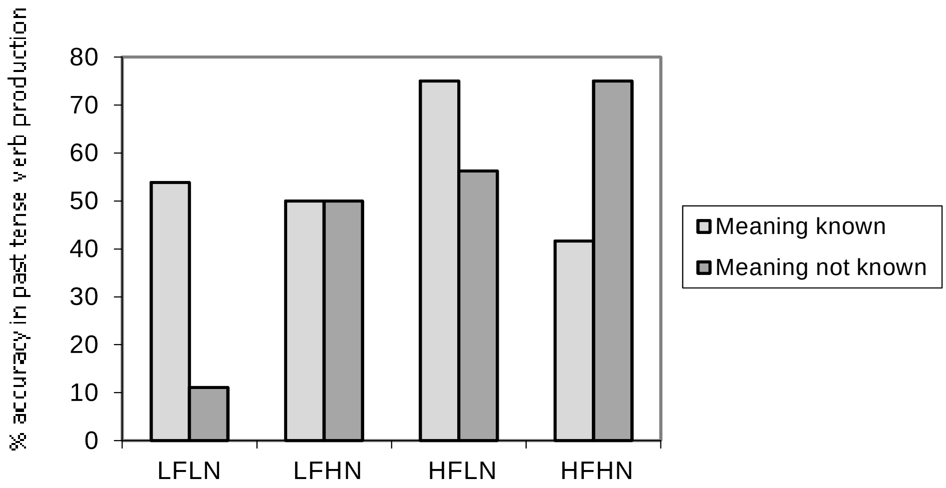


Figure 2.3.1: Percent production accuracy on meaning known and meaning not known irregulars across frequency and neighbourhood sets.

To ensure that these patterns of performance did not arise from differences in scores on the meaning-match items across the sets, the percentage of meaning-known items in each condition was examined. There were no systematic differences in meaning knowledge accuracy (Figure 2.3.2).

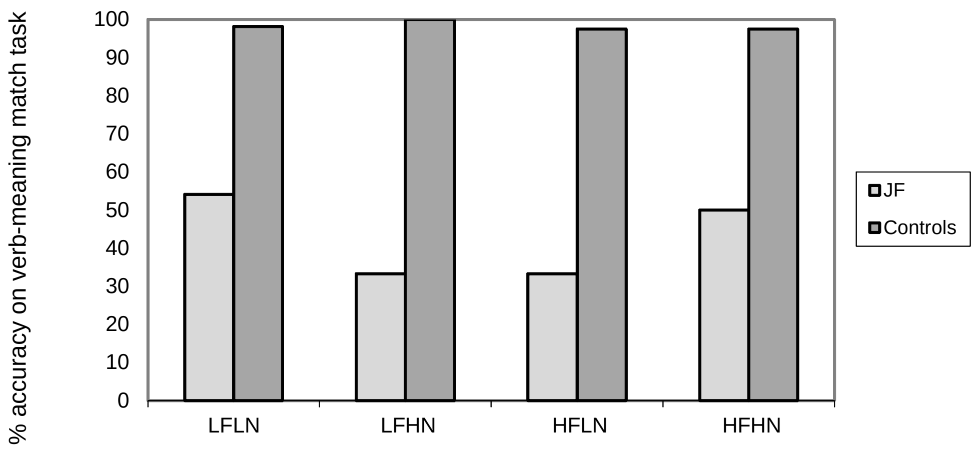


Figure 2.3.2: Percent meaning task accuracy on irregulars across frequency and neighbourhood sets for JF and controls (mean percentage)

Although JF made some (8) false positive errors on nonwords in the lexical decision task he scored perfectly on real word items, achieving 100% accuracy on both meaning known and meaning not known irregulars. Therefore, the association between meaning knowledge and production accuracy depicted in Figure 2.3.1 cannot be explained in terms of missing and present irregular items in the lexical store. The control group also scored perfectly on the real word irregulars in the lexical decision task. The 5 controls made a total of 20 false-positive errors on the nonword items, and JF's performance on nonwords was within one standard deviation of the control mean.

These findings are consistent with the single mechanism account which holds that low frequency irregulars that may not receive high levels of neighbourhood support are disproportionately error-prone for JF, and that success on these items could be mediated by semantic knowledge. Patterson et al. (2001) found associations between overall scores on synonym judgement tasks and on production of the same verbs. Through investigating production accuracy on irregular verbs for which meaning is not known, the impact of frequency and neighbourhood support was explored where meaning knowledge is degraded. The current results suggest that production success on low neighbourhood size irregular verbs may be predictable by the level of semantic knowledge that remains about each individual verb.

Study 4: Sentence priming study: frequency and neighbourhood group

Stemberger (2004a) used a visual primed sentence production task to explore the production of irregular verbs in normal participants. In their experiment nouns were used to prime verbs in 3-word noun-verb sentences such as 'The ball fall'. The nouns either primed the vowel sound in the present tense (the ball fall), the past tense verb (the bell fall) or an unrelated vowel (the pill fall). Stemberger found facilitatory effects of priming the present tense verb, where maintenance of the vowel caused an increase in over-regularisation errors compared with an unrelated prime condition, but an apparent inhibitory effect with past-tense verb primes (priming the bell fall made the production of 'fell' less likely compared with an unrelated noun condition). The finding that phonological priming may induce pronunciation errors is consistent with evidence from phonological processing studies in aphasia (Lecours and Lhermitte, 1969), and in normal participants (Shattuck-Hufnagel, 1979), and suggests an effect of competition between two vowel sounds crossed with base form dominance (see below). The present study assessed whether JF, who makes frequent over-regularisation errors on irregular verbs in free production, would be influenced by phonological priming in Stemberger's task.

Linguistic studies on child and adult regularisation errors in free speech have found that some irregular verbs are more likely to be over-regularised than others (Bybee and Slobin, 1982, Marcus, Pinker, Ullman, Hollander, Rosen and Xu, 1992). Stemberger (1993) asserts that over-regularisations are predictable by vowel dominance: if the vowel of a base form is dominant and its past form is recessive, the likelihood of an over-regularisation error is increased. In contrast, over-regularisation errors are less likely in cases where the vowel of a base form is recessive and its past form is dominant. Dominance has largely been measured in terms of phoneme frequency. The verbs were roughly equivalent in terms of present and past dominance within each set. More importantly, the stimuli were closely matched on the numbers of past- and present- vowel dominant verbs between sets (see Table 2.4.1 for frequencies of present and past vowel dominant items in each condition).

Table 2.4.1. Number of base and past-dominant verbs in each condition

Set	N	Vowel dominance	Frequency
LFLN	20	Present	12
		Past	8
LFHN	20	Present	12
		Past	8
HFLN	20	Present	12
		Past	8
HFHN	20	Present	13
		Past	7

The effects of lexical frequency on priming were also assessed. The impact of lexical frequency on over-regularisation rates have been examined in linguistic studies of free speech (e.g. Marcus et al., 1992). Bybee and Slobin (1982) found that more regularisations were produced when the past tense form was low frequency compared to a high frequency irregular form. This suggests that low frequency irregular verbs may be less assisted by frequency and neighbourly support might be more influenced by priming in JF's case, where semantic knowledge is impaired. Irregular verbs that are low frequency and have low numbers of neighbours might be more affected by priming than other verbs.

Method

Participants

Participants were the same as those detailed in Study 1.

Materials

There were 480 'The-noun-verb' structured sentences, where the noun was designed to prime the present tense verb vowel (e.g. The ball fall), the past tense verb vowel (e.g. The bell fall) or an unrelated vowel (e.g. The bill fall). Sentences comprised the same 80 irregular verbs, which were re-used across the three priming conditions, totalling 240 sentences. The four irregular verb sets were 20 LFLN irregulars (less than 3 neighbours), 20 LFHN irregulars (4 or more neighbours); 20 HFLN irregulars and 20 HFHN irregulars. The sets were closely matched on a range of lexical variables (Table 2.4.2). A further 240 sentences were comprised of regular verbs with unrelated noun primes that were used as filler items. As in Stemberger (2004), many of the sentences were semantically unusual. However, this was equally true in all sets.

Table 2.4.2. Mean scores on a range of lexical variables for irregular verb stimuli across conditions

Set	N		Celex frequency	Log frequency	Familiarity	Letter length	Imageability	Phonological present-past pairing neighbours	Phonological pairing neighbourhood size
LFLN	20	Present	8.82	0.89	337.30	4.70	331.90	0.95	1.95
		Past	21.57	0.90	213.75	4.45	184.20		
LFHN	20	Present	8.87	0.91	268.85	4.85	256.50	6.20	7.20
		Past	20.68	0.99	218.65	4.80	159.30		
HFLN	20	Present	120.86	2.02	479.00	4.20	361.70	0.55	1.55
		Past	134.38	1.94	291.70	4.60	201.80		
HFHN	20	Present	121.96	1.94	482.25	4.30	405.60	4.75	5.75
		Past	101.24	1.70	296.40	4.25	231.90		

Procedure

Initially a visual priming task was conducted as in Stemberger (2004). However, JF frequently misread the preceding noun resulting in a different vowel sound to the phoneme prime. For this reason an auditory version of the task was conducted, where the experimenter read out each sentence (e.g. 'the feet teach') and asked JF to repeat the entire sentence, but to change it to the past tense ('as if it happened yesterday'). The examiner recorded each response. JF completed the experiment twice in order to increase data points. Testing was conducted within two, one hour-long testing sessions separated by one week.

Results and Discussion

JF was significantly impaired on the primed sentence production task, scoring 11.14, 5.5 and 7.8 standard deviations below the control mean in the present, past and unrelated conditions respectively. JF's production accuracy data across conditions are provided in Figure 2.4.1.

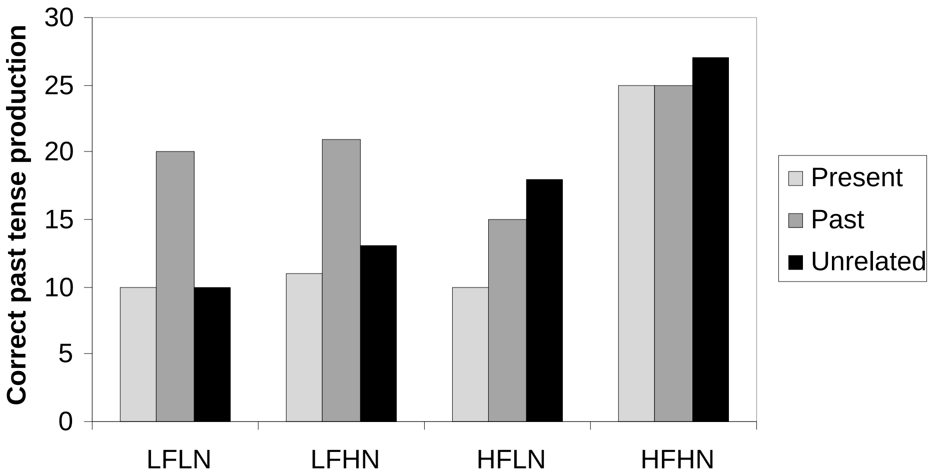


Figure 2.4.1. Accuracy scores across frequency and neighbourhood conditions in each priming condition: Patient JF (max 40)

Two separate four-way loglinear analyses were performed on JF's data with prime type, frequency, neighbourhood size and production score entered as factors. One analysis used the present prime and unrelated condition data, and the other used the past prime and unrelated scores. The present prime analysis produced a model that retained two two-way interactions, between production accuracy and frequency (X^2 (1) = 19.684, p < .001), and production accuracy and neighbourhood size (X^2 (1) = .660, p = .001). A three-way interaction between production accuracy, frequency and neighbourhood size approached significance (X^2 (1) = 3.427, p = .064). The loglinear analyses performed on the past prime and unrelated data generated a three-way interaction between production accuracy, prime type and frequency (X^2 (1) = 8.133, p = .004), and a significant two-way interaction between production accuracy and neighbourhood size (X^2 (1) = 7.860, p = .005, Figure 2.4.1).

To decompose the interactions, separate three-way loglinear analyses were conducted on the data from each neighbourhood and priming combination (low *N* present primes, low *N* past primes, high *N* present primes and low *N* past primes) with production accuracy, frequency and prime type entered as factors. For the low neighbourhood present prime and unrelated data, no significant three-way interactions emerged (all p's > .1), although a two way interaction between production accuracy and frequency was highly significant (X^2 (1) = 7.408, p = .006). For the low *N* past prime and unrelated data, the analysis returned a significant three-way interaction between production accuracy, prime type, and frequency (X^2 (7) = 15.450, p = .031).

For the high *N* present prime and unrelated data set, the analyses returned no significant interactions involving prime type (all p's > .1), and a highly significant two-way interaction between production accuracy and frequency (X^2 (1) = 20.162, p < .001). For data with the high neighbourhood unrelated and past primes, the analyses returned a three-way interaction between production accuracy, prime type and frequency that approached significance (X^2 (7) = 12.900, p = .075).

Production score did not interact with prime condition in either the low neighbourhood or the high neighbourhood present-prime analyses. Although inhibitory effects of present prime can be observed for the LFHN, HFLN, and HFHN sets, they failed to reach significance, and the unrelated data does show differences across condition.

For past tense primes, production score interacted with prime type and frequency for low neighbourhood verbs (X^2 (7) = 15.450, p = .031). In the past tense and unrelated data for the high neighbourhood set, an interaction between production score, prime type and frequency approached significance (X^2 (7) = 12.900, p = .075). These interactions reflect facilitatory effects of past primes with low frequency irregular verbs, and 'inhibitory' effects of past primes with higher frequency forms for both low and high neighbourhood sets (Figure 2.4.1). This suggests that lower frequency irregular verbs are more readily influenced by phonological information provided in past noun primes than higher frequency verbs, that may produce a pattern of performance analogous with normal participants (Shattuck-Hufnagel, 1979).

Controls achieved near-ceiling performance on all sets and no interactions existed when the control data were entered into loglinear analyses (Figure 2.4.2).

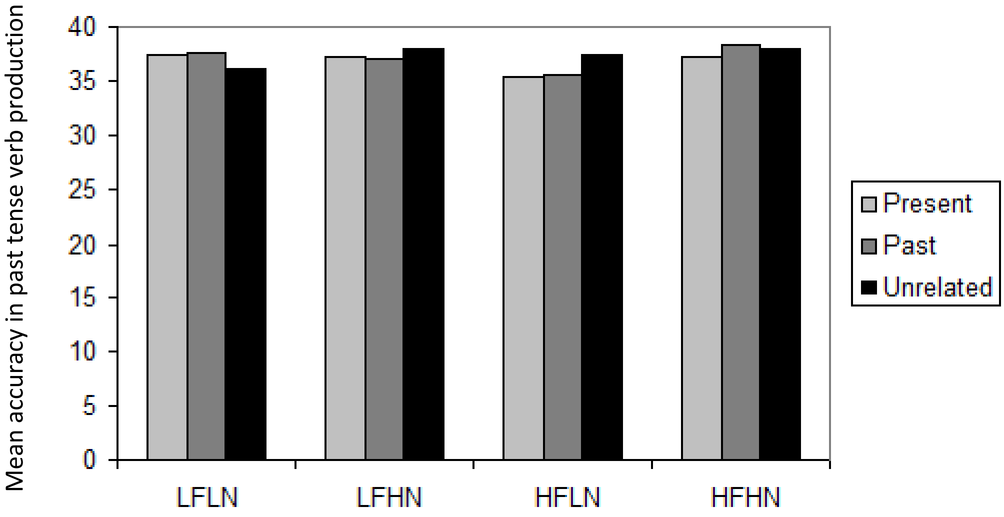


Figure 2.4.2: Mean accuracy scores across frequency and neighbourhood conditions in each priming condition: Controls (max 40).

Summary

Study 4 used a sentence completion paradigm to prime the vowel sound in the present irregular and past irregular verb. My aim was to investigate the role of phonological support in irregular past tense verb production. In the control literature (Stemberger 2004), visually presented present tense primes led to more over-regularisation errors (e.g. The ball fall = The ball falled), relative to an unrelated noun condition. From these results, it was predicted that JF (given his preponderance of over-regularisation errors in normal test conditions) would produce over-regularisation errors in the unrelated condition, would be further hindered by a present-tense priming condition, and be facilitated by past-tense primes. It was anticipated, though, that the strength of this priming effect would differ across combined conditions of high / low frequency and neighbourhood. Specifically, the prediction was that low frequency and neighbourhood size irregulars would be more susceptible to priming effects than other sets in patient JF, due to the potentially weaker representations for these items.

We used present-tense priming (e.g. The ball fall) to test whether this condition led to an increase in over-regularisation errors compared with the unrelated set. This was not the case, with accuracy rates being roughly equivalent in present prime and unrelated prime conditions. There were, however, significant interactions between production accuracy and frequency; production accuracy and neighbourhood size, and an interaction between production accuracy, frequency and neighbourhood size that was nearly significant (p = .064). These data reflect ‘neighbourhood’ and frequency effects on production accuracy in the present and unrelated primed production conditions, in that accuracy was poorer when support from these variables was weak versus strong.

Though both factors influenced JF’s performance, the findings showed that lexical frequency modulated the strength, and pattern of priming effects in the past prime analysis, over neighbourhood size. Production score, prime type and frequency interacted significantly in the loglinear analysis of the past prime vs. unrelated conditions data. This reflected a pattern of increased accuracy with past over unrelated primes that occurred with low frequency, but not with high frequency irregulars. In fact, higher frequency irregulars brought about more deleterious effects of past and facilitatory effects of present primes, with both priming conditions reducing accuracy relative to an unrelated prime condition. However, accuracy was improved to the degree of removing differences across priming conditions only when support was strong from both frequency and neighbourhood.

A dual mechanism approach can accommodate the finding of over-regularisation errors, as in some cases, the regular rule is applied before the lexicon is consulted, or where the irregular past lexical entry cannot be accessed (e.g. in the case of novel irregular verbs). A single mechanism model of past tense production accounts for over-regularisation errors through the pervasiveness of the regular inflection, which, in the context of degraded support from semantics, can override the irregular transformation where support from 'neighbourhood' and frequency is low. The findings here support this idea, in that priming production with the correct irregular past was more helpful in low relative to high frequency conditions. The finding that support from both frequency and neighbourhood was necessary to improve accuracy and remove priming effects is consistent with the other results in the current series, and more generally, with a single mechanism model.

General Discussion

Taken together, the current experiments show that neighbourhood support, frequency and semantic knowledge are key predictors of past tense verb production in patient JF. Many studies have assessed the contribution of frequency to verb production, but no study to my knowledge has systematically investigated the effect of irregular neighbourhoods. Previous investigations of the role of neighbourly support have been limited due to the small and unequal set sizes provided by irregular neighbourhood groupings (e.g. McClelland and Patterson 2002a). In classifying irregular verb neighbourhood size in terms of the number of phonologically similar / rhyming irregular pairs each item has, substantial and roughly equivalent set sizes of items were used, whilst close phonological similarity was maintained in grouping items. In using these stimuli it has been possible to demonstrate that neighbourhood size and lexical frequency predict production to a similar degree: strong input from high counts on either variable was sufficient to produce near-ceiling accuracy rates in a sentence completion task (Study 2).

It was hypothesised that, if the single mechanism account of verb production is correct, then effects of frequency and neighbourhood size should emerge even with regular verbs (i.e. there should be cases where the tendency to produce the regular suffix is offset in cases where regular verbs are less assisted by frequency; also regular verb stems may be analogous to irregular verbs with high numbers of phonologically similar past-present pairings; e.g. sleep-slept, weep-wept). Whether JF's production of regular verbs was susceptible to effects of frequency was investigated in Study 1b. Accuracy in the regular set was significantly poorer with low frequency than high frequency regular verbs, and there were instances of 'irregularisation' errors that only occurred with low frequency items. This finding is surprising given JF's strong regular > irregular dissociation coupled with the high prevalence of the regular inflection. The results emphasise the role of frequency in past tense verb production, while interference from high-neighbourhood irregular verbs was observed in the irregularisation errors. The control participants made no errors in regular verb production in either frequency condition. It may be that impaired meaning knowledge leads to an increased dependency on, and sensitivity to, lexical frequency and neighbourhood size.

The data from the meaning match and production comparison (Study 3) underlined meaning knowledge as the source of JF's deficit in production for LFLN irregulars. The study also demonstrated that the relationship between semantic knowledge and production accuracy in JF's case was highly specific: Meaning knowledge about an individual verb predicted accuracy of past tense production of that verb. Previous studies had shown that production accuracy and accuracy in making synonym judgements about the produced verbs were highly correlated, but had based their analyses on overall scores. The finding that effects of meaning existed only for LFLN irregulars suggests that input from frequency or neighbourhood support is sufficient to elicit correct productions without useful semantic input. Another finding was that production accuracy increased on meaning-not-known irregulars with increases in frequency and neighbourhood size. This finding reflects the strong influence of neighbourhood support and frequency in facilitating production where input from semantics is weak. Data from a lexical decision task on the items, which showed that JF's lexical knowledge about the items was relatively good, suggested that the effect of meaning knowledge on production accuracy may not be due to 'missing' lexical representations in the lexical store, as indicated by dual mechanism models. Again, the finding that JF's production accuracy was modulated by meaning knowledge for the LFLN irregulars, is consistent with a single mechanism view.

In Study 4 a primed sentence production paradigm (Stemberger, 2004a) was used to explore priming effects across different types of irregular verbs. Some studies have suggested that the use of a previous vowel can lead to an increased error rate on a subsequent token of the same vowel (Shattuck-Hafnagel, 1979). For example Stemberger (2004) found an inhibitory effect of a past prime condition with young controls where past tense primes lessened the likelihood of the target past tense vowel appearing (e.g. 'The bell fall' made 'fell' less likely). There were also facilitatory effects of present prime, where the likelihood that a present tense vowel appeared in the output was increased by priming, leading to over-regularisation (e.g. 'The ball fall' made 'falled' more likely). Here a group of older controls achieved a near-ceiling performance in all conditions, and so did not show any effect of priming on verb production accuracy. This is surprising given Stemberger's findings but could be due to the present use of auditory rather than visual presentation.

Our expectation was that low frequency and neighbourhood size irregulars would be more susceptible to priming effects in patient JF. A facilitatory effect was found for low frequency irregulars for past primes, but not for present primes. Due to JF's frequent over-regularisation errors in this condition, any significant increase in the number of over-regularisation errors with present primes when compared to the unrelated condition were not expected. Of prime interest instead was whether past primes were able to offset the tendency to over-regularise, and expected that the past prime condition would be more successful reducing over-regularisation errors with low frequency / neighbourhood verbs than in other conditions. The significant 3-way interaction between production score, prime type and frequency in the loglinear analysis of the past prime vs. unrelated conditions data showed that increased accuracy with past than unrelated primes occurred with low frequency, but not with high frequency irregulars. In fact, higher frequency irregulars brought about more inhibitory effects of past and facilitatory effects of present primes, with both priming conditions reducing accuracy relative to an unrelated prime condition. Although there were key differences between the current study and that employed by Stemberger (2004a), this pattern of results is more in keeping with those produced by the younger controls in Stemberger's study. Where both frequency and neighbourhood support predicted accuracy to equivalent extents in free production, frequency rather than neighbourhood seemed to mediate priming effects in Study 4. However, both high frequency and neighbourhood size was necessary to improve production accuracy and remove differences across priming conditions. This result suggests that the strength of the lexical representation and its relationship to others in the lexicon modulates performance.

In sum, there were significant effects of irregular neighbourhood size, lexical frequency for both irregular and regular verbs, and degree of meaning knowledge for low frequency, low neighbourhood irregular verbs on verb production in a patient with degraded semantic knowledge. Low frequency irregulars were also more facilitated by past tense primes than higher frequency irregulars in primed sentence production, and higher frequency verbs brought about an inhibitory effect of past primes.

There was evidence of differences in the extent to which lexical, semantic and phonological factors contribute to JF's production accuracy in the irregular past tense: Study 3 showed that neighbourhood was an important factor when meaning knowledge was manipulated, whereas neighbourhood was less important than frequency when priming production using nouns that bore phonological similarity to the irregular past tense in Study 4. Findings from Study 3 could be explained in terms of frequency being more helpful when it is free to interact with meaning knowledge (intuitively, JF's knowledge of high frequency verbs might be better than for lower frequency verbs), though this was not the case within the stimuli sets (Figure 2.3.2). The pattern of performance in Study 4 may speak to the idea that support from neighbourhoods and phonological primes may operate in a similar nature (that part of the neighbourhood effect may be due to phonological 'interference' from similar rhyming-sound pairings). For instance, a prime of the correct past tense may override, or simply negate, support from other past tense irregular verb neighbours. These conclusions are however tentative, as differences in contributions from these factors over different tasks was not the key focus of the current Chapter. Further work is needed to establish whether the extent and nature of neighbourly support and frequency differ across different production tasks.

More work with patients showing an irregular < regular dissociation is needed to establish whether impaired meaning knowledge causes increased dependency on neighbourhood and frequency effects in modulating production.

Chapter 4: Rehabilitation of spelling in a patient with a graphemic buffer impairment: The role of orthographic neighbourhood in remediating the serial position effect

Abstract

A rehabilitation study of a patient with a graphemic buffer (GB) deficit associated with primary progressive aphasia is presented. The patient (JF) showed benefits when spelling high orthographic neighbourhood words (an '*N*' effect) along with a characteristic bow-shaped accuracy curve across letter position. Using an Anagram Copy Treatment (ACT) rehabilitation method the spelling of a set of words was treated, and generalisation was tested to orthographic neighbours with shared or changed medial position letters relative to words in the treated set. Generalised improvement was found for words with shared middle letters. This result is attributed to top-down support in rehabilitation from learned lexical representations that exert a facilitative effect on untreated neighbours with distinguishing letters at the initial and final letter positions, set against a deleterious effect of training on the spelling of untreated neighbours with a medial letter position change.

Introduction

The Graphemic Output Buffer (GB) is conceptualised as a short-term memory mechanism responsible for holding the orthographic representation of a word as it is spelled. Within dual route models of spelling (e.g., Rapcsak, 1997, Caramazza, Miceli, Villa and Romani, 1987 and Ellis, 1982), the GB is typically thought to be sited at the point of convergence of sublexical and lexical spelling processes. Consistent with this, both real and non- word spelling are often similarly affected in these patients (though see Sage and Ellis, 2004).

A graphemic buffer deficit may manifest itself in a letter length effect, an abundance of segmental errors (transposition, substitution, addition and deletion), and the absence of lexical influences on spelling (Caramazza et al., 1987). Early studies of GB patients also described a bow-shaped accuracy curve across serial letter positions (lower accuracy rate in the middle sections, LB: Caramazza et al., 1987). More recently however, patients demonstrating a pattern of accuracy that declines linearly across the word (i.e. more errors at final letter positions), who also have the characteristic error types, have been reported too (patients JH and PB: Schiller, Greenhall, Shelton and Caramazza, 2001; BA: Ward and Romani, 1998 and HR: Katz, 1991). Patients showing a linear decline in letter accuracy across word position also frequently demonstrate lexical influences (BS: Ward and Romani, 1998, PB and TH: Schiller et al., 2001 and MD: Bormann, Wallesch and Blanken, 2008). These observations have led to the suggestion that there may be two classes of GB patient; those who show the bow-shaped accuracy curve and a lack of lexical effects (type A), and those demonstrating a linear decrease in accuracy across letter position and an effect of lexical variables similar to those reported in deep dysgraphia (type B). The type A deficit is commonly attributed to intact representations arriving at a damaged graphemic buffer, while the type B deficit is explained in terms of damaged processes at the GB along with deficient upstream semantic systems, leading to damaged representations entering the buffer (Glasspool, Shallice and Cipolotti, 1999).

Many neuropsychological patients fitting a type A graphemic buffer classification have been described (Patient LB: Caramazza, Miceli, Villa and Romani, 1987; AS: Jonsdottir, Shallice and Wise, 1996; SE: Posteraro, Zinelli and Mazzucchi, 1998; JRE, RSB and BWN: Buchwald and Rapp, 2004, and PR, RH and AC: Haslam, Kay, Tree and Baron, 2009). The deficits in these patients can be attributed to damage to a single, post-lexical component involved in the written production of both words and nonwords. The bow-shaped accuracy takes an exacerbated form of the spelling lapses observable in normal participants (Wing and Baddeley, 1980) with vulnerability of the middle positions being caused by several possible factors: for example (i) the greater number of neighbouring letters (and thus the greater competition for selection) in these medial positions compared to the initial and final letters, with this increased competition at the grapheme level being more problematic in the case of impaired graphemic buffer (Bormann, Wallesch and Blanken, 2008), or (ii) differential weighting of end relative to middle letter positions within the buffer (Glasspool and Houghton, 2005).

Hillis and Caramazza (1989) first reported a GB patient who made more errors on the final letters in a word, which they attributed to an attentional deficit (right neglect) that modulated use of the GB. However, a pattern of increased errors at the final positions has been reported in the absence of an attentional deficit: Patient HR, described by Katz (1991). The authors demonstrated an increase of spelling errors across serial positions. To assess whether these difficulties were mnemonic in nature, Hillis and Caramazza tested HR's ability to spell words backwards so that the final letters were written first (e.g. TABLE - ELBAT). In this backwards spelling modality, errors occurred at the last letters to be spelled rather than in the last positions of the graphemic representation itself (e.g. CHAIR – RIAHF). The authors framed these results within a ‘rapid decay’ account, postulating an abnormal decay of the GB in their patient and hence the letters reported last tend to be the ones that suffer. A similar pattern was reported by Schiller et al. (2001) in their dysgraphic patients PB and TH.

Although it was initially argued that GB deficits should be independent of lexical factors, such effects have since been reported (e.g. effects of frequency, word class and concreteness: PB and TH: Schiller et al., 2001 and lexical frequency: RSB and BWN: Buchwald and Rapp, 2004). In many cases, these effects co-occur with semantic errors and a failure in writing nonwords, consistent with a classification of deep dysgraphia (Bub and Kertesz, 1982, though see Katz, 1991 and Miceli, Capasso, Ivella, and Caramazza, 1997 for reports of type B GB patients who do not show semantic errors).

Effects of lexical variables and an advantage for real over non- words might emerge if activation is presented to the GB in a cascading form, so that a semantic level might influence processes at the graphemic level (e.g. Sage and Ellis, 2004). Ward and Romani (1998) described the linear decline in accuracy across the word produced by their deep dysgraphic patient (BA) in terms of a lexical activation rather than a graphemic buffer deficit, as this serial position effect was evinced only with tasks requiring lexical access.

The prior results then demonstrate lexical effects in type B GB patients. However, some type A graphemic buffer patients too show an effect of lexical variables (e.g. orthographic neighbourhood, age of acquisition and imageability, Sage and Ellis, 2004). One rehabilitation study attempted to use this orthographic neighbourhood effect therapeutically (Sage and Ellis, 2006). The intervention produced a generalised treatment effect to untreated words from the same orthographic neighbourhoods as treated items. The authors interpreted these findings as evidence that the buffer receives input from the lexicon, and that word representations strengthened through training improve accuracy on other items from the same orthographic neighbourhoods as treated items. Orthographic neighbourhoods can take a letter change at any position in the word, though, and it is not clear whether there is differential generalisation according to the positions of letters within a word. For example, retraining could help re-establish links between letters and position-holders with a GB (Glasspool, Shallice and Cipolotti, 2006). Given that central letters can be the most difficult to retrieve for GB patients, then training might produce the most effect (and perhaps generalise most effectively) for words that share their central letters.

Rehabilitation was performed using an Anagram Copy Treatment (ACT), which involves rearranging anagrams with subsequent repeated copying of the target word, and is designed to strengthen written word representations. This kind of method is suitable for the treatment of a Graphemic Buffer impairment because it trains the capacity of the GB to hold written word representations in memory, while the anagram is solved. Previous studies using this procedure have shown improvements in treated word spelling, with some generalised improvement, particularly to sections of untreated words that are similar to the treated words (Raymer, Cudworth, and Haley, 2003). In the current study, one set of words was used for training and another three sets of words were reserved to test generalisation: One set comprised words that were orthographic neighbours of the treated items with shared middle letter positions, though could contain distinguishing letters at the start or end of the words (e.g. clock-block, or break - bread), the second set included orthographic neighbours of the treated items who were distinguished by letter changes at the initial or final positions (e.g. clock-click), and the third set involved words that bore no neighbourhood relation to words from the trained set (e.g. clock-puppy). Is there evidence that generalised improvement to orthographic neighbours would be greater than to unrelated words, and that untreated neighbours with shared middle letter sections might particularly benefit?

Both spelling and reading in English typically show a bias in the report of letters as a function of position. In reading there can be greater priming from the external than the internal letters in strings (Humphreys, Evett & Quinlan, 1990). In spelling, the initial letters are often reported well and there can be a bow shape function in which the middle letters are produced/reported less well (e.g. Wing and Baddeley, 1980, Jensen, 1962). Competitive Queuing models of GB disorders are able to accommodate these primacy and recency effects in serial position curves in spelling conditions (e.g. Glasspool, Shallice and Cipolotti, 2006, Glasspool, Shallice and Cipolotti, 1999). This is due to a number of factors: i) items earlier in the sequence are held to have higher absolute activation, leading to better separation of consecutive items, ii) throughout the sequence, when items are activated and selected they are subsequently inhibited, reducing errors at the terminal sequence positions and iii) where models include dynamic cueing of serial position, end effects occur because items at the start or end of a sequence have fewer adjacent positions than items occupying central positions (Glasspool, Shallice and Cipolotti, 2006).

The idea that initial and terminal letters in a sequence may be at an advantage relative to other words in a sequence could have several consequences for rehabilitation. For instance, if activation strength is greater for initial letters, while end-position letters benefit due to a linear reduction in competitors from a pool of letters for a given item, direct training on a set of words will benefit middle position letters more so than elsewhere in the word. In the case of direct training of a restricted word set, learned representations of letter combinations at the middle positions may have a stronger top-down impact relative to those at other positions, as letters at these positions are less assisted by activation and inhibition processes, due to the nature of the type A GB deficit. This could lead to facilitative effects of training to untreated orthographic neighbours with the same letter combinations at the middle positions (those defined by an initial or terminal position change), but deleterious effects of training to untreated orthographic neighbours with distinguishing letters at the central positions, with obvious implications for rehabilitation. This possibility that learned representations may increase competition for untreated neighbours that take a letter change at the middle section when treating a type A GB deficit was not tested by Sage and Ellis (2006).

In addition to examining effects of orthographic neighbourhood, the present study is also of interest because it examines rehabilitation in a dementia case. Haslam, Kay, Tree and Barona (2009) showed that GB deficits were present in the three patients presenting with dementia that they studied. Two of their patients exhibited performance consistent with a GB impairment in delayed copying, suggesting GB damage independent of problems at the level of orthographic representations. However, no studies have investigated the efficacy of using spelling rehabilitation to overcome GB deficits in these patients. Can spelling be improved in such patients with degenerative deficits?

Case Description

JF is a 64 year old, right handed male, who pre-morbidly worked as a self-employed mechanical engineer and also managed his own consultancy firm. When he was first admitted to the centre, he presented with a three-to-four year history of problems in spoken and written language, with no associated memory or behavioural symptoms. JF was given a provisional diagnosis of primary progressive aphasia. A structural MRI scan revealed bilateral posterior atrophy, particularly around the region of the left intra-parietal sulcus, and minimal changes in the medial temporal regions. A more recent scan conducted in February 2009 revealed new small lesions in superior and inferior parietal gyri and a small lesion in the rolandic operculum when compared with the earlier scan (Figure 3.1). A coronal image from JF's standard scan is provided in Figure 3.2, which indicates bi-parietal atrophy. This result, coupled with his emerging visuospatial problems outlined below, suggests that JF presents with Posterior Cortical Atrophy.

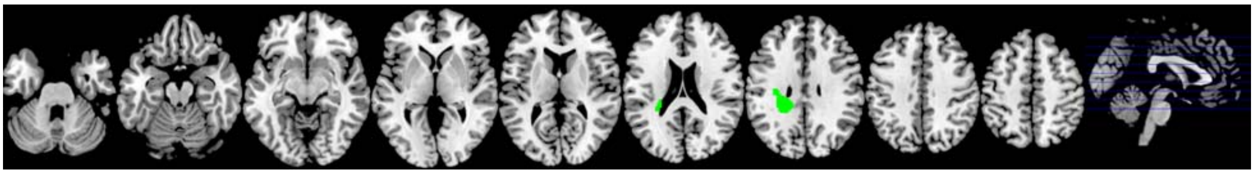


Figure 3.1. JF brain images from the recent (2009) scan, uncorrected. N.B. Grey matter lesion appears in red and white matter lesion in green. The lesion was created in SPM and added as an overlay onto a standard multi-slice template in MRIcron. The SPM analysis was a one sample t- test with the covariates healthy (140 brains aged 40+) vs. patient, age and gender.

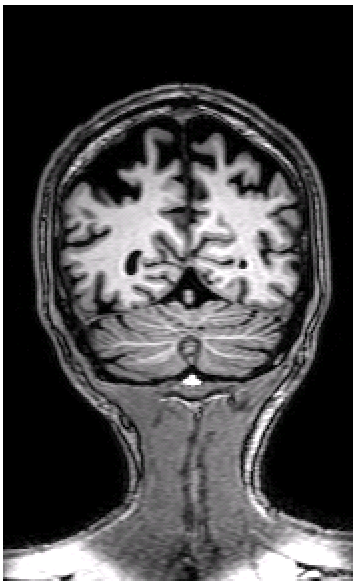


Figure 3.2. Coronal image from JF's standard clinical scan

Though marred by word finding difficulties, JF's free speech was grammatically well-formed and contained no semantic errors. He performed well on tests of auditory processing (72/72 on discrimination of word and non-word pairs, PALPA 1 and 2, Kay, Lesser and Coltheart, 1992) and real word reading (36/36 on sentence reading, PALPA 37, Kay et al., 1992), although there were difficulties with reading nonwords (11/24, PALPA 36, Kay et al., 1992). JF also showed some signs of degraded semantic knowledge. Of the semantic processing tests, his most impaired performance was on the tests of semantic association, which required JF to select a word that is closest in meaning to the target word amongst a range comprising the target and four foils (one semantically related and 3 unrelated, PALPA 51). See Table 3.1 for a summary of his results.

Table 3.1. JF's performance on language processing tests relative to normal cut-off scores.

Task	***Published Normal Cut-off scores***
Auditory processing	
Discrimination of minimal pairs (PALPA 1 and 2)	
Words: 72/72	*M* = 35.19, *SD* = 1.68
Nonwords 72/72	
Word – rhyme judgements (PALPA 15)	
Auditory: 23/30	Norms not available
Written: 26/30	
Phonological segmentation of final sounds (PALPA 17)	
Words: 30/30	*M* = 29.29, *SD* = 0.69
Nonwords: 15/15	*M* = 14.17, *SD* = 1.20
Reading	
Mirror-reversed letter discrimination (PALPA 18)	
34/36	*M* = 35.44
Cross-case matching of single letters (PALPA 19 and 20)	
Upper to lower: 26/26	
Lower to upper: 26/26	
Sentence reading (PALPA 37)	
36/36	
Nonword reading (PALPA 36)	
3-Letter 5/6	*M* = 5.77, *SD* = 0.71
4-Letter 2/6	*M* = 5.89, *SD* = 0.43
5-Letter 2/6	*M* = 5.57, *SD* = 0.90
6-Letter 2/6	*M* = 5.65, *SD* = 0.85
Tests of semantic processing	
Auditory synonym judgement (PALPA 49)	
High imageability: 28/30	Norms not available
Low imageability 26/30	
Semantic association (PALPA 51)	
High imageability: 8/15	*M* = 13.43, *SD* = 1.26
Low imageability: 6/15	*M* = 12.25, *SD* = 1.82
Picture Naming × Oral Reading (PALPA 53)	
39/40	*M* = 39.80, *SD* = 0.35
Pyramids and Palm Trees (Howard and Patterson, 1992)	
3 pictures: 49/52 (89%)	*M* = 98-99%
2 pictures, 1 word: 46/52 (83.63%)	Norms not available

Spelling to dictation

JF's handwritten production of letters was generally good, resulting in well-formed letters indicating no additional impairment of graphomotor skills. In terms of whole-word accuracy, JF demonstrated effects of length, imageability, frequency and orthographic neighbourhood size (see Table 3.2), along with dramatically reduced performance in delayed relative to direct copying, JF's spelling performance was assessed across written and oral response modalities. Though whole-word accuracy scores were at floor performance, qualitative analysis of JF's responses showed similar error types, where often most of the word was preserved but an error occurred in a medial position (e.g. bonus = botus). The roughly equivalent performance across oral and written spelling conditions renders an allographic disorder unlikely.

Table 3.2. JF's performance on spelling tests relative to normal cut-off scores.

Task	***Published Normal Cut-off scores***
Copy: Effect of delay	
Direct copy: 69/72 (95.83%)	Norms not available
Delayed copy: 33/72 (45.83%)	
Transcoding	
Upper to lower case: 50/50	Norms not available
Lower to upper case: 50/50	
Spelling to dictation: Letter length (PALPA 39)	
3-letter 5/6	Norms not available
4-letter 3/6	
5-letter 2/6	
6-letter 0/6	
Spelling to dictation: Imageability × Frequency (PALPA 40)	
HIHF: 6/10	$M = 9.68$, $SD = 0.67$
HILF: 2/10	$M = 9.25$, $SD = 0.75$
LIHF: 2/10	$M = 9.11$, $SD = 1.37$
LILF: 3/10	$M = 8.36$, $SD = 1.97$
Spelling to dictation: Grammatical class (PALPA 41)	
Nouns: 1/5	$M = 4.79$, $SD = 0.42$
Verbs: 2/5	$M = 4.82$, $SD = 0.39$
Adjectives: 3/5	$M = 4.82$, $SD = 0.48$
Functors: 0/5	$M = 4.68$, $SD = 0.55$
Spelling to dictation: neighbourhood (Lavidor and Ellis, 2002)	
Low neighbourhood: 14/60	Norms not available
High neighbourhood: 29/60	
Spelling to dictation: Written × Oral response	
Written: 6/20	Norms not available
Oral: 3/20	

Analysis of error types on a corpus of 468 words (letter length ranging from 3-6) revealed that the majority (85.9%) of JF's errors were due to substitution, insertion, deletion and transposition, although the most frequent error produced was that of deletion (Table 3.3). The finding that most errors could be categorised in this way is consistent with reports of both type A and type B graphemic buffer patients.

Table 3.3: Distribution of error types in spelling 3, 4, 5, 6 and 7-letter words. N = 468 N.B. Transposition = order errors on two target letters, Substitution = a target letter is replaced by a non-target letter, Deletion = the omission of a target letter, Addition = insertion of a non-target letter in the word.

Error type	***Frequency***	***Percentage of errors***
Transposition	45	11.7%
Substitution	91	23.6%
Deletion	167	43.4%
Addition	27	7.03%
Combination errors (e.g. substitution and transposition)	65	16.5%
Orthographically related real word substitutions	11	2.7%
Unrelated real word substitutions	4	1.01%

Caramazza and Miceli's (1990) guidelines for scoring accuracy across serial positions were employed: For substitution responses, an error was documented at the point where the error occurred. A deletion error was counted at the position from which the deleted letter came from. Addition errors received half an error point in letter positions before and after the insertion, and transposition errors were scored as one point for each of the exchanged letters. With this scoring JF showed a bow-shaped accuracy curve, with the medial positions of words being most vulnerable (when spelling responses of various letter lengths were normalised to 5 letters according to Wing and Baddeley, 1980, Figure 3.3).

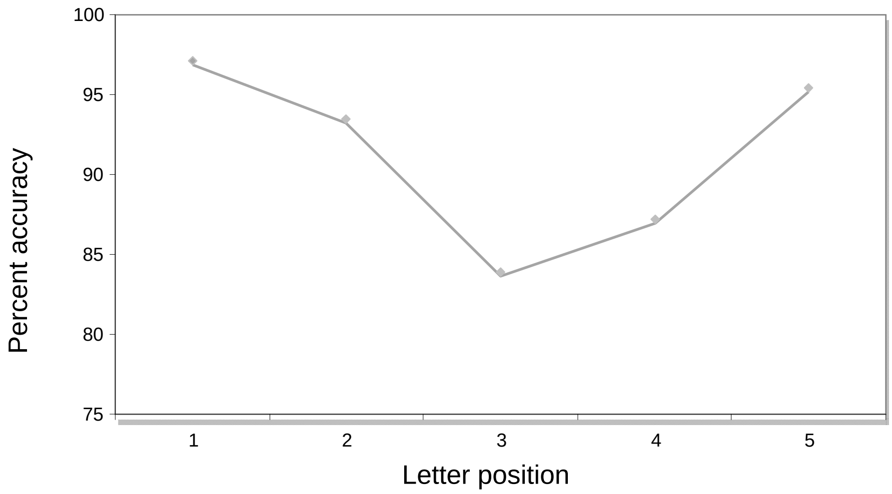


Figure 3.3. Accuracy as a function of letter position summing across letter length (normalised using Wing and Baddeley, 1980). N = 468.

Relative to position 1 there was a drop in accuracy at positions 2, 3, and 4 ($X^2(1)$ = 27.64, 5.02 and 22.0 respectively, all *p*'s < .025). There was no significant difference between accuracy at positions 1 and 5 ($X^2(1) = 3.801$, $p > .05$). The magnitude of JF's bow shaped accuracy curve over position was smaller than in other Graphemic Buffer patients reported in the literature: JF showed a 13% drop in accuracy from positions 1 and 3. Compared with a ranged of 22% - 31% in patients JRE, RSB and BWN (Buchwald and Rapp, 2004), though JF's remediation in accuracy between positions 3 and 5 was within the range elicited by Buchwald's patients (12% compared with the 10-14% range reported in Buchwald and Rapp, 2004).

The influence of lexical variables on JF's spelling performance

JF's spelling performance on real and nonwords was compared, when accepting all plausible spellings for the nonword items. JF demonstrated a clear effect of lexicality on spelling performance, scoring 62.3% on real words and 23.4% on nonwords ($X^2(1)$

= 54.442, $p < .0001$). Similar results have been found in other GB patients (AM: de Partz, 1995, FM: Tainturier and Caramazza, 1996, and AZO (Miceli, Capasso, Ivella and Caramazza, 1997), and is frequently explained in terms of an impaired nonword spelling route as well as GB damage.

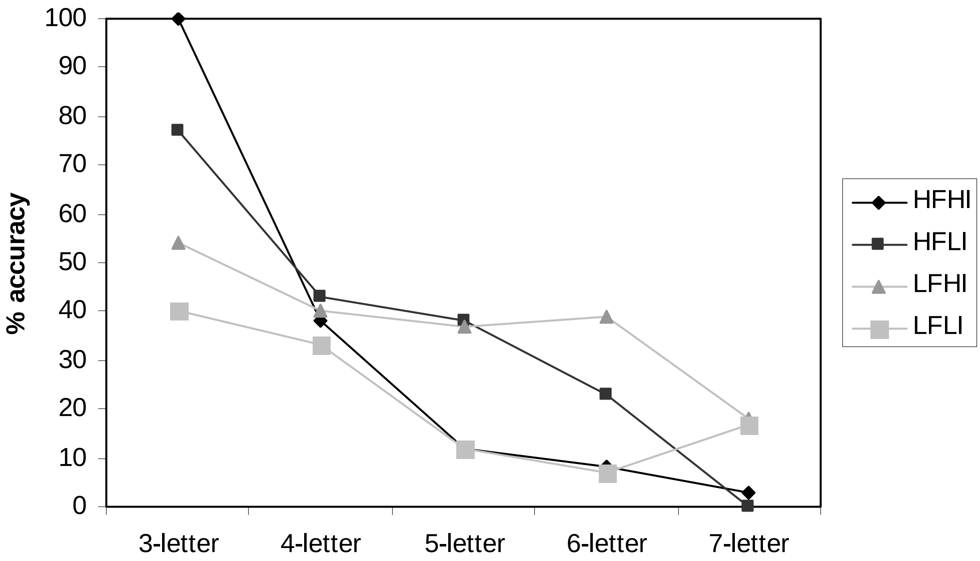


Figure 3.4. Accuracy as a function of imageability and frequency across letter-lengths. Overall N = 430.

To further evaluate the effects of imageability, frequency and length, JF's spelling was tested on a larger word set (N = 10/25 in each condition, see Table 3.4 for the raw scores and individual set sizes).

Table 3.4. Percentage whole word accuracy and N for the imageability × frequency × length stimuli

Letter length	*LFLI*	*LIHF*	*HILF*	*HIHF*	*Mean score*
3	40% (*N*=10)	100% (*N*=10)	53% (*N*=15)	75% (*N*=15)	67%
4	32% (*N*=25)	36% (*N*=25)	40% (*N*=25)	44% (*N*=25)	38%
5	12% (*N*=25)	12% (*N*=25)	32% (*N*=25)	32% (*N*=25)	22%
6	8% (*N*=25)	8% (*N*=25)	32% (*N*=25)	24% (*N*=25)	18%
7	12% (*N*=25)	4% (*N*=25)	12% (*N*=20)	0% (*N*=10)	7%
Mean score	20.8%	32%	33.8%	35%	

A loglinear analysis was conducted with the factors letter length (3-7), imageability (low and high) and frequency (low and high). A significant three-way interaction between accuracy, frequency and length was returned ($X^2(1) = 13.981$, $p = .007$), and a borderline interaction between imageability and accuracy ($X^2(1) = 3.849$, $p = .050$). The accuracy data are shown in Figure 3.4. The accuracy × frequency × length interaction was due to a better performance for high than low frequency words ($X^2(1) = 4.784$, $p = .029$) particularly across the middle letter positions. In addition, across the different lengths and frequencies, performance tended to be better on words with high relative to low imagery ($X^2(1) = 5.263$, $p = .022$, when summing across frequency and lengths 5 and 6).

JF showed a highly significant effect of regularity in spelling (regular words 63.3%, exception words 22.3%, $X^2(1) = 121.665$, $p < .0001$). Regularity effects in GB patients have been reported in a study that attributed the effect to post-lexical processes as exception words were not regularised by their patient (Annoni, Lemay, de Mattos Pimenta, and Roch Lecours 1998). Inspection of JF's errors showed that irregular grapheme combinations frequently remained in the responses, though they were often subject to a GB error (e.g. could = cuold). JF made no regularisation errors.

JF also demonstrated an effect of orthographic neighbourhood size (low neighbourhood *M* = .64, high neighbourhood *M* = 5.84, Figure 3.5).

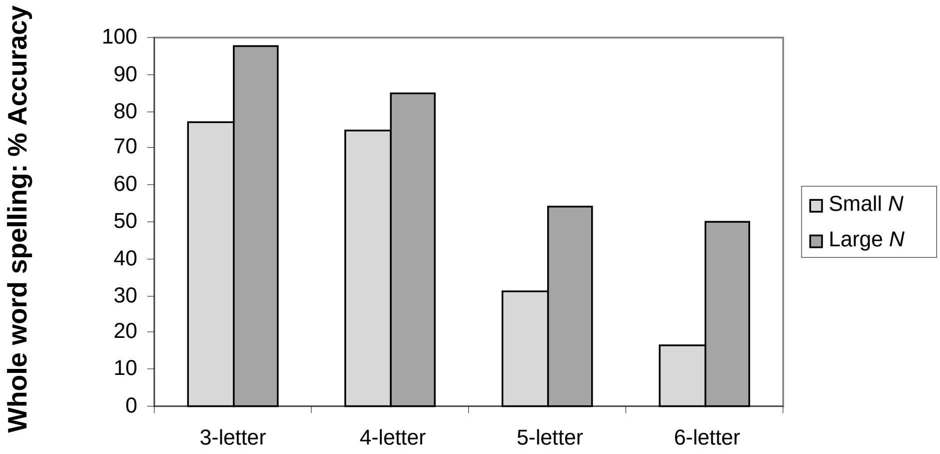


Figure 3.5. Accuracy as a function of imageability and frequency across letter-lengths. Overall N = 430.

A loglinear analysis was performed with the factors letter length (3-6), and orthographic neighbourhood (small and large). Significant two-way interactions were found between accuracy and length ($X^2(1) = 80.851$, $p < .0001$), and accuracy and neighbourhood size ($X^2(1) = 17.754$, $p = .003$, Figure 3.4). The advantage for words with large *N* was maintained across the word lengths. Similar results were reported in patient BH detailed in Sage and Ellis (2006).

Summary

JF presents as a GB patient in terms of his error types, his pattern of effects over word length and his bow-shaped serial position curve in spelling. He also shows a marked decline in delayed relative to immediate copy conditions, suggesting difficulties in maintaining orthographic representations independent of problems associated with word representations arriving at the buffer. However, JF's performance does not easily fit into the type A or type B classifications of graphemic buffer disorders. He demonstrated the lexical effects in spelling that are characteristic of type B GB patients (lexicality, frequency, imageability, regularity and neighbourhood size), but the bow-shaped serial position curve observed in type A patients. Similarly, an advantage for real word over nonword spelling has been found in individuals characterised as type A GB patients (FM: Tainturier and Caramazza, 1996, and AZO (Miceli, Capasso, Ivella and Caramazza, 1997). The influence of lexical factors and a failure to spell nonwords are consistent with a deep dysgraphic profile. However, JF made no lexical or semantic errors in writing (cf. Cipolotti, Bird, Glasspool and Shallice, 2004). Therefore, JF can be characterised as an intermediate GB patient, showing aspects of both the type A and type B performance. He also presented with primary progressive aphasia, with increasing word finding difficulties in free speech along with deteriorating visual spatial attention.

Following the review of JF's basic language and spelling abilities, he entered into a rehabilitation study aimed at improving his spelling. Given the presence of an orthographic neighbourhood effect on his writing, the study focussed on whether this could be harnessed to produce generalisation in learning, if words shared letters. The study also tested if the position of the overlap was important for generalisation.

Method

Participants

Patient JF was entered into the rehabilitation study.

Materials

Seventy words that were spelled incorrectly on a previous occasion at the pre-therapy baseline were entered into the treatment. Orthographic neighbours for each word were found and reserved as untreated items. There were three sets of untreated items, shared middle neighbours (clock-block), different middle neighbours (clock-click), and unrelated items (clock-puppy). Neighbours sharing middle sections with the treated words were added to the intact middle untreated set, neighbours with different middle sections were added to the different middle sections untreated set. An untreated-unrelated set was also used. This comprised of words that were not in the treated set and were not neighbours of those words. Each untreated set were matched as far as possible in terms of set size, and also on a range of lexical variables (CELEX frequency, log frequency, mean number of orthographic neighbours, imageability and regularity, Table 3.5).

Table 3.5. Mean scores on lexical variables across stimuli sets

	Treated items	Untreated orthographic neighbours with shared middle sections	Untreated orthographic neighbours with different middle sections	Untreated unrelated items
N	70	105	92	95
CELEX frequency (mean)	74.91	18.10	13.81	23.31
Log Frequency (mean)	1.27	.84	.63	1.06
Word length	5.26	5.30	5.28	5.40
Mean *N* of orthographic neighbours	4.32	4.86	3.91	3.30
Imageability (Bristol/MRC)	362.01	266.55	199.38	235.95
% Regular	58.5	52.3	64.5	66.6

It was also important to control for consonant-vowel structure in the treated set, given the potential for regularity differences at different positions: e.g. if beginning and end consonants have regular consonants, while middle positions feature vowels (that have a limited set of alternatives), strengthening of the middle positions could be due to strengthened phoneme-grapheme conversion for vowels rather than strengthened whole word or letter combinations per se. Therefore, the treated set contained 38 words with vowels at the medial positions, and 31 contained consonants at the middle positions. The remaining item was a 6-letter word involving both a consonant and a vowel at the middle section (upheld).

Procedure

The current method employed an Anagram Copy Procedure (ACT). Anagrams of the treated items were presented to JF in Arial font size 14, and the experimenter read out the target word. JF was instructed to make the given word from the presented anagram. JF verbally pronounced each letter in the sequence before he wrote it down, If JF read out a letter out of sequence, the examiner advised him that it was incorrect, and invited JF to select another letter from the anagram. When JF solved the anagram, he was asked to copy the item directly. Once copied successfully, JF was required to reproduce the word after a 10-second delay. If JF committed any error in his response, including a correct letter in the incorrect order, he was corrected online, shown the item again, and asked to reproduce the word again after a 10-second delay. When JF was successful in his delayed copy, the next anagram was presented and JF repeated the procedure again. All baseline spelling tests were conducted using spelling to dictation. The intervention was performed across 10 once-weekly 1.5 hour sessions, and the performance on the untreated sets was assessed over three 1.5 hour sessions conducted before and after the treatment phase.

Analysis

During the rehabilitation and baseline stages, as well as recording correct and incorrect responses and percent accuracy within items at each stage of the procedure, information pertaining to serial position was also documented: To explore whether improvement was produced across all positions, or if it was specific to some positions (e.g. whether the middle positions or first and last positions improve more), the number of target letters that appeared anywhere in JF's spelling was noted, as well as the serial position where the letters appeared. As with the baseline assessments, Caramazza and Miceli's (1990) guidelines for scoring accuracy across serial positions were followed throughout. Substitution responses counted one error at the point where the error occurred. A deletion error was counted at the position of the omission, and transposition errors were scored as one point for each of the exchanged letters. Due to the restrictive nature of the anagram items, insertion errors were not possible during the anagram copy treatment. The frequency of each category of each error type produced before and after therapy was noted to provide additional evidence on how any rehabilitation effect might come about.

Results

Whole word accuracy on the treated word sets improved from 10/70 pre-therapy to 69/70 post-therapy (McNemar $p < .0001$, Figure 3.6). The post-therapy ceiling performance suggests that the treated word representations were well-learned.

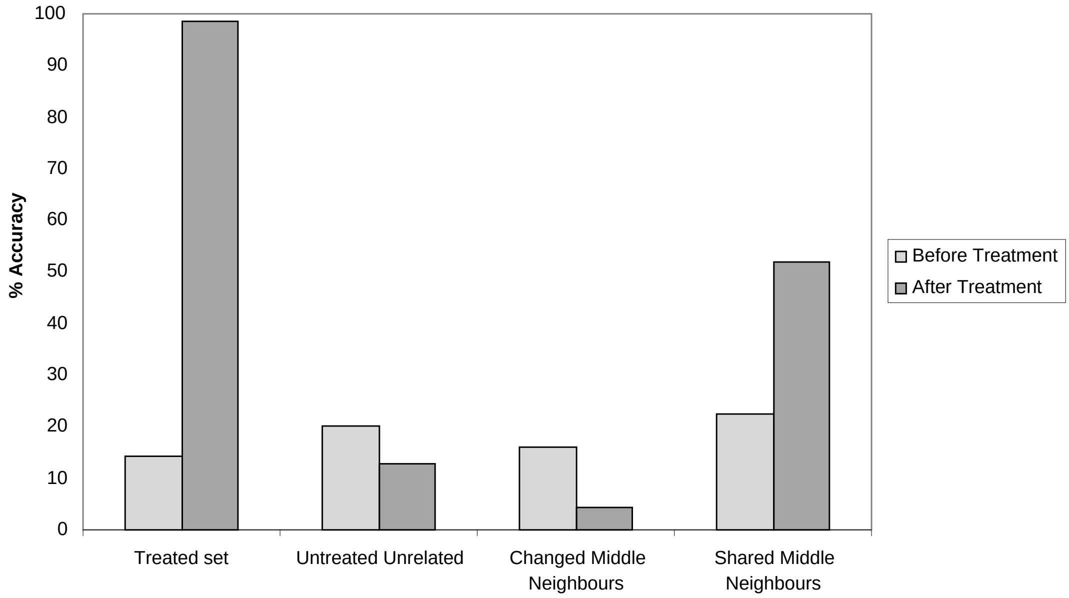


Figure 3.6. Whole word accuracy on treated and untreated sets before and after treatment

Performance on the treated and untreated sets did not differ overall before therapy ($X^2(1) = 1.614$, $p = .204$). Whole word accuracy data on the untreated sets are provided in Figure 3.5. There was no significant change on the untreated unrelated set following treatment (McNemar 14/70 vs. 9/70, $p > .05$). There was a significant decline in performance on untreated neighbours of the treated items taking a letter change at the medial position (McNemar 11/68 vs. 3/68, $p = .008$), a finding returned to in the Discussion. The only significant generalised improvement post-therapy was on orthographic neighbours with intact middle sections (McNemar 19/86 vs. 44/86, $p < .0001$).

Accuracy across letter position

For each untreated set, a loglinear analysis was performed on the accuracy data. As the words were comprised of 5-6 letter words, the results were summed across lengths to normalise the data. Each analysis contained three factors: Letter report accuracy (correct / incorrect), position in word (1-5, using Wing and Baddeley, 1980), and baseline (before / after).

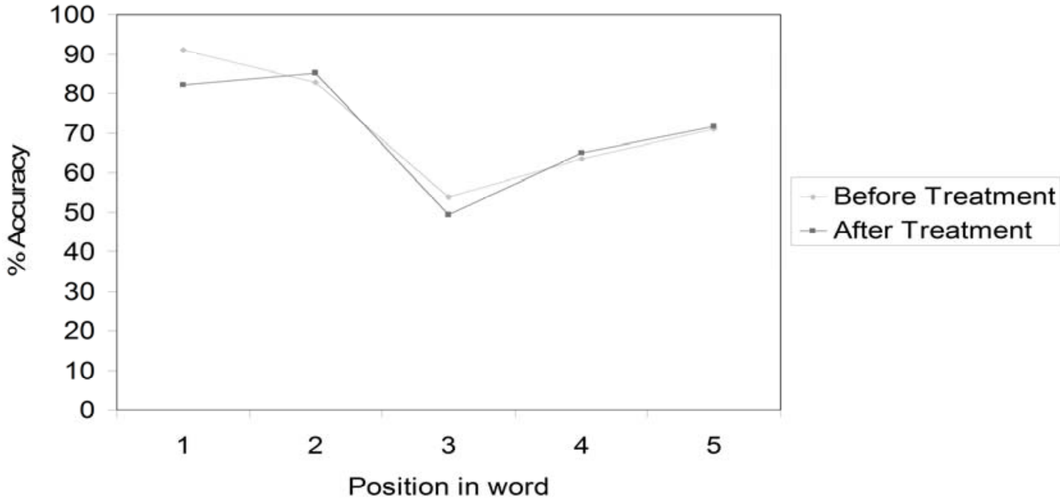


Figure 3.7. Percent accuracy across letter position on the untreated unrelated set.

For the untreated unrelated set, there was a significant interaction between accuracy and position, reflecting the advantage for other positions over the medial section (X^2 (4) = 87.625, p < .0001), but no effect of baseline (p = 1.0), and these factors did not interact (accuracy, baseline test and position: X^2 (4) = 2.860, p = .582). These results show that accuracy at the central letter positions was lower than at other positions in the word, and that this pattern did not change after therapy. The absence of a baseline effect shows that letter accuracy did not change after treatment (Figure 3.7).

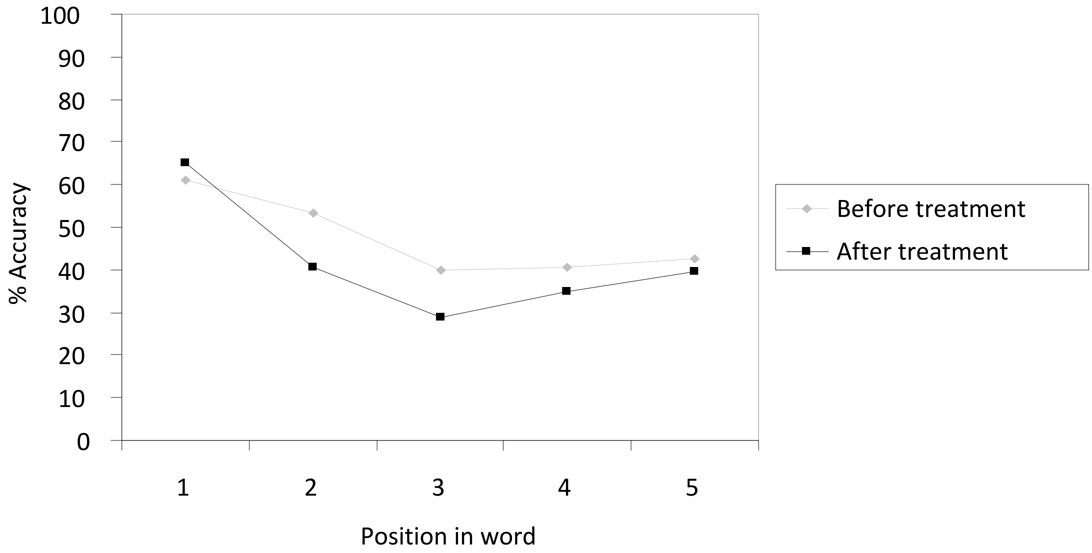


Figure 3.8. Percent accuracy across letter position on the untreated changed middle section set.

Analysis of the untreated changed middle section neighbours returned significant two-way interactions between accuracy and position (X^2 (4) = 98.933, p < .0001) and accuracy and baseline (X^2 (1) = 5.005, p = .025). However, the effect of baseline reflected a *decline* in performance post-therapy (see Figure 3.8). As with the untreated unrelated analysis, the untreated changed middle position neighbours data showed no significant three-way interaction between baseline accuracy, baseline and position (loglinear, X^2 (4) = 6.323, p = .176); the bow-shaped serial position curve did not alter post-treatment.

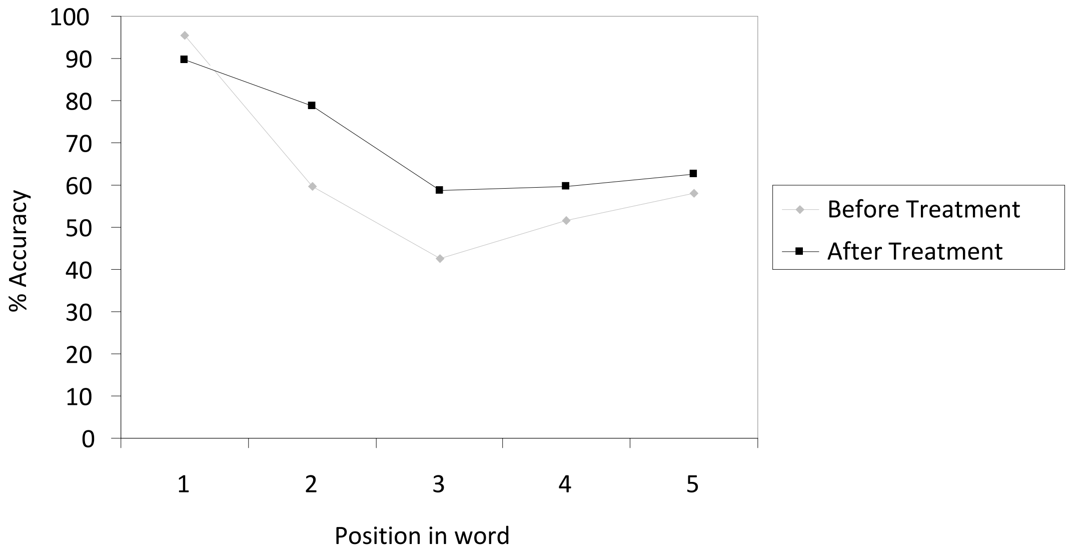


Figure 3.9. Percent accuracy across letter position on the untreated intact middle section set.

When letter accuracy scores for the intact middle section neighbours of the treated items were analysed, there was a three-way interaction between accuracy, baseline, and position (X^2 (4) = 10.326, p = .035). This result indicates a change in accuracy patterns over word position post-therapy. For positions 1, 4 and 5 there was no treatment effect (X^2 (1) = .270, p = .874). For positions 2 and 3 there was a treatment effect (X^2 (1) = 13.441, p = .001, Figure 3.9).

Error types

Error types were assessed before and after the treatment phase. Since there was a detrimental effect of treatment on untreated changed-middle section neighbours post-therapy, errors were inspected for neighbourhood competition. The number of errorful responses that bore similarity to a competing neighbour were tallied and contrasted with the frequency of other error types in the pre- and post- treatment data. *N* – similar errors were defined as responses containing the letter distinguishing an item from its treated neighbour. Therefore, *N*-similar errors were not always whole-word *N* substitutions (e.g. cattle = castle), but did always contain the distinguishing letter between the target and the treated neighbour (e.g. cattle=catsle). As well as transpositions, responses that contained a letter belonging to a competing neighbour that were subject to deletion errors elsewhere in the word were included (e.g. cattle = catsl). This was done to verify that the letter insertion came from the treated neighbour. Any responses that did not contain this distinguishing letter, or contained more than one letter insertion were categorised as *N* – dissimilar responses. A loglinear analysis was performed on the raw data, with the factors error type (*N* - similar / *N* – dissimilar), baseline (before treatment / after treatment), neighbourhood set type (changed middle section / intact middle section), and accuracy (correct or incorrect).

Table 3.6. Proportion of errors similar to orthographic neighbours relative to other errors before and after treatment for the untreated neighbourhood sets. Untreated changed middle section neighbours, left panel, intact middle section neighbours, right panel,

		Before therapy (expressed as a proportion of total errors)	*After therapy (expressed as a proportion of total errors)*
Untreated changed middle section neighbours	Error similar to a treated neighbour	.30	.52
	Other error	.75	.41
Untreated intact middle section neighbours	Error similar to a treated neighbour	.39	.9
	Other error	.68	.71

Table 3.6 shows the amount of each type of error, as a proportion of total errors, before and after treatment for both untreated neighbourhood sets. When the data were entered into a loglinear analysis, a highly significant three-way interaction

between error type (similar to treated N versus other error), baseline (pre-therapy versus post-therapy), and neighbourhood set (changed middle versus intact middle) was found ($X^2(2) = 15.875$, $p < .0001$). To decompose this interaction, we entered the data for the untreated changed middle section neighbours, and untreated intact middle section neighbours into separate loglinear analyses.

For the loglinear analysis of the untreated changed middle section neighbours, the factors were baseline, accuracy, and error type. The analyses produced a significant two-way interaction between error type (similar to treated N and other error) and baseline ($X^2(2) = 8.886$, $p = .012$). Chi Square analyses of the error types in this set before and after treatment revealed that, proportionately, more N-similar errors were made after treatment ($X^2(1) = 10.004$, $p = .002$), and fewer 'other' errors ($X^2(1) = 25.064$, $p < .0001$).

For the untreated shared middle section neighbours, a significant 2-way interaction between error type, and baseline was found $X^2(2) = 7.594$, $p = .022$. There was a decrease in N-similar errors post therapy ($X^2(1) = 25.976$, $p < .0001$), while there was no change in the number of 'other' errors pre vs. post therapy ($X^2(1) = .024$, $p =.878$).

Consistency of general performance

JF was a progressive aphasic case, and so it was crucial to provide some measure of JF's general performance before and after therapy. The Birmingham Cognitive Screen (BCoS, Humphreys, Riddoch, Samson and Bickerton, in press) was employed. This is a battery comprised of a range of short assessments of cognition (e.g. memory, language, problem-solving). Summarised methodological details for each task are provided in Appendix E. Impaired performance is defined as any score more than 3 standard deviations away from the control mean for each condition. JF's data were compared with the scores from a group of 34 age-matched control subjects. Broadly, his presentation fits with an emerging visuospatial deficit, as is consistent with his biparietal atrophy (Figure 3.2). His performance across tasks before and after treatment is given in Table 3.7.

*Table 3.7. JF's performance on the Birmingham Cognitive Screen (BCoS) in 2009 and 2010. N.B. * denotes tasks where performance changed from spared in 2009 to impaired in 2010.*

	Pre-treatment, 18/01/2009	Post-treatment, 16/07/2010	Difference
LANGUAGE			
Picture Naming*	Spared 14/14	Impaired 10/14	-4
Sentence Construction*	Spared 8/8	Impaired 5/8	-3
Instruction Comprehension	Spared 3/3	Spared 3/3	Ceiling
Reading	Sentences Spared 42/42 Nonwords Impaired 0/6	Sentences Spared 41/42 Nonwords Impaired 0/6	-1 Floor
Writing	Impaired 1/5	Impaired 0/5	Floor
PRAXIS / CONTROL AND PLANNING OF ACTION			
Visuo-Constructional Abilities (Complex Figure Copy)	Impaired 23/47	Impaired 20/47	-3
Multi-Step Object Use	Spared 12/12	Spared 11/12	-1
Gesture Recognition	Spared 4/6	Spared 4/6	0
Gesture Production	Spared 10/12	Spared 11/12	+1
Meaningless Gesture Imitation	Impaired 5/12	Impaired 8/12	+3
LONG-TERM MEMORY			
Orientation*	Spared *personal information* 7/8 *time and space* 6/6	 *personal information* spared 7/8 *time and space* Impaired 5/6	 0 -1
Episodic Memory (Newly Acquired Knowledge)	*verbal memory* Impaired immed. recall 2.5/15 immed. recogn. 11/15 delayed recall 3.5/15 delayed recogn. 11/15 *task recognition* spared 8/10	*verbal memory* Impaired immed. recall 1/15 immed. recogn. 11/15 delayed recall 1.5/15 delayed recogn. 7/15 *task recognition* Spared 9/10	 -1.5 0 -2 -4 +1
Spatial Attention	*cancellation tasks* asymmetry scores Impaired key -8 *extinction tasks* visual Spared L unilateral 4/4 L bilateral 8/8 R unilateral 4/4 R bilateral 8/8 tactile Impaired L unilateral 3/4 L bilateral 8/8 R unilateral 1/4 R bilateral 8/8	Spared *cancellation tasks* asymmetry scores key 0 *extinction tasks* Spared visual L unilateral 4/4 L bilateral 8/8 R unilateral 4/4 R bilateral 8/8 tactile L tactile Spared L unilateral 4/4 L bilateral 4/8 R unilateral 4/4 R bilateral 8/8	NA
Controlled Attention	Impaired *auditory attention* accuracy 46/54 word recalled 1/3 *rule finding and switching* accuracy 1/18, rule 0/3	Impaired *auditory attention* accuracy 33/54 word recalled 1/3 *rule finding and switching* accuracy 1/18, rule 0/3	 -13 0 0
MATHEMATICAL /NUMBER ABILITIES			
Number Reading	Impaired 6/9	Impaired 3/9	-3
Number Writing*	Spared 5/5	Impaired 0/5	-5
Calculation	Impaired 1/4	Impaired 0/4	Floor

JF's performance deteriorated on a number of measures from 2009 to 2010. Notably, many of the tasks where a decline was observed involve verbal production (e.g. picture naming, sentence construction and orientation). However, some abilities remained preserved (sentence reading, gesture recognition and production, object use, and spatial attention). Interestingly given the degenerative nature of his presentation, performance on the spatial attention assessments improved post-therapy. The improvement on the cancellation task could be due to the visual aspects of the treatment, although this does not explain the improved performance on tactile extinction.

Discussion

The current study found that an error-reducing Anagram Copy Treatment was successful in improving the spelling of seventy treated words in a GB patient. Training was successful in producing a dramatic improvement on a fairly substantial set of treated words, and these effects were produced in a patient presenting with a degenerative condition. The treatment used a closed set of items and also an ordering procedure, and therefore may have operated by both training word representations and improving ordering and retention capacities at the graphemic buffer. The therapy produced generalised treatment effects to untreated orthographic neighbours of the treated items where letters in the middle sections remained constant. In contrast the treatment had no impact on unrelated, untreated words and it actually had a deleterious effect on untreated orthographic neighbours with overlap at the start and end (rather than middle) positions. At least part of this deleterious effect was due to letters from trained words incorrectly appearing in the response to the untreated words.

It is possible that the treatment could have enhanced the maintenance of graphemes in the buffer, as practice involved holding the stimuli under conditions of support from the anagram. Improved maintenance in the GB should have generalised to all types of untreated word. However, it needs to be borne in mind that treatment was conducted in the context of a degrading condition. A battery of tests designed to test different cognitive abilities (the BCoS, Humphreys et al., in press) suggested deterioration in verbal production, and in one written production modality (number writing, Table 3.5). It may be that this decline offset any potential post-therapy improvement in the untreated unrelated word condition. The finding that word representations were successfully reinforced while therapy failed to produce any more general effect of improving the functioning of the buffer may also indicate that improving word representations in an error-reducing paradigm is more beneficial in this degenerative case than enhancing maintenance in the buffer.

The finding that improvement after training generalized to untreated neighbours, coupled with the pre-therapy advantage for high *N* word sets, to the GB receiving support from other neighbourhood items, which may be consistent with a theory of interactivity between the buffer and the lexicon (e.g. Sage and Ellis, 2006), and are similarly coherent with existing reports suggesting the influence of lexical effects on the spelling of GB patients (e.g. Caramazza and Miceli, 1990; Sage and Ellis, 2004, 2006).

That any generalization at all was observed is inconsistent with strict unitary views of lexical representations (positing no decomposable parts). It may be that treatment, with its emphasis on re-ordering letters, may have reinforced phoneme-grapheme associations within the word. However, the effect is unlikely to be due to learning a finite set of phonemes, as the treated set included the full range of letters occurring at all positions, and were equally comprised of vowels and consonants. At the very least, the finding that this generalized improvement was specific only to untreated words with shared middle sections in the data suggests that the locus of the neighbourhood effect in graphemic buffer patients may be due to consistency at the problematic section of words, and therefore, that large neighbourhoods are not always supportive in these patients.

Thus, the data can also be accommodated within a competitive queuing model (e.g. Glasspool, Shallice and Cipolotti, 2006, Glasspool, Shallice and Cipolotti, 1999) in that the learned representations exerted a disproportionate effect over letter positions. Learned middle sections of words had a stronger impact on untrained items – leading to improved performance at these positions in the untreated sets for the intact middle section neighbours, and increased competition from learned letter combinations at these positions for the changed middle section neighbours. These findings are consistent with the main principles of a competitive queuing model, where activation strength is greater for initial letters, while end-position letters benefit due to a linear reduction in competitors from a pool of letters for a given item and are more distinctive.

No beneficial change was observed at the initial and final letter positions for the changed middle section neighbours, nor a deleterious effect at these positions in the intact middle section set, even in conditions where pre-therapy performance was not at ceiling. Because words with changes at both start and end positions were included in this set, any effects could have been neutralized. However, when separating these data by end and start position consistency, there was still no effect

(Chi Square, $p > .1$). This feature may indicate that lexical representations are weighted differentially over word position, with initial and final positions in the word being more resilient to change. Therefore, a 'position skeleton' provided by the initial and terminal letters in a sequence may remain constant, even in the context of a rehabilitation study. This could produce a 'frame' of a word's start and final positions, with letters at the middle section being determined by activation strength, which may fluctuate as a consequence of learning.

Tests were not conducted of whether improvement was due to learning similar combinations of letters (neighbourhood items) or strengthened lexical representations. However results from a previous priming experiment with GB patients indicates that support from nonword primes may be less facilitative than support from real word primes, though both priming effects could be due to top-down lexical processes (Sage and Ellis, 2006). The current finding, that improvement generalized only to untreated neighbours of treated items with constant middle sections, however, suggests that this top down 'support' can exert a detrimental effect where the most problematic serial position does not remain constant across *N*.

N effects in the literature

Having a large number of orthographic neighbours can be beneficial or disruptive to performance. For instance, lexical decision has been found to be facilitated for words with many neighbours, whereas words with large neighbourhoods impede semantic categorisation (Bowers, Davis and Hanley, 2005). This pattern of findings has been attributed to the requirement of unique identification in the case of semantic categorisation, whereas tasks sensitive to overall lexical activity enjoy a facilitative effect (Grainger and Jacobs, 1996).

Similarly, dissociations have been made in the direction of the *N* effect between tasks requiring serial and parallel processing. For instance, letter by letter readers presenting with highly serial, length-sensitive reading have shown detrimental effects of high neighbourhood words where visual processing demands are high, leading to serial processing, and facilitative effects of high *N* words in words with less visual ambiguity allowing parallel processing (Arguin and Bub, 2005). Little is known about the effect of orthographic neighbourhood on spelling processes, although the spelling process intuitively involves activation of individual graphemic representations for each word over any more generalized operation such as a lexicality judgment. Spelling in the case of impaired ordering processes at the graphemic buffer may be a serial process, and so competing high orthographic neighbourhood words, or those with highly frequent, or well-learned neighbours, may produce a detrimental effect on spelling. However, the findings indicate that where competing neighbours remain intact at the most problematic position in the word, neighbourhoods may be supportive. This is consistent with previous data suggesting that large neighbourhoods are beneficial when uncertainty is low, but problematic when uncertainty is high (e.g. Arguin and Bub, 2005)

Conclusions

Sage and Ellis reported that treating a set of words promoted improvement on orthographic neighbours of the items. Here, generalised improvement was mediated by intact middle section neighbours rather than orthographic neighbourhoods more generally. The finding that performance declined on untreated neighbours of the treated items suggests that therapists should select treated items from neighbourhoods that are i) large and ii) share consistent letter combinations at the problematic middle sections of words, when treating graphemic buffer patients presenting with a bow shaped accuracy curve across letter position. Future work might investigate whether generalised improvement to neighbourhoods is modulated by consistency at the most problematic letter position in GB patients showing different patterns of accuracy across word position (e.g. in Type B GB patients demonstrating the linear decline in accuracy across the word).

Chapter 5: Overcoming the effect of letter confusability in letter-by-letter reading: A rehabilitation study

Abstract

Recent work has suggested that letter-by-letter dyslexic patients demonstrate effects of lexical variables in words comprised of low confusability letters, suggesting that these patients are capable of processing low-confusability words in parallel (Fiset, Arguin and McCabe, 2006). Here a series of experiments is presented investigating letter confusability effects in MAH, a patient experiencing letter-by-letter dyslexia as well as expressive and receptive aphasia, and DM, a relatively 'pure' alexic patient. Two rehabilitation studies were employed: one aimed to promote parallel processing, and another to improve serial reading. The parallel processing (word-level) treatment produced generalised improvement to low-confusability words only, but the serial processing strategy (letter-level) produced improvement on both high and low confusability words. Neighbourhood effects at initial baseline were facilitative with low-confusability words, but deleterious in high confusability sets in both patients. Across the therapies, MAH showed facilitative effects of *N* size on accuracy with high confusability words only after the letter level treatment (Treatment 2). However, for DM, facilitative effects of *N* size on RT emerged after the word level treatment (Treatment 1), and the effects persisted after Treatment 2. The results add support to the hypothesis that letter confusability plays a key role in letter-by-letter reading, and suggest that a rehabilitation method aimed at reducing ambiguities in letter identification may be particularly effective for treating letter-by-letter reading.

Introduction

Letter-by-letter (LBL) reading refers to a difficulty in identifying orthographic information that manifests itself in a pronounced letter length effect in reading (Dejerine, 1892; Geschwind, 1965a and b). Researchers have posited various loci for the disorder, including defective simultaneous processing of multiple objects (Kinsbourne and Warrington, 1962, Levine and Calvano, 1978), impaired access to intact orthographic forms (Patterson and Kay, 1982), disconnection of intact visual processes in the right hemisphere from language areas of the left (Coslett and Saffran, 1989), damaged lexical orthographic representations (an impaired word form system, Warrington and Shallice, 1980), and deficient visuo-perceptual analysis leading either to a general difficulty in processing visual information (Behrmann, Plaut and Nelson, 1980), or a selective impairment for encoding orthographic information (Warrington and Shallice, 1980). Common to these accounts is the idea that there is a low-level deficit in processing letters in parallel across words. For example, within a dual route model of reading (e.g. Dual Route Cascaded: Coltheart, Curtis, Atkins and Haller, 1993, Coltheart and Rastle, 1994 and Coltheart, Rastle, Perry, Langdon and Ziegler, 2001), damage is posited either to the orthographic analysis system, or to the orthographic input lexicon. Both stages are sited before access to word meaning. Similar notions are asserted by other reading models (e.g. the Triangle model: Plaut, McClelland, Seidenberg and Patterson, 1996 and Rosazza, Appollonio, Isella and Shallice, 2007).

However, several studies have shown that, despite the apparent peripheral nature of the disorder, letter-by-letter dyslexics can show sensitivity to lexical properties of words even before serial processing of the letters takes place. For example, studies of implicit reading have found that LBL patients can make above-chance lexical and semantic decisions despite exposure durations that are insufficient for explicit word naming (Shallice and Saffran, 1986; Coslett and Saffran, 1989; Coslett, Saffran, Greenbaum and Schwartz, 1993). Recent work into the effect of letter confusability (a measure of the structural similarity between the capital letters of the alphabet) in these patients has further suggested that there is an activation of lexical representations not based on serial processing. Arguin and Bub (2005) reported interactive effects of confusability and orthographic neighbourhood (*N*) size on reading in three alexic patients: When letter confusability was low there was a facilitatory effect of *N* size; faster reading latencies were recorded for words with high orthographic neighbourhoods. With high confusability words there was no significant effect of *N* size. There is evidence that effects of *N* reflect parallel processing of words. For example, normal participants do not show facilitatory effects of *N* when words are presented in sequential fragments (Snodgrass and Mintzer, 1993). *N* effects in normal participants are also more pronounced with low frequency relative to high frequency words (Andrews, 1989, 1992; Sears, Hino, and Lupker, 1995; Arguin, Bub, and Bowers, 1998), suggesting that the effects emerge when word processing is made more difficult. The positive effects of *N* found with low confusability words in LBL patients indicate that there can be parallel processing of letters in words, but such parallel processing also remains relatively problematic.

Inhibitory effects of high *N* size have also been demonstrated. Pugh, Rexer, Peter and Katz (1994) showed that neighbours had a detrimental effect on reading in normal subjects when the letter distinguishing the item from its neighbours occurred after a 100ms delay. Also, in the neuropsychological literature, a patient presenting with (left) neglect dyslexia showed a detrimental effect of *N* size when a neighbour differed from the target word by the leftmost letters (first two letters in 4-letter words), particularly when the first letter was visually similar to that of a neighbouring item (Arguin and Bub, 1997). Therefore, it may be that under conditions of serial presentation (e.g. delayed letter presentation in normal participants), or processing (e.g. neglect dyslexia or letter-by-letter reading), high *N* words may exert a deleterious effect. Presumably this effect would be especially detrimental when neighbours of a given item differed by a letter at the terminal positions. Given that LBL readers show strong serial components in reading, it is possible that negative effects of *N* could be apparent under some conditions.

Rehabilitation studies

Much of the rehabilitation work reported with letter-by-letter patients has sought to improve residual capacities in parallel processing (Coslett, Saffran, Greenbaum and Schwartz, 1993). However, since pre-lexical deficits may also be present in these patients (Chialant and Caramazza, 1998), improvements may additionally be brought about by improving letter discrimination. Several rehabilitation studies have used letter-level training with LBL patents. Multiple Oral Reading (MOR) methods require the repeated reading of text in order to improve reading speed (e.g. Moyer, 1979; Moody, 1988; Tuomainen and Laine, 1991; Beeson, 1998). Both Tuomainen and Laine (1991) and Moyer (1979) applied this technique to LBL patients, whilst also incorporating training in letter discrimination and identification. The patients' reading rates improved following therapy, but it was unclear whether this treatment effect was due to the letter identification or word-level aspects of the training. Arguin and Bub (1994) more specifically targeted the letter processing capacities of their patient. Their method used speeded same-different letter matching and overt reading of pronounceable letter strings. They found that while identification did not improve at the letter-level, there was a generalised improvement to whole word reading. The authors concluded from this result that a visual processing deficit was a key contributing factor to letter-by-letter reading.

Sage, Hesketh and Lambon Ralph (2005) contrasted the effectiveness of parallel and serial processing treatments in their rehabilitation study of the LBL patient FD. Using an ABACA design the authors applied a word-level treatment (e.g. pairing orthographic representation of a word with its phonological form, as read by the therapist) which aimed to promote parallel processing, and a letter-level treatment which focussed on distinguishing letter forms using kinaesthetic information (tracing letters) and improving letter naming. While both treatments improved accuracy on directly treated sets, only the letter-level therapy produced generalised improvement to control items. At initial baseline, FD showed a preponderance of omission errors, though after the word-level treatment, visual errors dominated. The letter-level treatment reduced the number of visual errors when compared to the word-level treatment. These data suggest that letter-level treatment may be more effective in overcoming letter confusability than a parallel processing word-level approach.

Current study

The literature suggests that, in LBL readers, there can be parallel processing of words under low perceptual demands, along with word recognition based on serial letter processing. To assess whether treatment targeted at parallel processing or serial letter identification is most effective, the present study assessed the relative effects of word and letter-level therapy in two LBL readers, DM and MAH. Following Sage et al. (2005), performance was tested both with treated words and unrelated items, but in addition, effects were assessed with words comprised of high and low confusability letters. It was predicted that effects of word-level treatment would generalise better to low- than high- confusability words, while letter-level training might show generalisation to low- and high- confusability words. Possible interactive effects of confusability and *N* size were also examined.

Case Descriptions

Patient DM

Patient DM, a 55 year old female, experienced left medial and inferior occipito-temporal brain damage as a result of an abscess in her brain caused by multiple arteriovenous malformations (AVMs) in her lungs. DM could not have an fMRI scan as she has Aneurysm clips. She formally worked as an English teacher in a secondary school. Functionally she presented with homonymous hemianopia, alexia without agraphia (Osswald, Humphreys and Olson, 2002) and anomia with a particular impairment in naming living things (Humphreys, Riddoch and Price, 1997). Aside from some occasional naming difficulties, DM's free speech was fluent and grammatically well-formed. There was no evidence of a semantic impairment (Table 4.5). Data on DM's reading performance have been published previously in Cognitive Neuropsychology (Osswald, Humphreys, and Olson, 2002).

Patient MAH

MAH was referred to the Brain and Behavioural Sciences unit at the University of Birmingham in 2008, and was diagnosed with chronic receptive and expressive aphasia after a cerebrovascular accident. His spontaneous speech is frequently incomprehensible, though his written production is much better, and he often uses a writing pad when comunicating. MAH managed the family business in selling non-metallic fixings before he retired, and he had already retired at the time of his CVA. At the time of testing he was aged 76-77 years. Lesion analysis using SPM and MRIcron revealed damage to the left superior and middle temporal gyrus, left superior temporal pole, left insula and the head of left caudate (Figure 4.1).

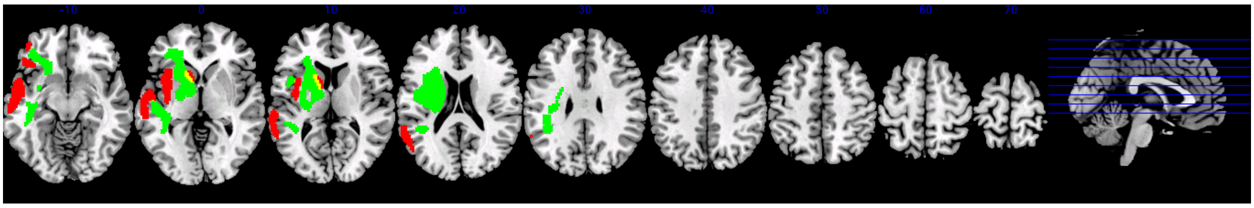


Figure 4.1. Images from MAH's fMRI scan. N.B. Grey matter lesion appears in red and white matter lesion in green. The lesion was created in SPM and added as an overlay onto a standard multi-slice template in MRIcron. The SPM analysis was a one sample t- test with the covariates healthy (140 brains aged 40+) vs. patient, age and gender.

MAH's free speech contained frequent semantic and phonological paraphasias, with frequent attempts to self-correct. Such self-correction tendencies suggest a problem not due to a semantic deficit per se but rather a problem in access to the phonological output system. This was also suggested by his relatively good performance on tests of semantic access, especially with written presentation (see Table 4.5).

Background tests

Tests of visual perception and recognition

Both DM and MAH performed well across a range of tests from the Visual Object and Space Perception Performance Battery (VOSP; Warrington and James, 1991) and the Birmingham Object Recognition Battery (BORB; Riddoch and Humphreys, 1993, Table 4.1). These results indicate preserved abilities in perceiving basic object features, viewpoint-invariant object representations and accessing stored knowledge of object shape.

Table 4.1. DM's and MAH's performance on visual perception and recognition.

	DM	MAH	Published normal cut-off scores
VOSP			
Object perception			
Screening test	18/20	20/20	15
Incomplete letters	20/20	20/20	16
Object decision	18/20	14/20	14
Space Perception			
Dot counting	10/10	8/10	8
Position Discrimination	20/20	20/20	18
Number location	10/10	10/10	7
Cube analysis	9/10	8/10	6
BORB			
Length match	29/30	26/30	24
Size match	29/30	27/30	23
Orientation match	24/30	23/30	20
Position of gap match	34/30	38/40	27
Minimal feature view match	25/25	25/25	19
Foreshortened view match	25/25	25/25	16
Object decision:			
Easy (version A)	28/32	31/32	24
Hard (version A)	32/32	26/32	23
Key cancellation	50/50	50/50	
Apple cancellation	50/50	49/50 (asymmetry score - 1 fewer on left)	

Tests of Auditory Processing

Both patients performed assessments from the PALPA battery (Psycholinguistic Assessments of Language Processing in Aphasia, Kay, Lesser and Coltheart, 1992). DM achieved a spared performance across a range of auditory processing tasks (Table 4.2). MAH was impaired in same-different judgements using nonwords (PALPA 1, Kay, et al., 1992) but not with real word stimuli (PALPA 2, Kay et al. 1992), or with real words when picture selection was required (PALPA 4, Kay et al., 1992) (see Table 4.2). This pattern of performance suggests facilitated auditory discrimination with meaningful stimuli. MAH showed some impairment at making fine auditory discriminations, generally spared auditory lexical access, but strong effects of syllable length on repetition.

Table 4.2. DM's and MAH's scores on tests of auditory processing (denotes impaired performance [> 3 standard deviations of the control mean])*

Tests of auditory processing	**DM**	**MAH**	**Published normal cut-off scores**
PALPA 1 Same-different discrimination task using minimal pairs – nonwords	Same: 36/36 Different: 36/36	Same: 24/36 Different: 28/36*	Same: *M* = 25.70 (0.56) Different: *M* = 35.09 (2.34)
PALPA 2 Same-different discrimination task using minimal pairs – real words	Same: 36/36 Different: 36/36	Same: 36/36 Different: 31/36	Same: *M* = 35.54 (0.78) Different: *M* = 34.83 (2.58)
PALPA 4 Minimal pairs test requiring picture selection	40/40	38/40	*M* = 39.00 (1.70)
PALPA 5 Auditory lexical decision task	HIHF: 20/20 HILF: 20/20 LIHF: 20/20 LILF: 20/20 Nonwords: 78/80	HIHF: 20/20 HILF: 20/20 LIHF: 18/20* LILF: 20/20 Nonwords: 76/80	HIHF: *M* = 19.86 (0.48) HILF: *M* = 20 (0) LIHF: *M* = 19.95 (0.22) LILF: *M* = 19.62 (0.67) Nonwords = *M* = 76.00 (4.27)
PALPA 7 Repetition: syllable length	1 syllable: 8/8 2 syllables: 8/8 3 syllables: 8/8	1 syllable: 5/8 2 syllables: 4/8 3 syllables: 1/8	NA

Reading

Table 4.3 summarises DM and MAH's reading performance pre-therapy. Due to the variability in reading performance across the two patients, accuracy was taken as a primary measure for MAH, and RT for DM. DM's responses were timed using a stopwatch, and timing commenced as the to-be-read word was presented and was stopped when DM began to pronounce the word.

Table 4.3. DM's and MAH's scores on tests of reading (denotes impaired performance)*

	DM	*MAH*	*Published normal cut-off scores*
Reading			
PALPA 29 Reading: letter length	3-letter: 6/6, *M* RT = 0.62 4-letter: 6/6, *M* RT = 1.13 5-letter: 6/6, *M* RT = 1.56 6-letter: 6/6, *M* RT = 2.65	3-letter: 5/6* 4-letter: 3/6* 5-letter: 2/6* 6-letter: 0/6*	3-letter: *M* = 6 4-letter: *M* = 6 5-letter: *M* = 6 6-letter: *M* = 6
PALPA 30 Reading: syllable length	1 syllable: 8/8, *M* RT = 1.44 2 syllables: 7/8*, *M* RT = 2.20 3 syllables: 8/8, *M* RT = 2.91	1 syllable: 1/8* 2 syllables: 2/8* 3 syllables: 1/8*	1 syllable: *M* = 7.83 (0.38) 2 syllables: *M* = 8 3 syllables: *M* = 7.90 (0.31)
PALPA 31 Reading: imageability × frequency	HIHF 20/20, *M* RT = 1.76 HILF 20/20, *M* RT = 2.05 LIHF 20/20, *M* RT = 2.58 LILF 19/20, *M* RT = 3.20	HIHF 17/20* HILF 13/20* LIHF 1/20* LILF 2/20*	HIHF: *M* = 19.94 (0.25) HILF: *M* = 19.94 (0.07) LIHF: *M* = 20 LILF: *M* = 19.52 (0.68)
PALPA 32 Reading: grammatical class	Nouns: 20/20 Adjectives: 20/20 Verbs: 20/20 Functors: 17/20*	Nouns: 2/20* Adjectives: 2/20* Verbs: 3/20* Functors: 2/20*	Nouns: *M* = 19.87 (0.43) Adjectives: *M* = 19.97 (0.18) Verbs: *M* = 19.97 (0.18) Functors: *M* = 19.93 (0.37)
PALPA 35 Reading: regularity	Regular words: 30/30 Irregular words: 30/30	Regular words: 15/30* Irregular words: 5/30*	Regular words: *M* = 29.96 (0.20) Irregular words: *M* = 29.85 (0.37)
PALPA 36 Reading: nonwords	3-letter 6/6, *M* RT = 3.14 4-letter 6/6, *M* RT = 3.11 5-letter 5/6, *M* RT = 4.75 6-letter 6/6, *M* RT = 7.64	3-letter 2/6* 4-letter 3/6* 5-letter 1/6* 6-letter 0/6*	3-letter: *M* = 5.77 (0.71) 4-letter: *M* = 5.89 (0.43) 5-letter: *M* = 5.57 (0.90) 6-letter: *M* = 5.65 (0.85)

DM performed well on most tests of reading in terms of accuracy, with the exception of reading function words (Table 4.3, PALPA 32, Kay, 1992). Reading was, however, slow and strongly modulated by word length (PALPA 29, Kay et al., 1992, Experiment 1; see also Osswald et al., 2002), as is consistent with LBL reading. There was no reliable increase in reading times (RT) as syllable length increased with letter-length matched stimuli (PALPA 30, Kay et al., 1992). We examined DM's correct RT data using Krushkal Wallis tests. In DM's real word and nonword letter length data (PALPA 29 and 36) there were significant main effects of length ($X^2(1) = 20.957$, $p < .001$), and an effect of lexicality that approached significance ($X^2(1) = 7.333$, $p = .062$). A similar analysis on the syllable length data found no significant effect of syllable length ($p = .179$). When a Krushkal Wallis test was used on the frequency and imageability RT data, there was a significant main effect of imageability $X^2(1) = 5.109$, $p = .024$, but not frequency ($p = .542$). There are no RT data or analyses for MAH due to poor accuracy.

MAH was impaired on accuracy for most PALPA tests of reading (Table 4.3). MAH's accuracy data on real words and nonwords over length were combined in a loglinear analysis. There was an effect of length ($X^2(3) = 10.974$, $p = .012$), but not lexicality ($X^2(1) = 1.011$, $p = .315$) on reading accuracy, though when summed across word length, there was an effect of lexicality (Chi Square test $X^2(1) = 16.000$, $p < .0001$). Nonword reading was poor even for short word lengths (6/24, Table 4.3). A loglinear analysis on the frequency and imageability accuracy data, revealed a significant two-way interaction between accuracy and imageability ($X^2(1) = 25.642$, $p < .0001$). No other interactions approached significance (all p's > .2).

As with DM, there appeared to be no effect of syllable length on reading when letter length was controlled (PALPA 30, Kay et al., 1992, Table 4.3, Chi Square test, $p = .135$). Though impaired in both conditions, MAH read regularly spelled words better than irregular items (15/30 and 5/30 respectively, Chi Square test $X^2(1) = 20.000$, $p < .0001$). Across the different stimuli the most frequent errors were phonologically related to targets (85.38%, nonwords 55.08% and real words 29.29%), and MAH made no semantic errors.

Spelling

DM performed within the normal range in all spelling assessments (Table 4.4).

Table 4.4. DM's and MAH's scores on tests of spelling (denotes impaired performance).*

	DM	***MAH***	***Published normal cut-off scores***
Spelling			
PALPA 39 Spelling: letter length	3-letter 6/6 4-letter 6/6 5-letter 6/6 6-letter 6/6	3-letter 2/6 4-letter 3/6 5-letter 2/6 6-letter 3/6	NA
PALPA 40 Spelling: frequency × imageability	HIHF 20/20 HILF 20/20 LIHF 20/20 LILF 20/20	HIHF 7/20* HILF 4/20* LIHF 0/20* LILF 0/20*	HIHF: M = 19.94 (0.25) HILF: M = 19.94 (0.07) LIHF: M = 20 LILF: M = 19.52 (0.68)
PALPA 41: Spelling: grammatical class	Nouns 10/10 Adjectives 10/10 Verbs 10/10 Functors 10/10	Nouns 1/10* Adjectives 0/10* Verbs 0/10* Functors 0/10*	Nouns: M = 4.74 (0.42) Adjectives: M = 4.82 (0.48) Verbs: M = 4.82 (0.39) Functors: M = 4.68 (0.55)
PALPA 44: Spelling: regularity	Regular 20/20 Exception 19/20	Regular 6/20 Exception 6/20	NA

MAH was impaired across a range of spelling conditions (Table 4.4). On a spelling test manipulating letter length, MAH scored 10/24. There was no effect of letter length on accuracy (PALPA 39, Kay et al. 1992, Table 4.4) but there were signs of frequency and imageability effects (7/20 for HIHF, 4/20 for HILF, and 0/20 on both LIHF and LILF words, PALPA 40, Kay et al., 1992, Table 4.4). However, there was no effect of regularity (regular 6/20; exception 6/20), a finding reproduced in a homophone spelling test (regular 2/20; exception 3/20). A test assessing performance on words of different grammatical classes produced very low performance (1/20, ‘bell’). MAH scored 7/30 on a lexical morphology spelling test, with no effects of item morphology. Again, successfully spelled words were highly imageable (hairy; bees; rug; mice; daisy; sailor and pram).

When all words spelled at baseline were combined, the majority of errors were phonologically related to targets (33.19%, nonword 17.89% and real word 15.3%). While there were no semantic errors produced in reading, a high proportion of spelling responses were semantically related to the target, contributing to 10.3% of the total errors. Twenty-five percent of the total errors were non-responses.

Tests of semantic processing

DM showed good semantic processing (Table 4.5).

*Table 4.5. DM’s and MAH’s scores on tests of semantic processing * denotes impaired performance.*

	DM	***MAH***	***Published normal cut-off scores***
Tests of semantic processing			
PALPA 51 Semantic association (High Imageablity)	15/15	12/15	*M* = 13.43 (1.26)
PALPA 52 Spoken word - written word matching	15/15	8/15*	*M* = 15
PALPA 53 Picture naming	40/40	18/40*	*M* = 39.80 (0.35)
Pyramids and Palm Trees Test (Howard and Patterson, 1992)			
3 Pictures	51/52	47/52*	49-52 (range)
3 Written words	51/52	51/52	49-52 (range)

MAH was impaired on a number of tests of semantic access (Table 4.5). On picture naming (18/40, PALPA 53 Kay, Lesser and Coltheart, 1992), most (22) errors were phonologically related to targets, resulting in 13 nonword and 6 real word productions. The remaining errors were nonword responses that were phonologically unrelated to the target. On the PALPA test of spoken word-written word matching (PALPA 52), MAH scored 8/15, with 5 errors due to selecting the synonym and 2 to selecting the semantic foil.

Performance on the Pyramids and Palm Trees Test (Howard and Patterson, 1992) was within the normal range when stimuli were three written words, but below the normal cut-off when three pictures were used (47/52).

MAH's performance across input and output modalities

The picture stimuli provided in both the PALPA (Test 53, Kay et al., 1992) and BORB (Test 14, long version, Riddoch and Humphreys, 1993) were combined to investigate MAH's naming performance across different input and output conditions. The same stimuli were used in each presentation condition, and MAH was required to name items that could be presented as pictures, written words or spoken words. He was asked to respond either verbally or in writing, depending on the condition. One condition was completed per session, and a period of two weeks separated each testing session to prevent learning of the items. The presentation order of the items was re-randomised for each new session. The results are presented in Table 4.6.

Table 4.6. MAH's naming performance across input and output modalities. N.B. Types of semantic error were defined thus: Associative = apple-pie; superordinate = pie-dessert; coordinate = pie-cake; synonymic = pie-cake.

	Spoken input, spoken output (repetition)	**Spoken input, written output (spelling to dictation)**	**Written input, written output (direct copy)**	**Written input, spoken output (reading)**	**Picture input, spoken output (verbal picture naming)**	**Picture input, written output (written picture naming)**
Correct (max 116)	35	73	116	37	18	57
Minor orthographic / phonological error	1	5	0	2	0	4
Semantic	9 (associative 6; superordinate 0; coordinate 2; synonymic 1)	13 (associative 6; superordinate 1; coordinate 5; synonymic 1)	0	1 (coordinate)	15 (associative 2; superordinate 1; coordinate 10; synonymic 0)	23 (associative 9; superordinate 2; coordinate 7; synonymic 5)
Mixed (semantically and phonologically related)	3	1	0	1	0	0
Phonologically related real word	29	6	0	21	16	4
Phonologically related nonword	30	5	0	47	27	7
Unrelated nonword	2	6	0	1	14	10
Unrelated pseudohomophone	1	1	0	0	1	4
Unrelated word	3	3	0	3	12	5
Morphological	0	0	0	1	1	2
Perseverative	2	3	0	2	0	2
No response	1	0	0	0	12	1

When written and spoken input and output data were entered into a loglinear analysis, a highly significant interaction between input, output and accuracy was returned ($X^2(1) = 162.432$, $p < .0001$). Separate Chi Square analyses on the accuracy data were used to unpack this interaction. There was a significant advantage for written over spoken output. With spoken and picture input, MAH was better at written than verbal responding ($X^2(1) = 25.016$, $X^2(1) = 29.968$, $p < .0001$ respectively, Table 4.6).

There were effects of stimulus presentation type on MAH's naming accuracy. He performed better with both spoken and written input than picture stimuli: When picture naming was compared with written, $X^2(1) = 54.696$, $p < .0001$, and with spoken presentation, $X^2(1) = 8.603$, $p = .003$, Table 4.6.

Error Types

Repetition produced a preponderance of phonologically related errors (59/81, nonwords 30/81, real words 29/81). In reading, most errors were phonologically related (68/79, nonwords 47/79, real words 21/79). MAH's poorest score was in verbal picture naming, where most errors were phonologically related (43/98, nonwords 27/98, real words 16/98), although there were also frequent unrelated responses (nonword 14/98; real word 12/98) and semantic errors (15/98).

Written output conditions produced a number of semantically related responses, accounting for 13/43 errors in spelling to dictation and 23/59 of errors in written naming of picture stimuli.

Summary: MAH's performance on naming across input and output conditions

Written response modalities produced semantic errors. However, semantic errors never occurred with written input, where phonologically related errors dominated. The absence of semantically related errors in reading suggests that responses were constrained by visual orthographic information about a target word. The number of semantic errors in spelling seen here, taken with the poor spelling performance with nonwords and abstract words demonstrated earlier (Table 4.4), suggests signs of deep dysgraphia.

Case summaries

DM represented a case of 'pure' alexia. She achieved a performance within the normal range on tests of auditory processing, spelling, semantic access and comprehension. Her reading accuracy was usually at ceiling, although her reading times showed effects of letter length. She demonstrated effects of imageability and frequency at pre-therapy baseline.

MAH showed poor comprehension in reading (Table 4.3) and with auditory input (auditory sentence comprehension, spoken word – picture, and spoken word - written word matching). Phonological processing was better preserved, although performance was impaired with meaningless stimuli (Table 4.2). Performance in reading, spelling and comprehension was also influenced by lexical variables (frequency and imageability). A regularity effect was observed in reading only (Tables 4.3 and 4.4), and there was an effect of letter length (linear decline in whole-word accuracy as letter length increases) in reading (Tables 4.3 and 4.4). Broadly, the results suggest that MAH suffers central deficits in language, with a profile of deep dysgraphia for spelling, and reading performance showing some aspects of phonological dyslexia (difficulties reading nonwords, the absence of semantic errors and a preponderance of visual paralexias). In addition, MAH's length-sensitive reading is consistent with a profile of reliance on serial letter processing. It is possible that MAH's ability to fall back on this reading strategy obscured the semantic errors that would characterise his reading as phonological deep dyslexic. The data suggest that MAH was a deep dysphasic, deep dysgraphic and phonological-deep dyslexic patient.

Experimental Investigation: Letter confusability and *N* size

The current rehabilitation aimed to test whether promoting letter identification skills led to increased parallel processing in DM and MAH. Parallel processing was assessed by testing for effects of orthographic neighbourhood (see Arguin, Bub, and Bowers, 1998). The effects of this variable, and of letter confusability, were initially tested as a baseline measure. Effects of orthographic neighbourhood size were tested in DM and MAH pre-therapy, and whether these effects were modulated by letter confusability was assessed.

Method

Materials

The experiment used 180 3-7 letter words that were categorised as low confusability, low neighbourhood size (LCLN, e.g. OTTER), low confusability, high neighbourhood size (LCHN, e.g. RUSTY), high confusability, low neighbourhood size (HCLN, e.g. THEIF), or high confusability, high neighbourhood size (HCHN, e.g. BLOCK). These sets were matched on a range of lexical variables see Table 4.7 for stimulus details).

Table 4.7. M scores on lexical variables for the confusability × N size stimuli used in the Experimental Investigation sections, and as the untreated set in the Rehabilitation Method

	Confusability	*Celex Frequency*	*Regularity*	*Length*	*Imageability*	*Orthographic N Size*	*N*
Low Confusability, Low Neighbourhood	0.026	13.68	20 regular	5.13	381.85	.035	108
Low Confusability, High Neighbourhood	0.027	15.53	18 regular	5.55	388.33	6.528	72
High Confusability, Low Neighbourhood	0.053	15.31	22 regular	5.23	336.89	0.341	82
High Confusability, High Neighbourhood	0.052	15.89	19 regular	5.80	351.04	5.568	81

Measures were based on the summed confusability (structural similarity with other letters) for each word. The confusability scores were taken from the letter confusability matrix provided in Gilmore, Hersh, Caramazza, and Griffin (1979), developed from rating data from normal participants. Words with < .45 summed confusability were categorised as low confusability, and > .53 as high confusability. Words with small orthographic neighbourhoods were determined at those with 0-1 neighbours, and the high *N* word set comprised words with 3 or more neighbours.

Procedure

The experimental stimuli were created using E-Prime stimulation software (E-Prime 1.2, Psychology Software Tools, 2002) and were displayed on a 1024 × 768 Samsung monitor. Words were presented to the left of fixation in point 18 Arial font to counter right hemianopia in both patients. The software recorded accuracy and response times. Vocal onsets were used as a measure of response latency, and these were recorded using a microphone attached to a SR box. After each trial the examiner typed in whether the patient's response was correct (1) or incorrect (0). The examiner manually recorded erroneous responses verbatim. Prior to analysing correct response latencies, the following values were excluded: (1) trials where there was technical difficulty with the response means (e.g. the voice key failed to register the participants' vocal onset), (2) values less than 200ms, and (3) values more than 3 standard deviations away from the mean in that condition for that patient. The words in the experiment served as the untreated word set, and the experiment was run with both patients at each baseline phase.

Results

Accuracy analyses

DM made no errors in any neighbourhood or confusability condition. MAH's accuracy data are provided in Figure 4.2.

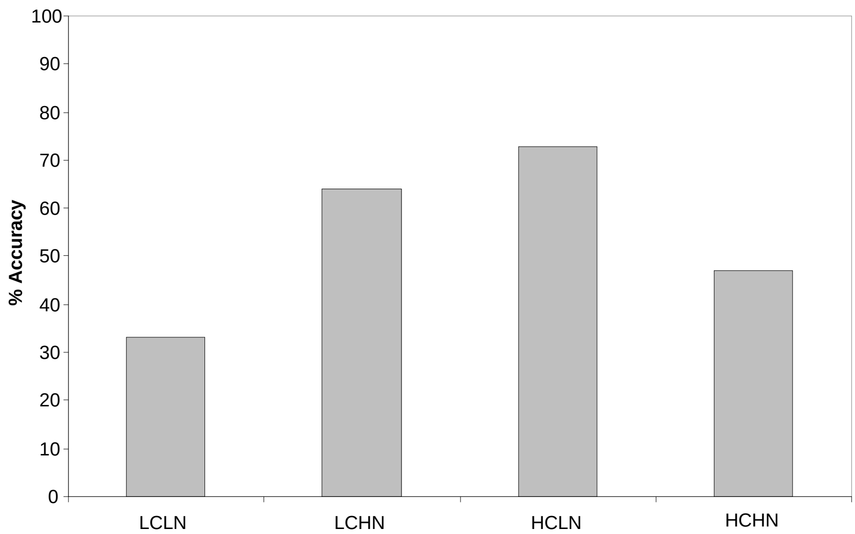


Figure 4.2. MAH's whole word accuracy in reading words (n size and letter confusability). N.B. LCLN = low confusability, low neighbourhood size, LCHN = low confusability, high neighbourhood size, HCLN = high confusability low neighbourhood size, and HCHN = high confusability, high neighbourhood size.

MAH's accuracy data were submitted to a loglinear analysis. The test revealed a significant three-way interaction between confusability, neighbourhood size and score ($X^2(1) = 28.029$, $p < .001$), reflecting a significant facilitative effect of neighbourhood for the low confusability items ($X^2(1) = 16.843$, $p < .001$), but a deleterious effect of neighbourhood for high confusability items ($X^2(1) = 11.562$, $p = .001$, Figure 4.2).

RT analyses

Due to the poor accuracy rates in MAH's performance in some conditions, his RT data were not analysed. DM's mean RT data are presented in Figure 4.3.

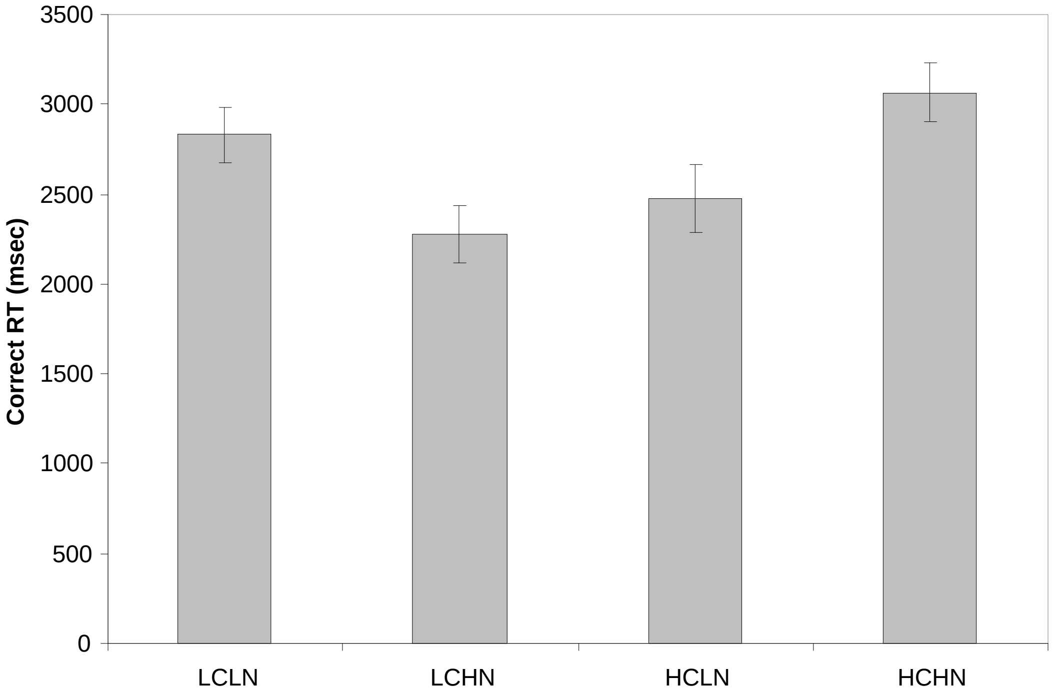


Figure 4.3. DM's response latencies in reading words (n size and letter confusability). Error bars are based on 95% confidence intervals.

Normality tests conducted on DM's RT data were not significant (Shapiro-Wilk test $p > .1$) and so the data were entered into a univariate ANOVA. There was a significant interaction between confusability and N size ($F(1, 338) = 53.433$, $p < .001$). The interaction reflected significantly better performance with high N relative to low N words in the low confusability set ($p = .001$), but a significantly poorer performance with high compared with low N words in the high confusability set ($p < .001$). There was a deficit for high neighbourhood words with high confusable letters ($t(159) = 4.906$, $p < .0001$) and a facilitation with low confusability letters ($t(179) = 5.450$, $p < .0001$). The pattern of facilitative effect of higher N in low but not confusability word is consistent with other work on the subject (e.g. Arguin, Bub and Bowers, 1998), and more generally with the idea that high *N* size is facilitative when letter discrimination is good, but is worsened with higher perceptual demands, where these other neighbours compete for selection.

Discussion

The effect of orthographic neighbourhood appeared to be mediated by letter confusability in MAH and DM. A facilitative effect of high orthographic neighbourhood was found only with words comprised of low confusability letters and for words with high confusability letters, there was a detrimental effect of high neighbourhood. This pattern of results was found in the accuracy data for MAH and the RT data for DM. This latter result could be attributed to disproportionately increased competition from other neighbours for high confusability high *N* sets. On the other hand, the better discrimination of low confusability letters may allow word representations to be activated in parallel, enabling positive effects of supporting (high neighbourhood) words to emerge. In sum, it appears that large neighbourhood is supportive when discrimination is good and increases competition when discrimination is poor.

Rehabilitation Study

The current study used two rehabilitation techniques based on the whole word and letter-level therapies described in Sage, Hesketh and Lambon Ralph (2005). It was hypothesised that the whole word therapy would promote parallel processing, but that this may only be manifested in selective sets with low perceptual demands. For instance, by producing parallel processing, primarily for words comprised of low confusability letters. In contrast, letter recognition therapy, focussed on distinguishing letters from visually similar competitors, aimed to improve letter-by-letter processing, which may in turn improve the reading of words with high confusability letters.

Method

Participants

Patients DM and MAH participated in the rehabilitation study.

Materials

Forty 4- and 5- letter words were selected for treatment. Half of the set contained high confusability (e.g. GABLE) and half low confusability letters (e.g. PROUD, see Table 4.8 for stimuli details).

Table 4.8. M scores across lexical variables for the Treated Set

	Confusability	*Celex Frequency*	*Regularity*	*Length*	*Imageability*	*Orthographic N Size*	*N*
Low Confusability	0.028	30.129	12 regular	M = 4.5 Range = 4-5	331.15	7.4	20
High Confusability	0.051	39.73	10 regular	M = 4.5 Range = 4-5	332.1	8.4	20

The stimuli described in the Experimental Investigations section were reserved as untreated items and were tested at each baseline (untreated confusability by neighbourhood set, see Table 4.7 for stimuli details). Another untreated set of high and low confusability words were also used (see Table 4.9 for details) to assess the effect of confusability over baseline phase overall. As before, words with < .45 summed confusability were categorised as low confusability, and > .53 as high confusability, using Gilmore et al. (1979). Words in each condition were matched across a range of lexical variables (see Tables 4.7, 4.8 and 4.9 for information on the stimuli sets).

Table 4.9. M scores across lexical variables for the untreated confusability experiment

	Confusability	*Celex Frequency*	*Regularity*	*Length*	*Imageability*	*Orthographic N Size*	*N*
Low Confusability	0.026	18.87	42 regular	M = 5.5 Range = 4-7	201.54	2.8	180
High Confusability	0.052	17.74	53 regular	M = 5.5 Range = 4-7	212.83	3.07	180

Procedure

An ABACA design was used. As in Sage et al. (2005), the whole-word treatment was employed first and the letter-level therapy second. An error-reducing paradigm was adopted to prevent reinforcement of errors (Wilson, Baddeley, Evans and Shiel, 1994).

Treatment 1. Word-level therapy

The first therapy aimed to promote parallel processing of words. In both treatments, words were presented on individual cards in upper case in point 16 Arial font. The treatment was conducted at home and also during one-hour weekly rehabilitation and assessment sessions at the University. Patients' partners were given detailed instructions about the methodology and were happy to assist with the treatment. Each treatment phase lasted for 10 weeks.

The procedure for the word-level treatment is outlined below:

(1) A treatment trial began with presenting the participant with a word card.

(2) The therapist repeated the word five times while the patient studied the word on a card. Patients were asked to listen rather than attempt production at this stage.

(3) The patient is then requested to repeat the word five times while still studying the word card. DM never made errors at this stage, but MAH made frequent phonological paraphasia, which may have interfered with the intervention (see Fillingham, Hodgson, Sage and Lambon Ralph, 2003, for a discussion of this in relation to errorless learning). In these instances, the therapist read the word aloud, and asked him to repeat the word a further five times.

Treatment 2. Letter level therapy.

The second treatment aimed to promote letter identification, and overcome letter confusability by emphasising differences between visually similar letters.

(1) The participant was presented with each word card, but this time only one letter was visible at a time, using a 'moving window' card positioned over the word card. While showing the participant each letter, the therapist read each letter name and asked the participant to repeat it. In addition, the participant was asked to trace the letter shape with his / her finger. If any letter was

repeated incorrectly, the therapist read the letter again and asked the participant to repeat it. When a letter name was repeated accurately, the next letter in the word was treated.

(2) When this step was completed successfully, the participant was asked to read each letter before reading the whole word. If an error was made, the participant was taken back to step 1.

Results

Treatment 1. Word-level Treatment

Treated set - accuracy analyses

DM's reading of words was at ceiling at both pre- and post- therapy baselines (20/20 in each confusability and baseline condition), and so her accuracy data were not analysed.

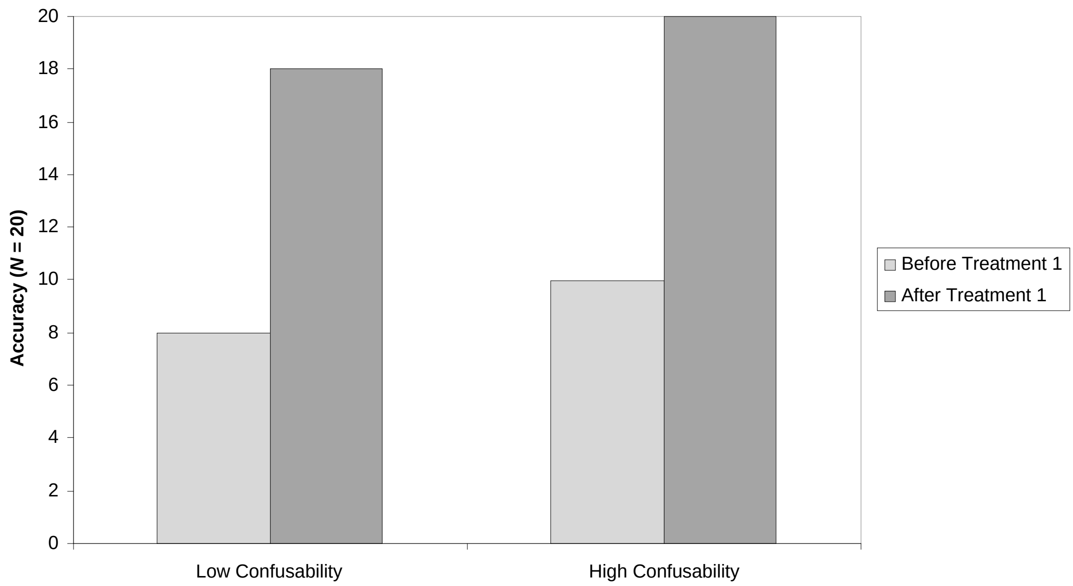


Figure 4.4. MAH's whole-word accuracy in reading high and low- confusability words from the treated set, before and after treatment.

MAH's reading accuracy on the treated words is shown in Figure 4.4. MAH's accuracy data on low and high confusability treated words were entered into a loglinear analysis. There was a significant main effect of baseline (before and after, $X^2(1) = 26.806$, $p < 0001$), but not confusability ($p = 1$). A McNemar test summing across low and high confusability words found a highly significant improvement post-treatment (McNemar test, 18/40 vs. 38/40, $p < .0001$).

Treated set - RT analyses

Reading times for the words in the treated set are provided in Figure 4.5.

Patient DM **Patient MAH**

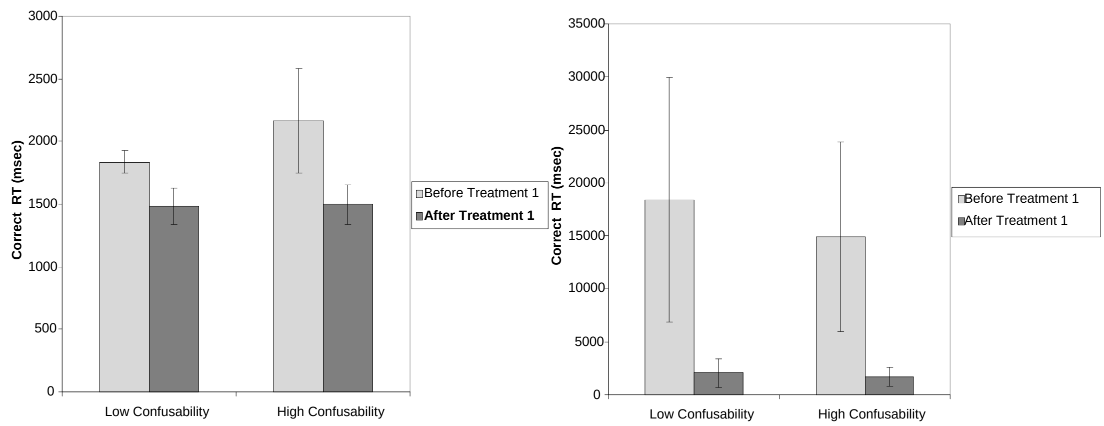


Figure 4.5. Patients' mean correct reading times for high and low- confusability words from the treated set, before and after Treatment 1 (word-level therapy, DM, left panel, MAH, right panel) Error bars are based on 95% confidence intervals.

Normality tests conducted on DM's correct reading time data were not significant (Shapiro-Wilk test $p > .8$), and so the data were analysed using univariate ANOVAs. DM's data revealed a significant effect of baseline ($F(1, 79) = 16.978$, $p < .0001$), but no significant effect of confusability ($p > .5$). MAH's data were not normally distributed (Shapiro-Wilk test p's $< .05$), and so were examined using

Friedman tests. MAH also showed a significant effect of baseline (X^2 (1) = 39.724, p < .0001), but not confusability (p > .4). The absence of an effect of confusability here may be because of the inclusion of both high and low neighbourhood words in this set, where high confusability words with large neighbourhoods elicited a deleterious effect on reading, relative to low *N* high confusability sets, and a facilitatory effect of high neighbourhoods in the low confusability set. Therefore, any effects of confusability may be neutralised, due to the mediation by neighbourhood size.

Treatment 2. Letter-level Treatment

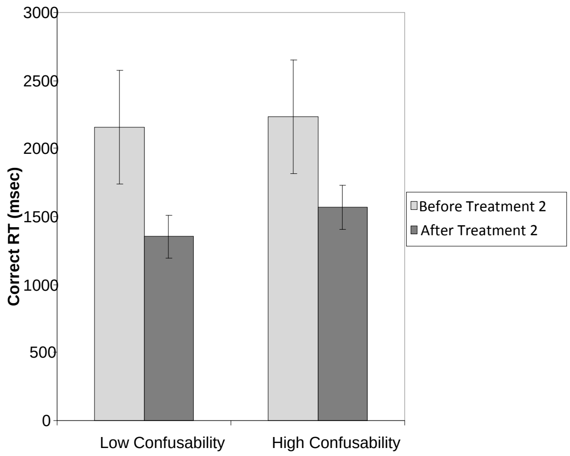
Treated set accuracy analyses

The letter-level treatment used a procedure that required patients to read aloud each letter in the word before they attempted to read the whole form. This error-reducing feature, coupled with the largely improved reading accuracy following the first treatment, led to pre-therapy ceiling accuracy for Treatment 2 that did not change throughout the treatment process. Therefore, accuracy rates on the treated items were not analysed for either patient after Treatment 2.

Treated set RT analyses

Mean RTs for each patient in each condition are shown in Figure 4.6.

Patient DM



Patient MAH

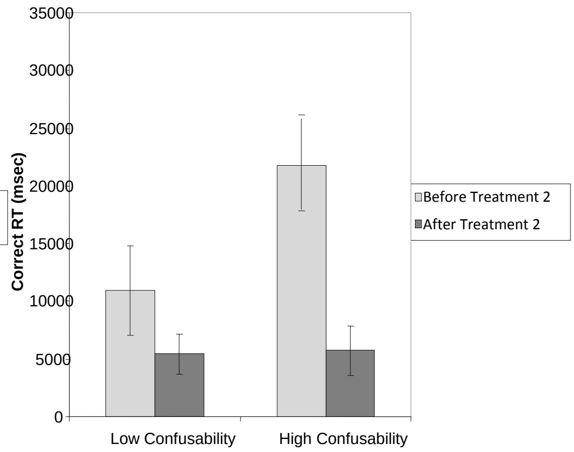


Figure 4.6. Patients' mean correct reading time for high and low- confusability words from the treated set, before and after Treatment 2 (letter-level therapy, DM, left panel, MAH, right panel). Error bars are based on 95% confidence intervals.

Results from normality tests conducted on both patients' correct reading time data were not significant (Shapiro-Wilk test p's > .05). RT data for both patients were entered into univariate ANOVAs. DM showed improved reading times following therapy ($F(1, 75) = 65.332$, $p < .0001$). There was no effect of confusability, and no interaction (p's > .4). MAH's data showed significant effects of treatment baseline ($F(1, 74) = 18.447$, $p < .0001$) and confusability ($F(1, 74) = 4.304$, $p = .042$). The baseline effect reflects facilitated RTs after therapy. The effect of confusability reflects a general improvement across words varying in confusability, after the letter level treatment.

Generalised improvement to untreated items following each treatment

Reading performance of the patients on untreated words was tested at each baseline phase, focusing on a comparison of generalised improvements with low and high confusability words.

Untreated confusability set: Accuracy analyses

DM performed at ceiling under all conditions, making just three errors in total, all of which were due to phonologically similar real-word substitutions. MAH's accuracy rates across low and high confusability words are provided in Figure 4.7.

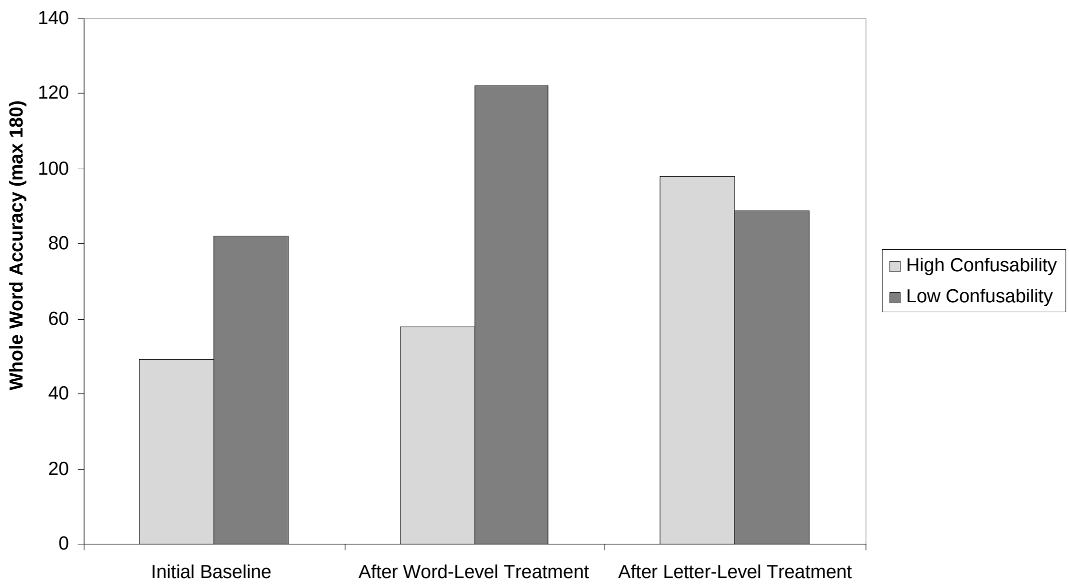


Figure 4.7. MAH's accuracy scores for high and low- confusability words from the untreated set, across baseline phases.

MAH's accuracy data over baseline phase and confusability was analysed using a loglinear analysis. There was a significant interaction between test time, confusability and score ($X^2(2) = 31.141$, $p < .0001$). When analysing performance on the high confusability words only, performance did not improve after the word level treatment (Chi Square test, p = .358); however performance improved significantly after the letter level treatment when compared against the initial baseline ($X^2(1)$ = 24.381, $p < .0001$) and against the phase after the word-level treatment ($X^2(1)$ = 16.386, $p < .0001$). For the low confusability words, performance improved reliably after the word level treatment ($X^2(1) = 15.421$, $p < .0001$), but not after the letter level treatment (p = .833), when compared to initial baseline, and there was a significant *decline* in accuracy on low confusability words between baseline performance after the two therapies ($X^2(1) = 13.837$, $p < .0001$).

Untreated confusability set: RT analyses

MAH's and DM's mean RT data are provided in Figure 4.8.

Patient DM **Patient MAH**

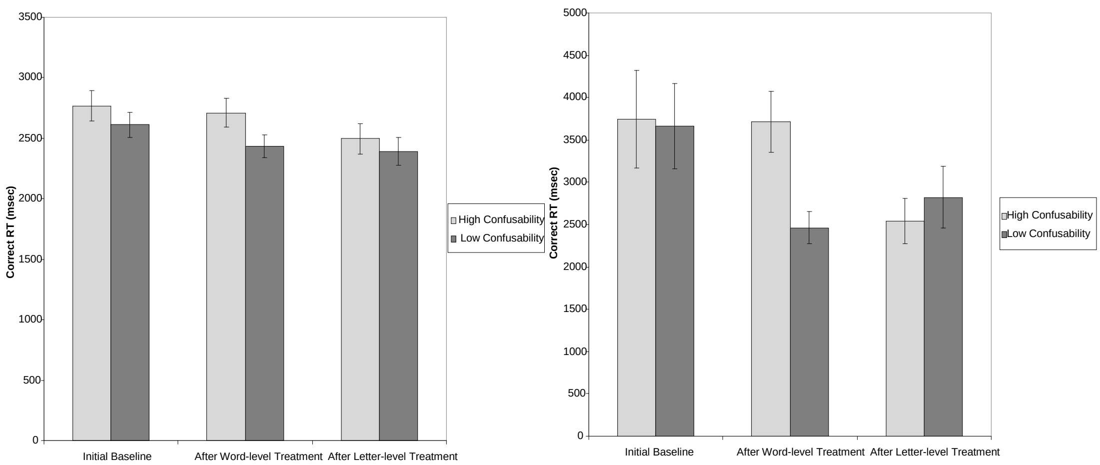


Figure 4.8. Patients' mean correct reading times for high and low- confusability words from the untreated set, across baseline phases (DM, left panel, MAH, right panel). Error bars are based on 95% confidence intervals.

Normality tests conducted on both patients' correct reading time data were not significant (Shapiro-Wilk test *p*'s > .8), and so both data sets were subjected to between subjects ANOVAs. DM's RT data were entered into a between-subjects ANOVA with the factors Confusability (Low and High) and Baseline Phase (Initial Baseline, Word-Level Treatment, Letter-Level Treatment). There was a significant interaction between confusability and baseline phase ($F(2, 194) = 12.565$, $p < .001$). Bonferroni post-hoc tests for the low confusability set showed significant differences between initial baseline and after both treatments (all *p*'s < .0001). There was also a significant difference in performance on the low confusability words between the baseline after the word-level and letter-level treatments ($p < .0001$). For the high-confusability words, there were no differences between initial baseline and after the word-level treatment ($p = .178$), but there was a significant improvement (reduced RTs) for the baseline after the letter-level treatment compared to both initial ($p < .0001$) and after the word-level ($p = .005$) treatment.

A similar analysis of MAH's RT data revealed a significant interaction between confusability and baseline phase ($F(2, 358) = 16.327$, $p < .0001$). Bonferroni post-hoc tests showed a significant improvement on the low confusable word set between initial baseline and after the word-level ($p < .0001$) and letter-level ($p < .0001$) treatments. No significant differences existed between the pre-therapy baseline and that following the letter-level treatment ($p = 1.00$). For the high confusable sets, no differences existed between initial baseline and the baseline that followed the word level treatment ($p = .058$), but there were significant improvements between baselines that followed the word-level and letter-level therapies ($p < .0001$), and between initial baseline and post-letter level therapy baseline ($p < .0001$). At initial baseline, no significant differences existed in reading times for high- and low- confusability words (p = .070), this pattern emerged after word-level treatment (p < .0001). After the letter-level treatment, no differences existed between high and low confusability sets (p = .233).

N effects in the untreated set across baseline tests

Untreated confusability by neighbourhood set
Accuracy analyses

DM scored at ceiling across baseline phases, and so her accuracy data were not considered further. MAH's percent accuracy scores at each baseline phase across neighbourhood and confusability conditions are provided in Figure 4.9.

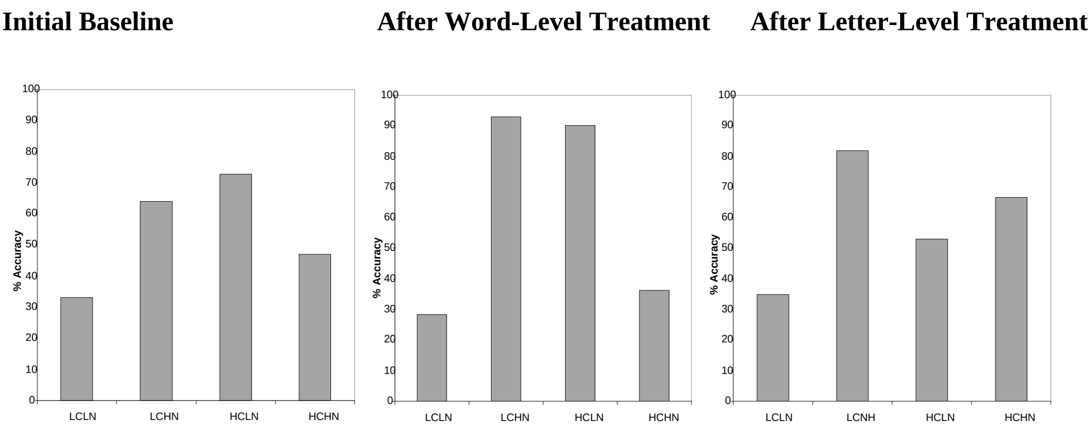


Figure 4.9. MAH's percent accuracy scores across confusability and neighbourhood conditions for each baseline test left to right - (initial baseline, after word-level treatment, after letter-level treatment).

MAH's data across baseline test and neighbourhood and confusability conditions were entered into a loglinear analysis. The analysis returned a significant 4-way interaction between baseline phase, confusability (high or low), neighbourhood size (high or low) and accuracy ($X^2(2) = 41.576$, $p < .001$). To unpack this interaction, separate loglinear analyses were performed on the data from each baseline phase. As detailed in the Experimental Investigations section, the initial baseline data showed a significant three-way interaction between confusability, neighbourhood size and accuracy ($X^2(1) = 28.029$, $p < .001$). This reflected the significant facilitative effect of neighbourhood for the low confusability items ($X^2(1) = 16.843$, $p < .001$), but the deleterious effect of neighbourhood for the high confusability items ($X^2(1) = 11.562$, $p = .001$, Figure 4.9, left panel).

When data from the baseline phase that followed the word-level treatment were subjected to a loglinear analysis, a significant three-way interaction between confusability, neighbourhood size and accuracy was found ($X^2(1) = 137.138$, $p < .001$). As with the initial baseline data, this interaction was due to facilitative effects of high *N* with low confusability words ($X^2(1) = 83.198$, $p < .001$), but detrimental effects of high *N* with high confusability words ($X^2(1) = 55.648$, $p < .001$, Figure 4.9, central panel).

When a loglinear analysis was performed on the data from the baseline phase after the letter-level treatment, there was also a significant 3-way interaction ($X^2(1) = 10.676$, $p = .001$). This was due to the facilitative effect of *N* size with low confusability words ($X^2(1) = 41.026$, $p < .001$), and a facilitative effect of *N* size with high confusability size that did not reach significance ($X^2(1) = 3.192$, $p = .074$, Figure 4.9, right panel).

<u>*RT analyses*</u>

MAH achieved fairly low scores in some of the neighbourhood by confusability conditions (Figure 4.9) and so correct RT analyses were not possible. DM's mean reading latencies over confusability and neighbourhood conditions across baseline phases are provided in Figure 4.10.

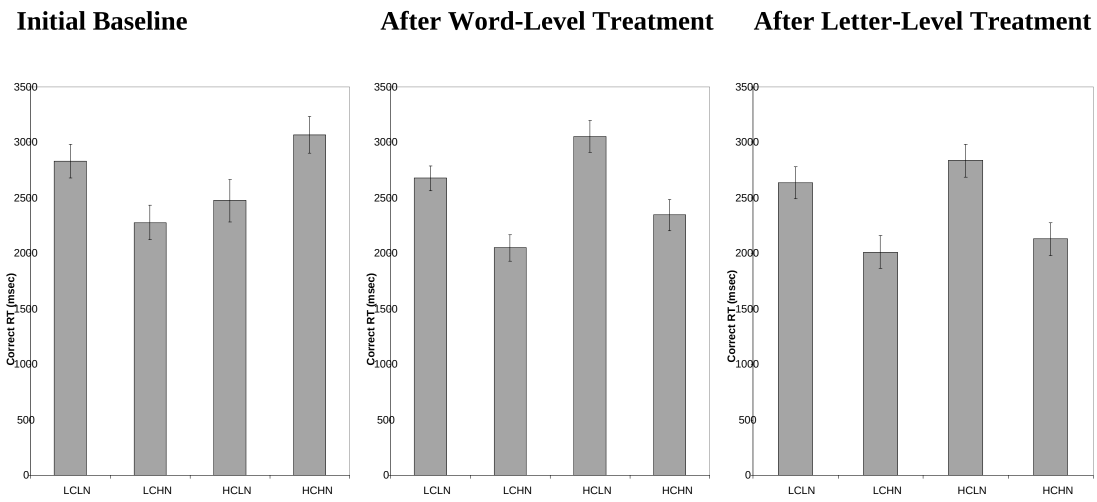


Figure 4.10. DM's mean correct RT across confusability and neighbourhood conditions for each baseline test left to right - (initial baseline, after word-level treatment, after letter-level treatment). Error bars are based on 95% confidence intervals.

Normality tests conducted on DM's correct reading time data were not significant (Shapiro-Wilk test *p*'s > .2) The same words were repeated at each baseline, and so DM's RT data from each condition and at each baseline phase were entered into a repeated measures ANOVA, using the factors confusability, neighbourhood size and baseline phase. The analysis revealed a significant 3-way interaction between time, confusability and (*N* size, $F(2, 126) = 44.762$, $p < .001$.

Bonferonni tests revealed that, at the initial baseline, high *N* words were read faster than low *N* words for the low confusability sets ($p < .001$), but the opposite pattern was seen with high confusability words ($p = .027$). After both word and letter-level treatments, RT's were significantly faster for high *N* words in both low and high confusability sets (all *p*'s < .001, Figure 4.10).

Letter length effect across therapies

In order to test for changes in the length effect over the treatments, performance on 4, 5, 6, and 7- letter words was analysed in both cases over baseline phase. The sets were matched for frequency, confusability, and neighbourhood size (Table 4.10).

Table 4.10. M scores across lexical variables for the untreated letter length experiment

Letter Length	*Confusability*	*Celex Frequency*	*Regularity*	*Imageability*	*Orthographic N Size*	*N*
4	0.038	15.16	40 regular	292.9	4.18	80
5	0.043	12.77	50 regular	257.5	2.06	80
6	0.037	16.58	45 regular	195.3	1	80
7	0.047	12.702	42 regular	164.9	0.41	80

MAH's accuracy data across length and baseline phase are provided in Figure 4.11.

Initial Baseline **After Word-Level Treatment** **After Letter-Level Treatment**

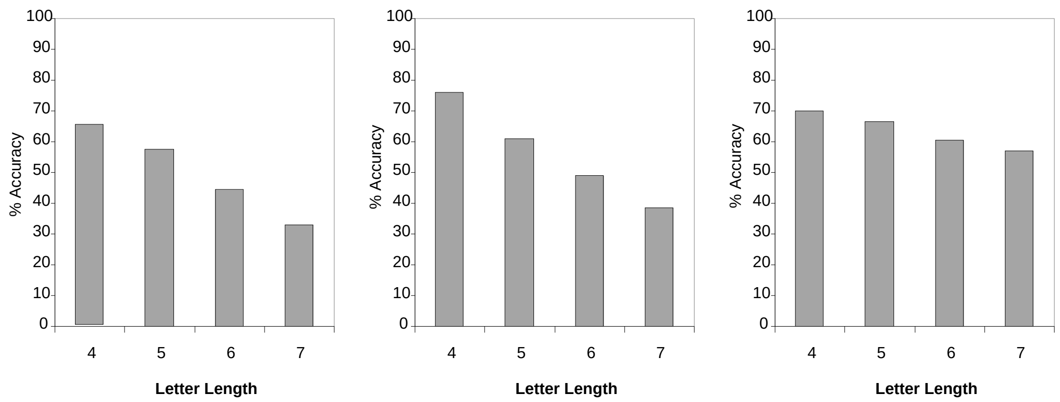


Figure 4.11. MAH's percent accuracy scores across letter lengths for each baseline test left to right - (initial baseline, after word-level treatment, after letter-level treatment).

A loglinear analysis was performed on MAH's accuracy data across length and baseline phase. Though there appeared to be a levelling out of the length effect in accuracy at the post- word-level therapy baseline (Figure 4.11), there was no significant 3-way interaction between accuracy, length and baseline phase (p = .564). Overall, in the three-way analysis, there were significant two-way interactions between accuracy and length (X^2 (6) = 49.463, $p < .0001$), and accuracy and baseline phase (X^2 (2) = 13.572, p = .001). The former interaction reflects the decline in accuracy over length, observed to some extent, at all baseline phases (Figure 4.11). To unpack the accuracy × baseline phase interaction, Chi Square tests were used to compare performance between baseline phases. There was no significant improvement in accuracy overall between the initial baseline and the baseline that followed the word-level (p = .127). However, there were significant improvements between scores at the initial baseline and after the letter-level treatment (X^2 (1) = 12.168, $p < .0001$) treatments, and between the baselines that followed the word-level and letter-level treatments (X^2 (1) = 4.271, p = .039). This indicated that, for MAH, letter-level therapy was more successful, relative to a word-level therapy in improving reading accuracy.

The data from each baseline phase were entered into three separate loglinear analyses, with length and accuracy as the factors. A loglinear analysis on the data from the post-letter therapy baseline, showed no significant interaction between accuracy and length (p = .416), despite significant two-way interactions at initial baseline (X^2 (4) = 24.114, $p < .0001$) and at the post-word therapy baseline (X^2 (4) = 28.169, $p < .0001$). These results suggested that letter-level therapy facilitated remediation of the length effect, and supports the idea that a letter-level therapy may promote parallel processing on all words, irrespective of letter confusability. This result might be attributable to better letter discrimination, which enables parallel processing in these patients. This finding contrasts with the results from the word-level treatment, which appeared to promote parallel processing only when perceptual demands allowed (e.g. low confusability).

DM’s mean correct RT data across letter length and baseline phase are provided in Figure 4.12.

Initial Baseline **After Word-Level Treatment** **After Letter-Level Treatment**

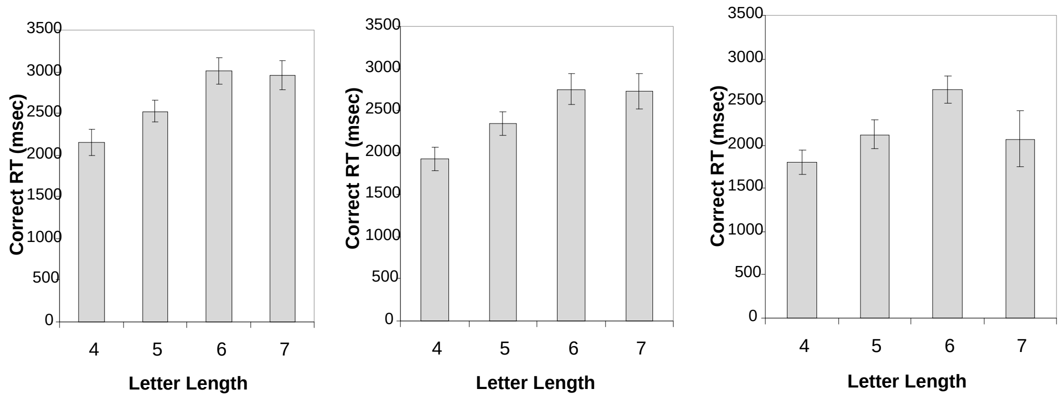


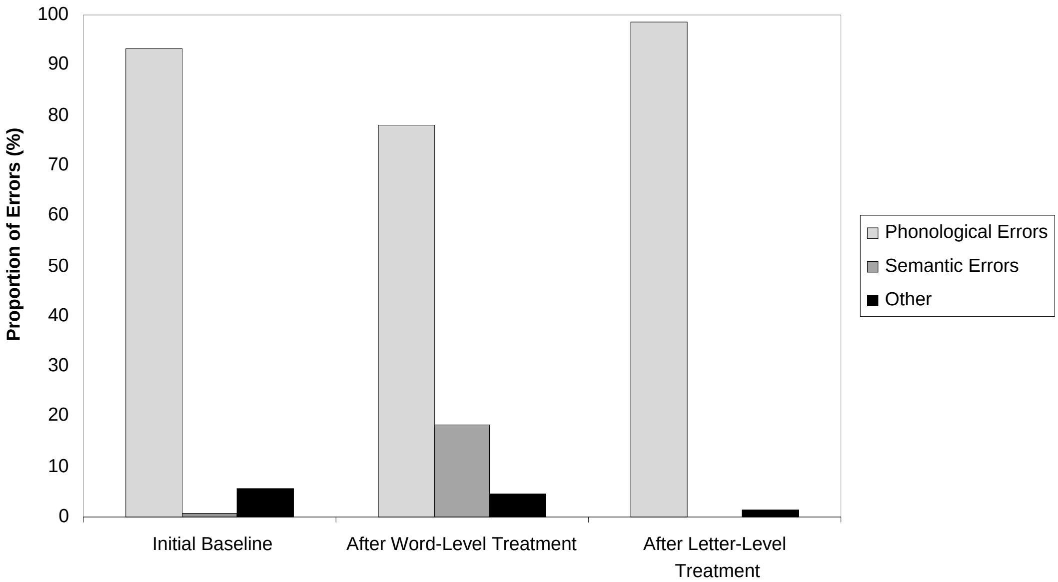
Figure 4.12. DM’s correct RT scores across letter lengths for each baseline test left to right - (initial baseline, after word-level treatment, after letter-level treatment). Error bars are based on 95% confidence intervals.

Normality tests conducted on DM’s correct reading time data were not significant (Shapiro-Wilk test $p > .1$) and so DM’s RT data were analysed across lengths and baseline phase using a between subjects ANOVA, with the factors length (4, 5, 6, and 7) and baseline phase (initial baseline, after word therapy, after letter therapy). The analysis returned a significant interaction between baseline and length ($F(6, 1021) = 2.672$, $p = .014$). There were also highly significant main effects of both factors (Length: $F(3, 1021) = 45.541$, $p < .0001$, and Baseline Phase: $F(2, 1021) = 20.681$, $p < .0001$). The effect of baseline phase on RT reflects the facilitated mean

RT across baseline phases (Figure 4.12). Bonferroni post-hoc analyses performed on these data revealed a significant effect of length at initial baseline ($p < .0001$), post-word therapy baseline ($p < .0001$), but not post- letter therapy baseline ($p = .154$). Similar to MAH, DM's performance improved over baseline phase, and again, the letter-level therapy led to a remediation of the length effect post-therapy.

<u>MAH's Error Types across baseline tests</u>

Figure 4.13 shows the distribution of error types for MAH at each baseline phase.



<u>Figure 4.13. The distribution (% of total errors) of MAH's errors across baseline sessions. Data are from the untreated reading set.</u>

When the data were entered into a loglinear analysis, examining the effects of the factors error type, accuracy, baseline phase on examining the frequency of each error type ($X^2(4) = 28.906$, $p < .0001$).

Phonologically similar responses were the most frequent error type for MAH at each baseline. When the phonological error frequency data were entered into a loglinear analysis, there was a significant three-way interaction between error type, accuracy, and baseline phase ($X^2(4) = 28.906$, $p < .0001$). This interaction reflected a significant reduction in phonological errors between initial baseline and after the word-level treatment (loglinear analysis on the phonological error data with the factors accuracy and baseline phase $X^2(1) = 7.760$, $p = .021$), and no significant difference in the frequency of phonological errors following letter-level treatment, when compared to the baseline that followed the word therapy (loglinear analysis on the phonological error data with the factors accuracy and baseline phase $p = .098$), nor between letter level treatment and the initial baseline (loglinear analysis on the phonological error data with the factors accuracy and baseline phase, $X^2(1) = 1.347$, $p = .510$).

Similar analyses were performed using the total semantic errors. There was again a significant interaction between baseline phase, accuracy, and frequency of semantic errors ($X^2(2) = 18.236$, $p < .0001$). The increase in semantic errors following word-level therapy was significant (loglinear analysis on the semantic error data with the factors accuracy and baseline phase (initial baseline vs. word-level), $X^2(1) = 12.362$, $p = .002$), and there was a significant decline in semantic errors when comparing word level and letter-level baseline phases (loglinear analysis on the semantic error data with the factors accuracy and baseline phase (letter-level vs. word-level) $X^2(2) = 12.735$, $p = .002$), and comparisons of semantic errors at the initial baseline and after word therapy were not significant (loglinear analysis on the semantic error data with the factors accuracy and baseline phase (initial baseline vs. letter-level), $p = .999$).

Discussion

These results demonstrate differential effects of word-level and letter-level therapy on reading in two letter-by-letter dyslexics. The word-level treatment improved reading performance in terms of both reading accuracy and speed for MAH, and for reading speed in DM. This applied to all treated items and to untreated items containing low, but not high confusability letters. The letter-level treatment led to improvements in reading performance on both untreated low and high confusability words. From these findings, it was concluded that the word-level therapy improved parallel processing of letters, generating a general benefit for words with low confusability letters. The general effect of the letter-level treatment was a generalised improvement to words with both low and high confusability letters. These findings indicate that, in contrast to word-level therapy, letter therapy promoted the speed and accuracy of letter-by-letter reading. A key feature of the letter-level treatment was to highlight distinguishing perceptual features of letters, and this may have been crucial in overcoming the letter confusability effect in the patients.

DM

DM, the pure alexic case, showed highly accurate, but slow and length-sensitive reading at initial baseline. Word-level therapy did not facilitate her performance in every reading condition, producing facilitated RTs only on low-confusability words. Letter-level treatment, on the other hand, was extremely effective and improved DM's mean reading time substantially across different word types.

When initially tested, DM showed a pattern of facilitative neighbourhood size (*N*) effects with low confusability words, and a deleterious effect of high *N* with high confusability words. These initial effects were attributed to an ability to process words in parallel for words with low confusability letters (i.e. low perceptual demands), and reliance on a serial strategy with high confusability letters (high perceptual demands), and are consistent with existing literature (e.g. Arguin and Bub 1994, 2005). However, after both treatments there were facilitative effects of *N* size in both confusability conditions (Figure 4.10). This finding is at odds with the lack of generalised improvement to high confusability words after the word-level treatment observed in the overall analysis. As *N* effects are indicative of parallel processing (Arguin and Bub, 2005), these data suggest some improvement in parallel processing even on high confusability words after the word-level therapy. The divergent patterns of *N* effects evidenced between these patients over baseline sessions could be explicable by the differences inherent in the patients' clinical profiles. For instance, DM's reading performance was better relative to MAH's, and so the effect of the first treatment may have been effective enough to produce parallel processing even on the high confusability words, but only in high *N* conditions. Therefore, high confusability items are still problematic, but words with high *N* are recognised faster.

<u>*MAH*</u>

MAH had wider-spread language problems relative to DM, including some apparent semantic deficits and deficits in name retrieval, in addition to his reading impairment. MAH showed some treatment gains in reading accuracy for both treated and untreated items, after both word- and letter-level therapies. As with DM, generalised improvement was modulated by item confusability and treatment type.

The direction of the neighbourhood effect in MAH's accuracy data differed as a function of condition and baseline phase. While low confusability words showed a facilitative effect of *N* size throughout, MAH's reading of high confusability words was detrimentally affected by high *N* size. This pattern was present at initial baseline and remained after word-level treatment. However, letter-level treatment produced an advantage for high *N* words comprised of both low and high-confusability letters (Figure 4.9). So, for DM, high neighbourhood items were facilitated after word treatment and this was maintained after letter level therapy. For MAH, high neighbourhood words remained problematic after the word level treatment, but were facilitated after the letter-level therapy. This finding was not, however because there was no improvement after word treatment for MAH, as word treatment did benefit low neighbourhood words.

In both cases, effects of confusability were still strong after word-level, but not after letter-level treatment (Figures 4.8 and 4.9). MAH's results are consistent with an improved ability to process words in parallel that generalised only to low confusability words after the word-level therapy, but to both low and high confusability sets following the letter-level treatment. For MAH, only the letter level therapy was successful in bringing about *N* effects on both low and high confusability sets, and removing the length effect. DM's data were also consistent with this, as she showed facilitated RTs only on low confusable words after word-level therapy, but on both low and high confusable sets after letter-level therapy, and a more equivalent performance across letter lengths, which could suggest parallel processing.

The findings may have more general implications for theories on LBL reading. Pre-therapy there was evidence for parallel processing in these patients, as *N* effects interacted with confusability. Though this pattern of results followed those demonstrated in three alexic patients reported in Arguin and Bub (2005), but no other studies to the knowledge have directly investigated the *N* effect in alexic patients, though if the *N* effect interacts with confusability, mixed sets of confusability words may be obscuring any effect of *N* in descriptions of these patients.

A number of studies have reported improved parallel processing in LBL readers as a function of whole-word therapy (e.g. Coslett et al., 1993). However, investigations of parallel processing following letter-level treatment are relatively sparse. Arguin and Bub's (1994) letter-level rehabilitation study found improvement on whole word, but not letter-level reading, suggesting that improving a serial strategy may enable parallel processing. Sage et al. (2005) showed that the reading of untreated items only improved after the letter-level treatment. The current data are consistent with these previous findings. Both patients showed improved performance on only low confusable sets after a whole-word therapy, but both high and low confusable words after letter-level therapy. The results are consistent with the idea that letter-level treatment particularly improves reading of high confusability words. Further, the remediation of the length effect in both patients' reading behaviour indicates that letter-level, but not word-level, therapies promoted parallel processing of letters overall.

The current chapter may be informative in terms of the locus of the letter-by-letter reading behaviour in these patients. Most notably, the finding that the largest benefit observed was due to letter level treatment, along with the strong effects of confusability at initial baseline, points to a pre-lexical perceptual deficit.

Distribution of MAH's error types over baseline phases

While DM made very few errors throughout the sessions, MAH's profile was less clear-cut. At initial baseline he produced semantic errors with auditory and picture input but never when reading written words, where phonological errors dominated. At least one letter-by-letter reading patient reported in the neuropsychological literature has shown aspects of a deep dyslexic profile, suggesting that a serial processing strategy may have developed in response to partially recovered deep dyslexia (Buxbaum and Coslett, 1996). It has also been noted that alexic patients may initially produce semantic errors in reading which may decrease over time, as patients adopt a letter-by-letter reading strategy (Landis, Graves and Goodglass, 1982). Strikingly only the word-level treatment led to an increase in semantic errors in MAH. The appearance of semantic errors after word-level treatment to MAH suggests that the therapy reduced the likelihood of MAH using a letter-level reading process, responding instead to information coded across the whole word. Coslett and Saffran (1992) have argued that the semantic errors can reflect right hemisphere reading and letter-by-letter reading a compensatory left-hemisphere strategy. According to this argument, the word-level training encouraged the use of right hemisphere reading at the expense of the left hemisphere compensatory process. It would clearly be interesting to assess this possibility using functional brain imaging.

In contrast to the effects after word-level therapy, semantic errors were no longer produced after letter-level therapy. This is consistent with letter-level treatment leading to more accurate reading responses based on facilitated letter identification and naming, perhaps operating through the left hemisphere. It is interesting to note that irregular words were included in the treated set. This may have been important for MAH, forcing him to rely on letter identification rather than sounds to support his reading (as letter-sound processing would lead to regularisation errors – no regularisation errors were observed). Nevertheless, given MAH's generally poor spoken relative to written word production (Table 4.6), the data indicate a general problem with spoken output in addition to the apparent problems in visual aspects of reading, treated here. These additional output problems likely served to limit improvement – especially for the whole word reading strategy if this led to MAH using a lexical-semantic reading route reliant on impaired phonological retrieval.

Conclusion

The current chapter described two patients with signs of letter-by letter reading who showed effects of letter confusability and orthographic neighbourhood on reading at initial baseline. A whole word treatment improved performance on low-confusability words, but only the letter-level treatment was successful in producing facilitated performance for both high confusability and low confusability words in the patients. These results are consistent with research showing that letter confusability contributes significantly to letter-by-letter behaviour.

The deleterious effect of high *N* words in high confusability sets can be attributed to increased competition following the serial letter processing strategy required to decipher these words, where interference from orthographic neighbours with shared letters to those already processed exert a detrimental effect. In high confusability word sets, letter-level therapy, but not word-level therapy, led to a facilitative effect of *N* on accuracy in MAH, whereas the word-therapy was sufficient to produce a supportive effect of high *N* size on RT in DM, and this effect persisted after letter level treatment.

Only the letter-level treatment was successful in remediating the letter length effect in both patients. This implies parallel processing rather than a highly effective serial strategy. Taken together, the confusability effects at initial baseline, and the highly beneficial effect of the letter-level therapy, lead us to attribute the patients' letter-by-letter reading to damage at a pre-lexical level, i.e. in letter perception / discrimination processes that are needed before a word can be read.

The current study demonstrated that a therapy geared at overcoming confusability through training letter discrimination was more successful than a parallel processing approach, giving better generalisation over different word sets. The letter-based therapy also reduced RTs to the same degree as the whole word approach, and remediated the length effect, both results suggesting that it did not simply lead to more accurate serial reading. Improving perceptual abilities at the letter-level may therefore be the most effective strategy for therapists working with letter-by-letter readers.

Part 2: Exploring associations and dissociations of functions in the cognitive architecture: assessing Neuropsychological patients across therapeutic and experimental contexts.

Chapter 6: The link between STM and Sentence Comprehension: A Neuropsychological Rehabilitation Study

Abstract

A memory rehabilitation study was conducted with two patients with contrasting impairments in verbal short term memory (STM): one with deficient phonological STM (pSTM) and one fitting a profile of impaired semantic STM (sSTM). Two treatments were employed, designed to improve phonological and semantic STM (pSTM and sSTM, respectively). The pSTM treatment selectively improved sensitivity to phonological aspects of STM tasks, and the sSTM treatment similarly brought increased lexical effects in STM performance. To some extent, these findings were replicated in sentence comprehension. The findings are discussed in relation to theories on the components involved in STM, and the role of STM in sentence processing.

Introduction

The link between STM and sentence comprehension

The relationship between sentence comprehension and STM is much debated (Martin, 1990a). Some researchers posit that sentence comprehension uses the same working memory as is used in non-linguistically mediated tasks (domain-general, e.g., capacity constraint theories: Just and Carpenter, 1992), while others argue that there is a separate component in working memory that deals with sentence comprehension (domain-specific, e.g. Caplan and Waters, 1995, 1999).

Neuropsychological studies provide evidence both in support (showing for instance that STM deficits co-occur with poor comprehension of reversible passives; Caramazza, Basili, Koller and Berndt, 1981), and against a link between working memory and sentence comprehension (when patients demonstrate sound comprehension despite severely impaired STM; McCarthy and Warrington, 1987; Butterworth, Campbell and Howard, 1986).

In addition, there are some suggestions that any links between STM and sentence comprehension in neuropsychological data may be modulated by agrammatism. Martin (1987) compared the comprehension of syntax in agrammatic patients with those who were not agrammatic. Comprehension of syntax was much poorer in the agrammatic patients despite having similar deficits in STM. However, one of the non-agrammatic patients with a phonological store impairment showed poor comprehension for structures with difficult components at the beginning of the sentence. To account for these results, Martin suggested that a phonological store might retain words in phonological form as back-up when sentence processing is slow.

Many criticisms levelled at research into any association between memory and sentence comprehension with Broca's aphasic patients concern the difficulty in isolating agrammatic patients' STM problems from difficulties in syntactic processing. The current study contrasts STM deficits in agrammatic and non-agrammatic patients, and tests the assumption that sentence comprehension problems in agrammatic patients are due to syntactic processing difficulties, whereas comprehension problems in non-agrammatic patients are attributable to a phonological short-term memory deficit (Caramazza, Basili, Koller and Berndt, 1981).

It may be that sentence comprehension uses STM under some conditions but not others. Martin (2006) noted that patients impaired in semantic retention performed poorly on sentence anomaly judgements where several adjectives preceded the noun, but performance improved when adjectives followed the noun. Martin explained these results in terms of adjective integration: Where adjectives appear after the noun, adjectives can be integrated with the noun immediately whereas when adjectives precede the noun, each adjective must be retained until the noun is encountered. This argument fits with Dependency Locality Theory (DLT: Gibson, 2000), which emphasises integration complexity (defined as the number of new discourse referents that separate items that need to be linked) as a main factor influencing comprehension.

Martin and colleagues have suggested that there may be distinct semantic and phonological components of STM. Their main evidence for this comes from dissociations between memory for semantic and phonological information. Phonological and semantic STM patients differ from cases of impaired phonological and semantic processing as they show impaired memory for semantic or phonological representations despite spared semantic and phonological processing at the single word level. For instance, a pSTM patient will achieve an impaired score when probed to recall which items in a memory set *rhymed* with a probe word, while sSTM patients will perform poorly with category probes (e.g. which items in the list were animals?), despite being able to make rhyme and category judgements when demands on memory resources are low (Martin and He, 2004). Work by Hoffman and others has suggested that semantic STM patients' deficits may be explained in terms of deficient semantic control processes: Semantic STM tasks place greater demands on semantic control, relative to standardised tests of semantic knowledge, and that patterns of performance of semantic STM patients is consistent with the idea that they occupy the mildest end of the continuum of semantic control impairments, with semantic aphasia representing the more severe end (Hoffman, Jefferies and Lambon Ralph, 2011). The authors assessed three semantic STM patients using tasks designed to manipulate the amount of semantic control required, and found that all three patients showed signs of mild deficits on the tasks, even when stimuli were made visually available throughout the testing session, a modality that lessens STM load. Also, patients were more impaired with increased demands on control, even though the amount of semantic information to be retained was held constant across conditions.

Martin's group have proposed similar phonological and semantic distinctions in sentence comprehension. Martin, Shelton, and Yaffee (1994) reported a double-dissociation in comprehending sentences requiring syntactic-semantic memory and performance requiring phonological retention of sentences. One patient, EA, showed poorer retention on span tasks for phonological rather than for semantic information, while their second patient, AB, performed better on tasks demanding the retention of semantic information. Crucially, the patient with better retention of phonological information performed better at sentence repetition, whereas the patient performing better with semantic information comprehended sentences well. The authors suggest that semantic short-term memory is important for sentence comprehension where several word meanings have to be retained before they can be integrated into propositional representations. On the other hand, phonological short-term memory is crucial for sentence repetition, but plays no major role in answering comprehension questions or detecting sentence anomalies.

Elsewhere, Martin (1990b) differentiates sentences that require syntactic-semantic from those requiring phonological 'reactivation' (i.e. the initial part of the sentence that needs to be recalled for comprehension). Sentences involving a syntactic-semantic reactivation are relative clauses such as 'I met the girl that the grandma drew', and include a constituent (the girl) that is reactivated after the gap, (after 'drew'). In this example the reactivation at the position of the gap is semantic and not phonological, since the reactivation is of the word's meaning rather than its form is needed. In contrast, phonological reactivation is required in structures containing a temporary lexical ambiguity, where an incorrect initial interpretation may result and re-analysis is required.

Rehabilitation studies

Although many patients (particularly those fitting an agrammatic profile) present with both comprehension and STM difficulties, most rehabilitation studies have focussed on improving sentence comprehension directly, through promoting understanding of thematic roles (e.g. Byng, 1988, Byng, Nickels and Black, 1994, Schwartz, Saffran, Fink, Myers, and Martin, 1994). However, some theories suggest that STM deficits affect comprehension of sentences, and so treating STM should result in generalised improvement to sentence comprehension. Intuitively, by improving STM rather than a set of sentences, you might also expect greater generalisation, in that the comprehension of any sentence using STM should improve.

Generalised improvements following memory rehabilitation to other tasks thought to require STM have been found in arithmetic problem solving (e.g. Vallat et al. 2005). However, generalised treatment gains in sentence comprehension after memory training are rarely investigated. To my knowledge, only one study has explored the relation between gains in STM and sentence comprehension. Francis, Clark and Humphreys (2003) trained sentence repetition using a hierarchy of sentence difficulty. Treatment produced significant gains in working memory tasks, backwards but not forwards span, and significant generalised improvement in performance on the Revised Token Test (McNeil and Prescott, 1978). There was a marginally significant improvement on the Test for the Reception of Grammar (TROG, Bishop, 1989), but interestingly no generalised improvement to the Reversible Sentences Comprehension Task (Byng and Black, 1999). However, it is unclear whether the sentence repetition training they used involved some comprehension: It is plausible that the patient might, for instance, have extracted meaning from the sentences in order to retain them. Clearly, if this was the case then generalised gains in sentence comprehension (Revised Token Test, McNeil and Prescott, 1978) are not surprising and do not speak to the issue of an association between STM and sentence comprehension. In fact, an element of comprehension involved in sentence repetition may explain the post-therapy gain found in a LTM verbal memory task (Recognition Memory Test) which is thought to tap semantic rather than phonological short-term memory (Goldblum, Gomez, Dalla Barba, Boller, Deweer, Hahn and Dubois, 1998).

The current study

The aims of the current study were fourfold: i) to test the relation between STM and sentence comprehension using a rehabilitation approach, ii) to explore the extent to which STM contributes to comprehension in an agrammatic patient, iii) to test whether phonological STM training specifically improves sentences / memory span tasks thought to use pSTM (across pSTM and sSTM patients) and iv) to examine whether a rehabilitation of sSTM leads to improvements in sentences and memory tasks that employ sSTM (across pSTM and sSTM patients).

Repeated practice of memory span lists has been found to improve STM performance (Francis, Clark and Humphreys, 2003, Kohn, Smith and Arsenault, 1990). Here recall of memory lists was trained over two ten week periods. The first treatment trained the retention of nonwords, working under the assumption that these items would employ pSTM more than sSTM. The second treatment used real words that involve both phonological and semantic STM, and so may train both semantic and phonological STM. In the latter condition, patients were advised to think about the meaning of each word as they heard them, to promote the use of semantic representations in retention.

Case Descriptions

Patient DS

At the time of testing DS was a 73-year-old male who had suffered a Left Inferior Frontal CVA, resulting in non-fluent Broca's aphasia. Prior to his stroke DS worked in transport and his education history included a non-university diploma. A structural MRI scan was taken, and the lesion was created in SPM and added as an overlay onto a standard multi-slice template in MRIcron. The SPM analysis was a one sample t-test with the covariates healthy (140 brains aged 40+) versus patient, age and sex. The analysis revealed a lesion in the left inferior frontal gyrus, involving the pars opercularis, pars triangularis and pars orbitalis, left rolandic operculum, left insula, left middle frontal gyrus, left precentral gyrus, left postcentral gyrus, left caudate and putamen (Figure 5.1). DS presents with difficulties in STM (span of 2 words).

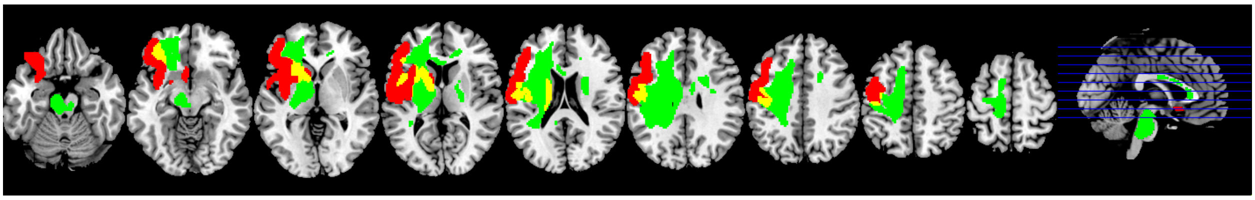


Figure 5.1. Brain images of patient DS. Extent threshold including only significant blobs containing ≥100 voxels. N.B. Grey matter lesion appears in red and white matter lesion in green. The lesion was created in SPM and added as an overlay onto a standard multi-slice template in MRIcron. The SPM analysis was a one sample t- test with the covariates healthy (140 brains aged 40+) vs. patient, age and gender.

Patient AK

Patient AK is a 74-year old male who was referred to the Behavioural Brain Sciences Centre at the University of Birmingham in 2008 following a suspected CVA in 2007. He had previously worked as an accountant and had studied Law, but retired from his work and studies after the CVA. At the time of referral, AK reported difficulties in memory and attention. AK reported that he had experienced chronic expressive aphasia acutely. These difficulties are now largely reconciled, and AK's free speech was fluent and grammatically correct at the time of testing. An fMRI scan suggested some damage to the occipital lobe (Figure 5.2). Despite an otherwise preserved presentation, AK demonstrates some difficulties in STM (span of 3 words).

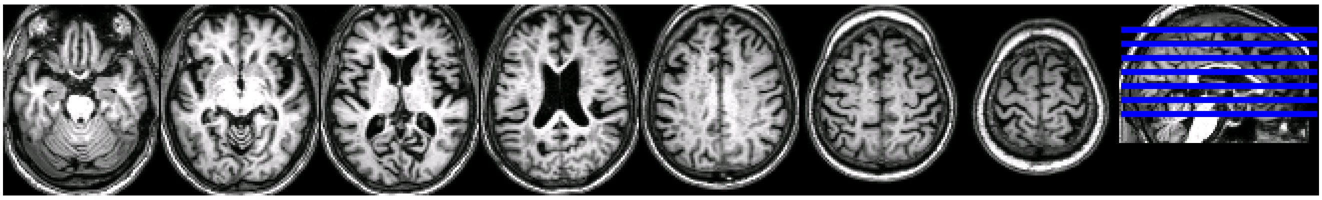


Figure 5.2. Brain images of patient AK. Extent threshold including only significant blobs containing ≥100 voxels. N.B. Grey matter lesion appears in dark grey and white matter lesion in light grey. The lesion was created in SPM and added as an overlay onto a standard multi-slice template in MRIcron. The SPM analysis was a one sample t- test with the covariates healthy (140 brains aged 40+) vs. patient, age and gender.

Assessment of auditory processing

DS scored within the normal range on a same-different discrimination task using nonword minimal pairs (same 30/36, different 35/36, PALPA 1; Kay, Lesser and Coltheart, 1992) and he scored 69/72 on a test of minimal pair discrimination with written selection (same 34/36, different 35/36, PALPA 3; Kay et al., 1992). He also performed well on auditory repetition of (single) words and nonwords (words 1 syllable: 8/8; 2 syllables: 7/8; 3 syllables: 8/8, nonwords 9/10; 9/10, 7/10 PALPA 7 and 8; Kay et al., 1992). Taken together, these results render an explanation of impaired span performance due to a phonemic processing deficit improbable. A ‘central’ phonological processing deficit also seems unlikely For example, DS scored 9/10 when required to ‘exchange the first parts of the words like key chain – chey kain’.

AK made no errors on any of the PALPA assessments of phonological processing conducted here (PALPA 1, 3, 7 and 8).

Assessments of effects in STM

General Methods

Following previous studies on the subject (e.g. Martin, Shelton and Yaffee, 1994, Barde, Schwartz, Thompson-Schill, 2006), performance was compared on span lists from different conditions using Chi Square tests to assess the effects of variables that may contrast phonological and semantic STM (e.g. the effects of item lexicality on memory).

The data reported from tests 1-3 and 5-7 (Table 5.1) used memory span tasks, where the examiner read out lists of items and asked the patients to recall as many items as possible. For each patient, the data are taken from accuracy rates at list lengths one item above word span with verbal presentation and report (unless reported otherwise). The data were summed across position, therefore each correct score denotes one item correctly recalled, irrespective of report order. Consistent with previous work (Jefferies, Hoffman, Jones and Lambon Ralph, 2008), primacy and recency effects in tests 1 and 5 were calculated from the number of items correctly recalled in the medial positions and the initial and final positions respectively. For patient DS lists of 3 items were used (span = 2) and for AK a list length of 5 was used (span = 4). The data in 4 and 8 are computed by contrasting results of a phonological probe task (e.g. name the items in the current list that rhymed with cat) with those from a semantic probe task (e.g. name the items in the current list that were items of furniture). For every patient, each trial randomly presented 6 items from three semantic / phonological categories, and probed one of these pairs: For example, the items chair, lion, chisel, tiger, table, hammer were presented followed by the probe 'name the animals', the correct response being lion and tiger. Where contrasts were made between real word targets, the stimuli were closely matched on any variables that were not specifically manipulated and real and non- word sets were matched on syllable and letter length.

The data for both patients are presented in Table 5.1.

Table 5.1. Patients' performance across tasks designed to assess semantic and phonological STM.[1]

Suggests semantic or phonological STM deficit	***Patterns of performance***	***DS***		***AK***	
		X^2		X^2	
Semantic	1) Reduced / absent primacy effect	List length:6 Primacy: 8/20 (40%) Medial positions: 44/80 (55%), $X^2(1) = 12.698$, $p < .001$	√	List length: 6 Primacy 16/20 (80%), Medial positions: 48/80 (60%), $X^2(1) = 2.778$, $p = .096$.	×
Semantic	2) Absent lexicality effect (words × nonwords)	Real words: 149/180 (82.77%) Nonwords: 96/180 (53.33%), $X^2(1) = 35.891$, $p < .001$	×	Real words: 242/280 (86.4%) Nonwords: 252/280 (90%), $X^2(1) = 1.718$, $p = .190$	√
Semantic	3) Absent frequency effect (high frequency × low frequency)	High frequency: 76/180 (42.22%) Low frequency: 72/180 (40%), $X^2(1) = .184$, $p = .668$	√	High frequency: 188/280 (42.14%) Low frequency: 156/280 (55.71%), $X^2(1) = 7.717$, $p = .005$	×
Semantic	4) Conceptual span is worse than phonological span	Phonological: 22/40 (55%) Conceptual: 10/40 (25%), $X^2(1) = 7.500$, $p = .006$	√	Phonological 8/40 (20%) Conceptual 25/40 (62.5%), $X^2(1) = 14.907$, $p < .001$	×
Phonological	5) Reduced / absent recency effect	List length: 6 Recency: 12/20 (60%) Medial positions: 44/80 (55%), $X^2(1) = 3.509$, $p = .061$.	×	List length: 6 Recency: 12/20 (60%) Medial positions: 48/80 (60%), $X^2(1) = .000$, $p = .1$	√
Phonological	6) Absent phonological similarity effect (phonologically similar × phonologically dissimilar)	Phonologically similar: 94/180 (52.22%) Phonologically dissimilar: 119/180 (66.11%), $X^2(1) = 7.186$, $p = .007$	×	Phonologically similar: 234/280 (83.6%) Phonologically dissimilar: 231/280 (82.5%), $X^2(1) = .114$, $p < .736$	√
Phonological	7) Absent word length effect (short words × long words)	Short words: 117/180 (65%) Long words: 101/180 (56.11%), $X^2(1) = 2.977$, $p = .084$	√	Short words: 251/280 (89.64%) Long words: 223/280 (79.64%), $X^2(1) = 10.770$, $p < .001$	×
Phonological	8) Phonological span is worse than conceptual span	Phonological 22/40 (55%) Conceptual 10/40 (25%)	×	Phonological 8/40 (20%) Conceptual 25/40 (62.5%) $X^2(1) = 14.907$, $p < .001$	√

The number of null effects of phonological and semantic variables in STM characterise AK as a pSTM, and DS as a sSTM patient, respectively. Though both

[1] √ denotes impairment (showing no significant effect of the manipulation), × denotes no impairment (showing a significant effect of the manipulation)

patients showed some effect consistent with the opposite STM type, they both also showed more aspects of one type of deficit over the other. In addition, patterns of performance in phonological and semantic STM patients are not unequivocal: In reporting on a large group of pSTM and sSTM patients, Barde et al. (2010) included cases where one type of STM patient may show a characteristic of the opposite type of deficit, but classified the different patient types on the basis of the magnitude of type-specific STM effects. Here DS was characterised with impaired semantic STM, and AK with a deficient phonological STM. In particular, the pattern of lexicality effects produced across the patients was against the other data, in that patients with a phonological STM deficit did not show lexicality effects, whereas DS (semantic STM) did show a lexicality effect. Another finding that might not be predicted in these patients is the presence of a phonological similarity effect in the reverse direction (an advantage for phonologically similar over dissimilar words) in DS (sSTM). However, this finding might be explicable by DS's difficulties in speech production, in that once key aspects of a phonological form are accessed and produced, subsequent access / production of phonologically related words might be easier. The effects of primacy, frequency, phonological similarity and word length, along with patterns of performance in phonological and semantic span tests were all consistent with the patients' diagnoses.

Experimental investigation: Can semantic / phonological STM deficits predict performance in sentence comprehension?

Sentences from Hanten and Martin (2000) were used to contrast performance in sentence repetition and sentence anomaly judgements, which have previously been found to correlate with phonological and semantic STM (Hanten and Martin, 2000).

Method

Participants

Patients DS and AK participated in the assessment. Data were also collected from 5 age- and education- matched control participants (mean age = 73, SD = 3.38). Control subjects were paid for their participation.

Materials

The set of 128 sentences used in Hanten and Martin (2000) were used. Half of the sentences were anomalous (involving a violation of semantic or pragmatic relationship). Examples of the sentence types are provided in Table 5.2. Hanten and Martin manipulated both the time at which integration could take place and semantic load. Integration refers to the ease at which one constituent may be mapped onto another to enable comprehension, and is frequently defined in terms of the number of items that separate these constituents. Half of the sentences allowed immediate integration and half only allowed a delayed integration. Within each condition, half of the sentences had a semantic load of one and half a semantic load of three. Semantic load was defined as the number of nouns (one or three) that appeared before or after a verb, or the number of adjectives (one or three) that appeared before or after a noun, depending on the condition. As a result, the load 3 sentences were longer than the load 1 structures. There were also forty filler sentences, again, half of these contained an anomaly and half did not. See Hanten and Martin (2000) for further details.

Table 5.2. Examples of sentences from the sentence comprehension tasks. Table taken from Hanten and Martin, 2000. [2]

Immediate integration
Load 1
Sensible They grew flowers behind their house.
Anomalous They grew rocks behind their house.
Load 3
Sensible They grew flowers, trees, and bushes behind their house.
Anomalous They grew flowers, trees, and rocks behind their house.
Delayed integration
Load 1
Sensible Flowers grew behind the house in the yard.
Anomalous Rocks grew behind the house in the yard.
Load 3
Sensible Flowers, trees, and shrubs grew behind the house in the yard.
Anomalous Rocks, trees, and shrubs grew behind the house in the yard.

Procedure

Sentences were read out loud by the examiner. In the repetition task (phonological STM), participants were required to repeat as much of each sentence as possible immediately afterwards. Response accuracy was recorded at the word level, resulting in two measures of accuracy, whole sentence accuracy (correct or incorrect), and proportion of words accurately read in the sentence (expressed as a percentage). In the comprehension condition (semantic STM), participants were asked to judge whether the sentence was acceptable or unacceptable.

Results

Data from the anomaly judgement task (semantic STM) for the patients and control group are presented in Table 5.3.

[2] Many thanks to Gerri Hanten and Randi Martin for supplying test materials

Anomaly judgement assessment (semantic STM)

Table 5.3. Correct responses by condition for the sentence anomaly task (ranges for the controls are shown in parentheses).

	Immediate integration				***Delayed integration***			
	Load 1		*Load 3*		*Load 1*		*Load 3*	
	Anomalous	*Sensible*	*Anomalous*	*Sensible*	*Anomalous*	*Sensible*	*Anomalous*	*Sensible*
Controls	15 (14-16)	16	12 (10-15)	16	11 (10-15)	15 (14-16)	13 (11-16)	16
DS	8/16	11/16	9/16	7/16	9/16	8/16	9/16	8/16
AK	14/16	16/16	10/16	16/16	10/16	16/16	11/16	16/16

DS's scores were below the normal range in all conditions of the sentence anomaly judgement task. AK's performance was within the normal range in every case (Table 5.3). The only reliable interaction in the control data was between sentence type (sensible or anomalous) and accuracy ($X^2(1) = 13.469$, $p < .0001$). A loglinear analysis was performed on the patients' data, with the factors accuracy (correct or incorrect), patient type (pSTM-impaired (AK) or sSTM-impaired (DS)), load (1 or 3), type (sensible or anomalous) and integration (immediate or delayed). The analysis revealed a highly significant three-way interaction between accuracy, patient and sentence type (sensible or anomalous) ($X^2(1) = 24.305$, $p < .0001$), possibly reflecting the better performance of AK (pSTM) than DS (sSTM) on the task, and the trend for a better performance on the sensible relative to anomalous sentences for AK but not DS (Table 5.3), who performed at chance.

To confirm this, we performed separate loglinear analyses on the data for each patient. For DS (sSTM), no significant interactions were found (all p's > .3). When AK's data (pSTM) were submitted to a similar analysis, a significant interaction between accuracy and sentence type was returned ($X^2(1) = 30.198$, $p < .0001$). In sum, there was a deficit for comprehension of anomalous sentences in the pSTM but not sSTM patient, as is consistent with previous findings (e.g. Martin and Romani, 1994, Martin, Shelton and Yaffee, 1994).

Repetition assessment (phonological STM)

The accuracy scores on the repetition version of the sentence comprehension task for each patient and the control group are given in Table 5.4.

Table 5.4. Whole sentence accuracy by condition for the sentence repetition task (ranges for the controls are shown in parentheses).

	Immediate integration				**Delayed integration**			
	Load 1		*Load 3*		*Load 1*		*Load 3*	
	Anomalous	*Sensible*	*Anomalous*	*Sensible*	*Anomalous*	*Sensible*	*Anomalous*	*Sensible*
Controls	14 (13-16)	15 (13-16)	13 (12-16)	15 (14-16)	11 (13-16)	15 (14-16)	12 (10-14)	14 (13-16)
DS	3/16	2/16	2/16	2/16	5/16	4/16	0/16	0/16
AK	15/16	13/16	4/16	6/16	14/16	15/16	2/16	5/16

As in the control data for the anomaly task, there was a significant interaction between accuracy and sentence type (anomalous or sensible) ($X^2(1) = 5.183$, $p = .023$). DS's repetition scores were substantially poorer than those of the controls in every condition. AK's data were within the normal range for all conditions with a load of 1, but fell below the range for all load 3 sentence sets. When the repetition data for both patients were entered into a loglinear analysis, a 4-way interaction between accuracy, integration type, load and patient was found that approached significance ($X^2(1) = 3.480$, $p = .062$). Loglinear tests were then conducted with the factors accuracy, load and integration type for each patient separately. The analysis on DS's data found a significant interaction between integration type, load, and accuracy

($X^2(1) = 7.291$, $p = .007$). Chi Square tests were conducted on DS's data to unpack this interaction. There was an effect of integration type for sentences with a load of 3 ($X^2(1) = 5.812$, $p = .016$), but not for those with a load of 1 ($p = .224$), reflecting a more pronounced deficit for repeating load 3 sentences with delayed integration than for immediate. For AK, there was a highly significant two-way interaction between accuracy and load ($X^2(1) = 56.027$, $p < .001$), reflecting the trend in AK's data for poorer repetition for load 3 relative to load 1 sentences.

Discussion

Both patients DS and AK presented with difficulties in STM, though they can be characterised in different ways. The data for DS fit a profile of semantic rather than phonological STM due to the absence of frequency and primacy effects. Consistent with this, his performance was better in phonological compared with conceptual span conditions. In contrast AK presented as a phonological STM patient. He showed absent effects of phonological similarity and recency, and was also better at conceptual than phonological span.

In the sentence comprehension data, STM impairment type (phonological or semantic) matched type-specific difficulties in sentence comprehension. AK showed difficulties in repetition with load 3 sentences (a finding that appeared not to be a function of any language production deficit, given his fluency of speech, Table 5.2). In contrast, he showed a largely preserved performance in comprehension when compared with the control data: The effect of sentence type (anomalous or sensible) on AK's comprehension performance was replicated in the control data (Table 5.3).

These data are therefore consistent with the theory that phonological STM contributes to sentence repetition, but not comprehension (Hanten and Martin, 2000).

Analysis of DS's performance on the sentence processing tasks is less clear. His score on the sentence anomaly judgement task was impaired relative to controls across conditions, despite his good single word comprehension, as is consistent with a profile of impaired of semantic STM (Hanten and Martin, 2000). However, DS's sentence repetition was also poor, dramatically affected by his speech production problems. DS's repetition performance declined further for load 3 sentences with delayed integration, which may be due to a memory over a production impairment given that sentence length was matched across these delayed and immediate integration sets. Note that a detrimental effect of delayed integration on repetition performance is expected in STM-impaired patients more generally (i.e. is not characteristic of either phonological or semantic-impaired patients, Martin and Romani, 1994, Martin, Shelton and Yaffee, 1994).

Rehabilitation Study

Patients DS and AK were entered into a rehabilitation study to test whether treatments geared at improving phonological and semantic STM respectively brought about any changes in their STM profiles and their sentence repetition and comprehension performance.

Method

Participants

Patients DS and AK participated in the rehabilitation study.

Materials

Lists of 5-letter real words and non- words were created for the treatment sessions and at home exercises. To prevent long-term learning of the stimuli (e.g. reliance on long-term rather than short-term memory), the words used in the lists were changed after each session with other word sets matched for frequency, imageability, age of acquisition and concreteness in the real word case, and for length in the nonword stimuli set.

Procedure

An ABACA design was used, where A denotes baseline and B and C denote treatment. Thorough assessments of memory and sentence comprehension were administered at each baseline stage. The first treatment stage aimed to train phonological STM. The examiner read aloud lists of nonwords at a list length of one item above each patient's span for nonwords (e.g. 4 where a patients' span was 3). The patients were requested to recall as many words from each list as possible in the order they were presented. The examiner recorded the accuracy of each patient's responses after each trial. Patients' errors were always corrected by the examiner providing the correct list at the end of each trial.

In addition to the treatment sessions, patients were given nonword recognition exercises to complete at home. The exercises were provided in a booklet, and each trial involved the presentation of a list of words (one per page), at a length of the patient's pre-therapy nonword recognition span + 1, followed by a response page containing a list of every item encountered on that trial along with an equal number of distractor nonwords (in random order). The patients were instructed to mark on the page which items they had seen in the trial. After responding, the patients were encouraged to flick through the past trial list to check their accuracy.

The same procedure was adopted for real words, using each patient's span for real words +1 in the university sessions with the examiner, and real word recognition span +1 for the exercises completed at home. In addition, the patients were encouraged to think about the meaning of each word as they heard it, to encourage use of semantic rather than phonological STM when encoding the words.

The weekly treatment sessions were conducted in a testing room at the university and lasted approximately 90 minutes. Each week, the patients completed twenty trials of the recognition exercises at home.

Results and Discussion

Performance was analysed at each baseline phase to test if there were changes as a function of each treatment. The results are presented in three sections i) Nonword and real word recall, ii) Phonological and semantic STM profiles across treatments and iii) Generalisation to sentence comprehension. Firstly, data are discussed on the trained task before and after treatments (nonword and real word recall).

Nonword and Real word recall

Nonword recall accuracy before and after nonword list (pSTM) training

Nonword recall was assessed through testing verbal recall of 48 lists of items at lengths that were one item longer than each patient's nonword span as tested at baseline. No items repeated in each testing session, and the order of items was changed for each subsequent session. Incorrect responses were always corrected by the therapist.

Both DS and AK showed improved accuracy in recall at one level above span as a function of nonword (pSTM) therapy (Figure 5.3).

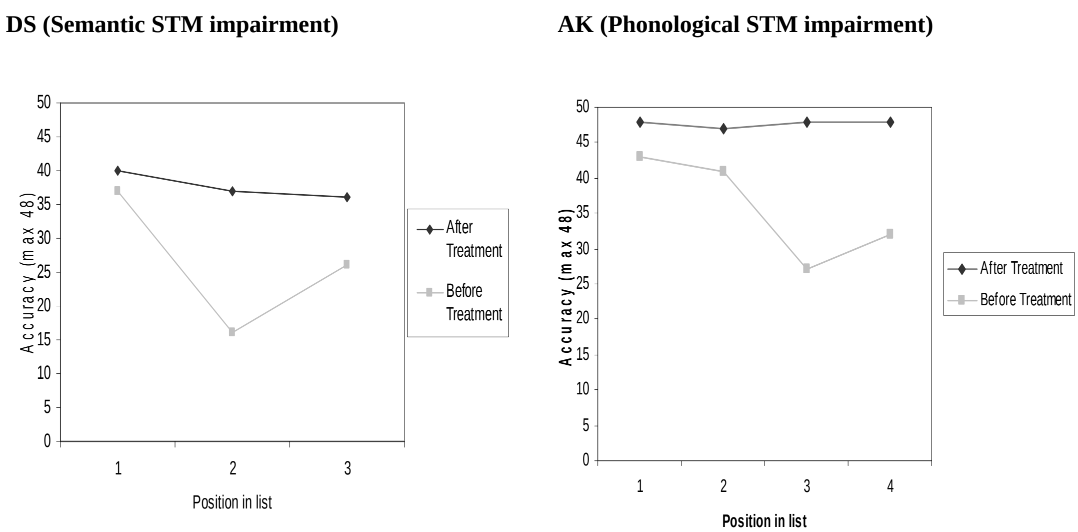


Figure 5.3: Nonword recall before and after nonwords (pSTM) treatment for each patient (DS left panel, AK right panel).

DS showed a significant improvement when data were summed across positions and entered into a McNemar test (153/576 and 212/576, $p < .0001$). Loglinear analyses were run with the factors accuracy, baseline phase (before or after) and position in list. There were significant interactions between accuracy and time ($X^2(1) = 26.299$, $p < .0001$), and accuracy and position ($X^2(2) = 31.071$, $p < .0001$). There was no significant three-way interaction between accuracy, position and baseline phase ($p = .779$). AK's data were summed across positions, and compared over baseline phases using McNemar tests. There was a significant improvement in the data overall following the nonword treatment (329/768 and 379/768, McNemar $p < .0001$). Results from loglinear analysis with the factors accuracy, position and baseline phase found significant two-way interactions between accuracy and time ($X^2(1) = 52.749$, $p < .0001$) and accuracy and position ($X^2(1) = 8.166$, $p = .043$). The three-way interaction did not approach significance ($p = .514$). Both patients showed improved nonword recall after nonword (phonological STM) treatment, but there were no reliable changes in the pattern of accuracy across position in the list for either patient.

Real word recall accuracy before and after real word list (sSTM) training

Real word recall was assessed through testing verbal recall of 48 lists of items at lengths that were one item longer than each patient's word span as tested at baseline. No items repeated in each testing session, and the order of items was changed for each subsequent session. Incorrect responses were always corrected by the therapist.

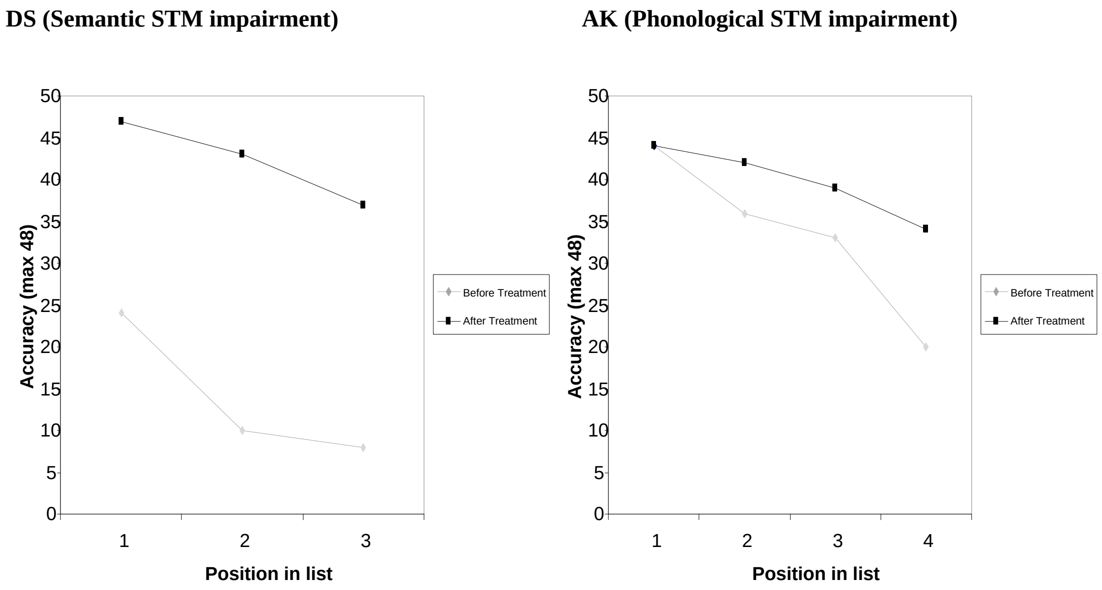


Figure 5.4. Real word recall before and after real word (sSTM) treatment for each patient (DS left panel, AK right panel).

When summing across positions, DS showed a significant improvement following the real word (semantic STM) therapy (120/576 and 250/576, McNemar $p < .0001$). A loglinear analysis with the factors accuracy, baseline phase (before or after), and position (1, 2 or 3) on DS's data revealed significant two-way interactions between accuracy and time ($X^2(1) = 145.850$, $p < .0001$) and accuracy and position ($X^2(2) = 48.001$, $p < .0001$), but no three way interaction ($p = .361$, Figure 5.4, left panel).

AK also showed a significant improvement in performance with real word lists overall (267/768 and 309/768, McNemar $p < .0001$). A loglinear analysis showed that accuracy interacted significantly with time ($X^2(1) = 11.137$, $p = .001$) and position ($X^2(3) = 78.662$, $p < .0001$), but there was no 3-way interaction ($p = .664$, Figure 5.4, right panel).

After the real word (semantic STM) treatment, both patients showed improved accuracy compared to initial baseline, but, as with the nonword treatment, there was no change in the pattern of accuracy over serial position.

Phonological and semantic STM profiles across treatments

The patients' scores on memory tasks were assessed at a list length of span level + 1 across conditions, at initial baseline, and at baselines that followed the phonological STM treatment, and the semantic STM treatment. Methodological details are provided in Appendix E. The data for DS and AK are presented in Tables 5.5 and 5.6, respectively.

Patient DS

Table 5.5. DS's (sSTM) performance on memory and sentence comprehension tasks at initial baseline and at baselines after each treatment phase.[3]

Suggests semantic or phonological STM deficit	Patterns of performance	Baseline		After Treatment 1 Phonological STM		After Treatment 2 Semantic STM	
		X^2		X^2		X^2	
Semantic	1) Reduced / absent primacy effect	List length:6 Primacy: 8/20 (40%) Medial positions: 44/80 (55%), $X^2(1) = 12.698$, $p < .001$	√	List length: 6 Primacy: 10/20 (50%) Medial positions: 32/80 (40%), $X^2(1) = .657$, $p = .418$	√	List length: 6 Primacy: 20/20 (100%) Medial positions: 28/80 (35%), $X^2(1) = 27.083$, $p < .001$	×
Semantic	2) Absent lexicality effect (words × nonwords)	Real words: 149/180 (82.77%) Nonwords: 96/180 (53.33%), $X^2(1) = 35.891$, $p < .001$	×	Real words: 72/90 (80%) Nonwords: 78/90 (86.66%), $X^2(1) = 2.707$, $p < .100$	√	Real words: 80/90 (88.88%) Nonwords: 56/90 (62.22%), $X^2(1) = 17.326$, $p < .001$	×
Semantic	3) Absent frequency effect (high frequency × low frequency)	High frequency: 76/180 (42.22%) Low frequency: 72/180 (40%), $X^2(1) = .184$, $p = .668$	√	High frequency: 66/90 (73.33%) Low frequency: 64/90, (71.11%) $X^2(1) = .111$, $p = .739$	√	High frequency: 66/90 (73.33%) Low frequency: 72/90 (80%), $X^2(1) = 1.118$, $p = .290$	√
Semantic	4) Conceptual span is worse than phonological span	Phonological: 22/40 (55%) Conceptual: 10/40 (25%), $X^2(1) = 7.500$, $p = .006$	√	Phonological: 32/40 (80%) Conceptual: 18/40 (45%), $X^2(1) = 10.453$, $p < .001$	√	Phonological: 18/40 (45%) Conceptual: 20/40 (50%), $X^2(1) = .287$, $p = .592$	×
Phonological	5) Reduced / absent recency effect	List length: 5 Recency: 12/20 (60%) Medial positions: 44/80 (55%), $X^2(1) = 3.509$, $p = .061$.	×	List length: 6 Recency: 16/20 (80%) Medial positions: 32/80 (40%), $X^2(1) = 10.256$, $p < .001$	×	List length: 6 Recency: 15/20 (75%) Medial positions: 28/80 (35%), $X^2(1) = 18.976$, $p < .001$.	×
Phonological	6) Absent phonological similarity effect	Phonologically similar: 94/180 (52.22%) Phonologically dissimilar: 119/180 (66.11%), $X^2(1) = 7.186$, $p = .007$	×	Phonologically similar: 70/90 (77.77%) Phonologically dissimilar: 80/90 (88.88%), $X^2(1) = 4.000$, $p = .046$	×	Phonologically similar: 72/90 (80%) Phonologically dissimilar: 84/90 (93.33%), $X^2(1) = 6.923$, $p = .009$	×
Phonological	7) Absent word length effect (short words × long words)	Short words: 117/180 (65%) Long words: 101/180 (56.11%), $X^2(1) = 2.977$, $p = .084$	√	Short words: 70/90 (77.77%) Long words: 62/90 (68.88%), $X^2(1) = 1.818$, $p = .178$	√	Short words: 62/90 (68.88%) Long words: 60/90 (66.66%), $X^2(1) = .102$, $p = .750$	√
Phonological	8) Phonological span is worse than conceptual span	Phonological 22/40 (55%) Conceptual 10/40 (25%)	×	Phonological: 32/40 (80%) Conceptual: 18/40 (45%), $X^2(1) = 10.453$, $p < .001$	×	Phonological: 18/40 (45%) Conceptual: 20/40 (50%), $X^2(1) = .287$, $p = .592$	×

[3] √ denotes impairment (showing no significant effect of the manipulation), × denotes no impairment (showing a significant effect of the manipulation)

DS's data, including Chi Square test statistics for each contrast, are provided in Table 5.5.

Performance was summed across task contrasts within each assessment type (semantic and phonological) and baseline phase and the data entered into a loglinear analysis. The factors were accuracy, baseline phase, and assessment type (phonological or semantic contrasts). The analysis revealed a significant interaction between accuracy and baseline phase ($X^2(1) = 66.709$, $p < .0001$), but no interaction between accuracy and test type ($p = .213$). Separate Chi Square tests then compared performance at each baseline phase, summing across phonological and semantic assessments. There was a highly significant improvement overall between performance at the initial baseline and after the phonological STM treatment ($X^2(1) = 57.829$, $p < .0001$), and after the semantic STM treatment ($X^2(1) = 29.195$, $p < .0001$). Interestingly, there was a significant decline in performance between post-phonological STM and post-semantic STM treatment baselines (752/1080 vs. 708/1080 $X^2(1) = 4.092$, $p = .043$). The improvement in STM performance after phonological treatment can be attributed to better use of phonological representations when encoding stimuli.

At the initial baseline, DS's performance fitted a profile of impaired semantic STM. Following the phonological STM treatment, this pattern remained largely unchanged. Broadly speaking, the only change was that the lexicality effect was levelled out due to an improved performance in nonword recall following treatment (using Chi Square tests, a significant improvement in recall of nonwords was found: $X^2(1) = 29.095$, $p < .0001$, while there was no change in real word recall $X^2(1) = 1.304$, $p = .253$).

In contrast, the semantic STM treatment produced several changes in DS's STM profile. Only one pattern characteristic of semantic STM patients remained, an absent lexical frequency effect ($X^2(1) = 1.118$, $p = .290$). A highly significant effect of primacy was observed after the semantic STM treatment ($X^2(1) = 27.083$, $p < .001$), along with equivalent performance on phonological and conceptual span tasks ($X^2(1) = .287$, $p = .592$). The lexicality effect that was present at initial baseline, but disappeared after phonological STM treatment, was also produced following semantic STM treatment ($X^2(1) = 17.326$, $p < .001$). Patterns of phonological effects remained much the same at post-phonological and post-semantic treatment baselines (Table 5.5).

There was an overall decline in performance between the post-phonological and post-semantic baseline phases despite the appearance of more normal semantic effects. Taken together, these two findings could indicate a promoted use of semantic information when encoding word stimuli (thus semantic effects), but that this is less efficient, overall, than a phonological strategy (752/1080 vs. 708/1080, $X^2(1) = 4.092$, $p = .043$).

Patient AK

Table 5.6. AK's (pSTM) performance on memory and sentence comprehension tasks at initial baseline and at baselines after each treatment phase.4

	Patterns of performance	Baseline		After Treatment 1 Phonological STM		After Treatment 2 Semantic STM	
		X^2		X^2		X^2	
Semantic	1) Reduced / absent primacy effect	List length: 6 Primacy 16/20 (80%), Medial positions: 48/80 (60%), $X^2(1) = 2.778$, $p = .096$.	√	List length: 6 Primacy: 20/20 (100%) Medial positions: 44/80 (55%), $X^2(1) = 14.062$, $p < .001$	×	List length: 6 Primacy: 20/20 (100%) Medial positions: 52/80 (65%), $X^2(1) = 9.722$, $p = .002$	×
Semantic	2) Absent lexicality effect (words × nonwords)	Real words: 242/280 (86.4%) Nonwords: 252/280 (90%), $X^2(1) = 1.718$, $p = .190$	√	Real words: 104/120 (86.66%) Nonwords: 106/120 (88.33%), $X^2(1) = .152$, $p = .696$	√	Real words: 106/120 (88.33%) Nonwords: 110/120 (91.66%), $X^2(1) = .741$, $p = .389$	√
Semantic	3) Absent frequency effect (high frequency × low frequency)	High frequency: 188/280 (42.14%) Low frequency: 156/280 (55.71%), $X^2(1) = 7.717$, $p = .005$	×	High frequency: 108/120 (90%) Low frequency: 102/180 (85%), $X^2(1) = 1.371$, $p = .242$	√	High frequency: 102/120 (85%) Low frequency: 106/120 (88.33%), $X^2(1) = .577$, $p = .448$	√
Semantic	4) Conceptual span is worse than phonological span	Phonological 8/40 (20%) Conceptual 25/40 (62.5%), $X^2(1) = 14.907$, $p < .001$	×	Phonological: 19/40 (47.5%) Conceptual: 24/40 (60%), $X^2(1) = 1.257$, $p = .262$	×	Phonological: 19/40 (47.5%) Conceptual: 20/40 (50%), $X^2(1) = .050$, $p = .823$	×
Phonological	5) Reduced / absent recency effect	List length: 5 Recency: 12/20 (60%) Medial positions: 48/80 (60%), $X^2(1) = .000$, $p = .1$	√	List length: 6 Recency: 12/20 (60%) Medial positions: 44/80 (55%), $X^2(1) = .162$, $p = .687$	√	List length: 6 Recency: 8/20 (40%) Medial positions: 52/80 (65%), $X^2(1) = 4.167$, $p = .041$	√
Phonological	6) Absent phonological similarity effect	Phonologically similar: 234/280 (83.6%) Phonologically dissimilar: 231/280 (82.5%), $X^2(1) = .114$, $p < .736$	√	Phonologically similar: 86/120 (71.66%) Phonologically dissimilar: 104/120 (86.66%), $X^2(1) = 8.185$, $p = .004$	×	Phonologically similar: 90/120 (75%) Phonologically dissimilar: 105/120 (87.5%) $X^2(1) = 8.185$, $p = .013$	×
Phonological	7) Absent word length effect (short words × long words)	Short words: 251/280 (89.64%) Long words: 223/280 (79.64%), $X^2(1) = 10.770$, $p < .001$	×	Short words: 112/120 (85%) Long words: 102/120 (93.33%), $X^2(1) = 6.154$, $p = .013$	×	Short words: 112/120 (93.33%) Long words: 100/120 (90%), $X^2(1) = 6.154$, $p = .016$	×
Phonological	8) Phonological span is worse than conceptual span	Phonological 8/40 (20%) Conceptual 25/40 (62.5%) $X^2(1) = 14.907$, $p < .001$	√	Phonological: 19/40 (47.5%) Conceptual: 24/40 (60%), $X^2(1) = 1.257$, $p = .262$	×	Phonological: 19/40 Conceptual: 20/40, $X^2(1) = .050$, $p = .823$	×

4 √ denotes impairment (showing no significant effect of the manipulation), × denotes no impairment (showing a significant effect of the manipulation).

Similar analyses were performed on AK's data (Table 5.6). Firstly, task contrasts were summed across within each assessment type (semantic and phonological) and baseline phase and performed a loglinear analysis on the data. The factors were accuracy, baseline phase, and assessment type (phonological or semantic contrasts). The analysis revealed a significant three-way interaction between accuracy, test type and baseline phase ($X^2(1) = 20.646$, $p < .0001$).

Separate Chi Square tests then compared performance at each baseline phase, and each assessment type. For the semantic assessment data, there were highly significant improvements between initial and post- phonological STM baselines ($X^2(1) = 14.508$, $p < .0001$), and post-semantic baselines ($X^2(1) = 17.959$, $p < .0001$). There was no difference in overall performance between post-phonological STM and post-semantic STM treatment baselines (527/660 vs. 535/663, $X^2(1) = .149$, $p = .699$).

The patterns of STM effects for AK at each baseline are given in Table 5.6. At the initial baseline phase, AK could broadly be characterised in terms of a phonological STM deficit, showing absent effects of recency ($X^2(1) = .000$, $p = .1$) and phonological similarity ($X^2(1) = .114$, $p < .736$), as well as a better performance on conceptual relative to conceptual span ($X^2(1) = 14.907$, $p < .001$). After phonological STM treatment, this pattern changed and AK produced many of the phonological effects in STM tasks that you would expect from a control participant. Notably, effects of phonological similarity ($X^2(1) = 8.185$, $p = .004$) and word length ($X^2(1) = 6.154$, $p = .013$) emerged, and phonological span was no longer poorer than conceptual span ($X^2(1) = 1.257$, $p = .262$), though the absent recency effect appeared more resistant to change ($X^2(1) = .162$, $p = .687$).

After the semantic STM treatment, AK's patterns of performance remained roughly the same.

Generalisation to sentence comprehension

Anomaly task (test of semantic STM)

The anomaly task involved auditory presentation of sentences used in Hanten and Martin (2000). The examiner read out each of the sentences, half of which were anomalous, and half were sensible. The patients were asked to judge whether the sentence was acceptable or unacceptable.

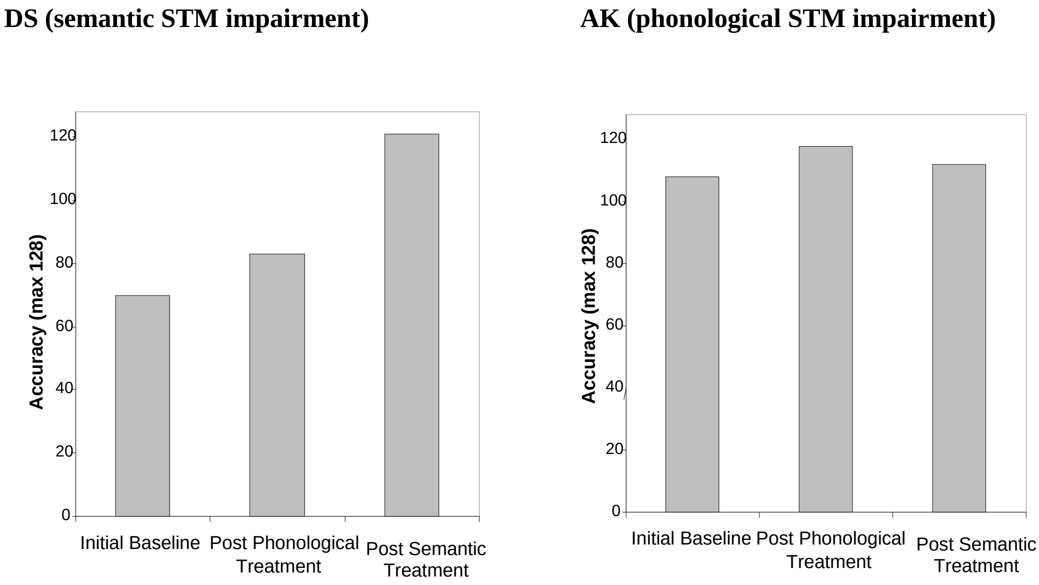


Figure 5.5. Performance on the anomaly task at each baseline phase for both patients (DS left panel, AK right panel).

A 3-way loglinear analysis was performed on the anomaly data with the factors accuracy, baseline phase, and patient. The analysis produced a 3-way interaction between accuracy, baseline phase and patient ($X^2(2) = 26.207$, $p < .0001$). Accuracy did not interact with baseline phase for AK ($p = .143$, Figure 5.5, right panel), but did for DS ($X^2(2) = 63.386$, $p < .0001$, Figure 5.5, left panel). Chi Square tests were used to unpack this interaction for DS. When contrasting scores at initial baseline and after the phonological treatment, there was no significant improvement ($X^2(1) = 2.745$, $p = .098$, McNemar test, 70/128 and 83/128, $p = .09$). However, there were significant differences when comparing scores at initial baseline and after the word therapy ($X^2(1) = 53.633$, $p < .0001$, McNemar test, 70/128 and 121/128, $p < .0001$), as well as between baseline phases after nonword and real word treatment ($X^2(1) = 34.848$, $p < .000$, McNemar test, 70/128 and 121/128, $p < .0001$, Figure 5.5, left panel).

Repetition task (test of phonological STM)

The sentence repetition task used the same sentences reported in the anomaly task (Hanten and Martin, 2000). The examiner read out each sentence and required the patients to repeat the sentence once the examiner had finished.

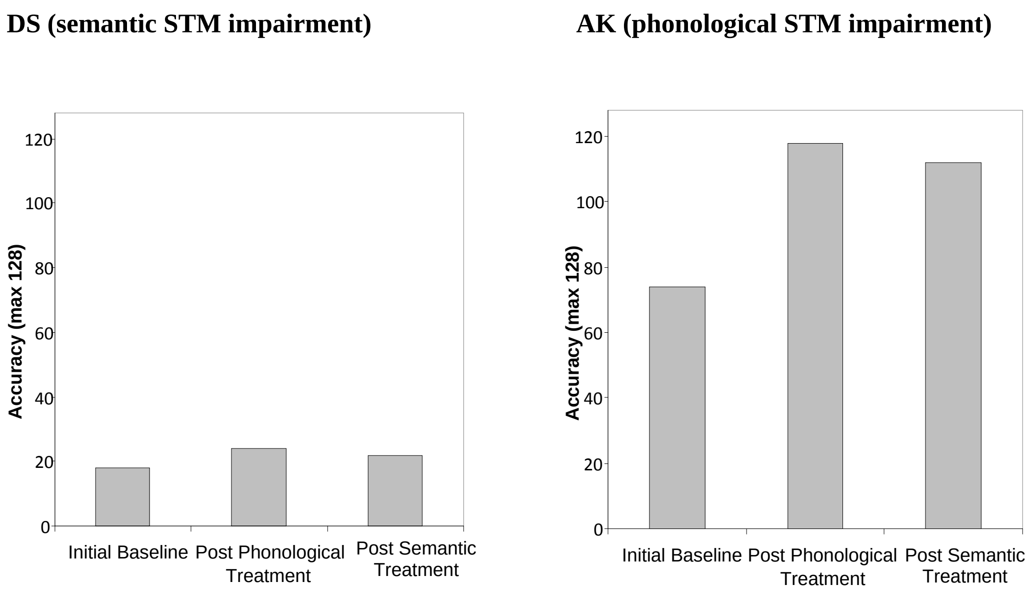


Figure 5.6. Performance on the repetition task (whole-sentence accuracy) at each baseline phase for both patients (DS left panel, AK right panel).

The repetition data (whole sentence accuracy) were entered into a loglinear analysis with the factors accuracy, baseline phase, and patient. The analysis returned a significant 3-way interaction ($X^2(2) = 17.994$, $p < .0001$). Separate loglinear analyses were subsequently conducted for each patient. For AK, there was a significant interaction between baseline and accuracy ($X^2(2) = 55.678$, $p < .0001$, Figure 5.6, right panel), but DS's performance did not change over baseline phases ($p = .587$, Figure 5.6, left panel). To examine the nature of this interaction for AK, separate Chi Square tests were run on the accuracy data at each baseline phase. There was a significant improvement between scores at initial baseline and after the nonword therapy ($X^2(1) = 42.635$, $p < .0001$), and between scores at initial baseline and after the word therapy ($X^2(2) = 30.177$, $p < .0001$), although no differences existed in accuracy scores after the nonword compared to the real word treatment ($p = .198$, Figure 5.6, right panel).

General Discussion

The current study described the STM and sentence comprehension capacities of two STM-impaired patients at baseline, and over two treatments designed to improve phonological and semantic STM. At initial baseline, the two patients could be characterised selectively in terms of, respectively, semantic and phonological deficits in STM. This is consistent with other neuropsychological data on the subject suggesting that there may be dissociable components of STM (Martin and He, 2004, Martin et al., 1994, Martin and Romani, 1994).

Performance on sentence comprehension assessments designed to employ phonological and semantic STM selectively largely followed these STM diagnoses. Patient AK (phonological STM impairment) performed well in sentence anomaly judgements, but was impaired in sentence repetition despite fluency of free speech. AK's impaired score on sentence repetition (defined as employing phonological STM, Hanten and Martin, 2000, Martin and Romani, 1994), taken with preserved speech production, might be attributed to impaired encoding of phonological representations rather than difficulties in production. Patient DS also showed impaired anomaly judgements, consistent with the diagnosis of deficient semantic STM. However, he presented with a nonfluent aphasic profile with agrammatic speech difficulties. Therefore, his performance on the sentence repetition task was probably confounded by problematic speech production.

The sentence comprehension assessment used conditions where the sentences took an immediate integration (‘They grew flowers behind their house’), or a delayed integration (‘Flowers grew behind the house in the yard’, Hanten and Martin, 2000). While the control group showed an advantage for sensible over anomalous sentences across both the anomaly and repetition assessments, anomaly (anomalous or sensible) only affected performance of the patients in the comprehension analyses. Therefore, anomaly had more of an impact on comprehension than repetition in the patients, though any effects in the anomaly tasks may have been obscured by chance level performance in DS’s case, and AK’s ceiling performance. Also, given the nature of the task, the effect of anomaly on the comprehension test could be due to an inclination to commit false-positive errors.

The treatments led to changes in the patterns of patients’ STM profiles. Broadly, semantic effects in STM emerged after semantic STM training in the sSTM patient (DS, Table 5.5), while phonologically-mediated effects were produced in AK (pSTM) after phonological STM treatment (Table 5.6). Following phonological treatment, DS (sSTM) showed a very similar pattern to his results as at initial baseline, with the exception of the lack of an effect of lexicality, due to improved nonword recall. However, after the real word (semantic STM) treatment, DS began to show effects of semantic variables, along with a primacy effect in serial position, described in the literature as a semantically-based effect (Jefferies et al., 2008).

For AK (pSTM), phonological STM treatment was effective in remediating effects of phonological similarity, word length, and the disadvantage for phonological over conceptual span.

There was some generalisation to sentence comprehension performance following STM treatment. AK showed a facilitative effect of phonological STM treatment on sentence repetition (Figure 5.6, right panel), the DS showed an improvement in sentence anomaly judgements following semantic, but not the phonological STM treatment (Figure 5.5, left panel). These findings have two main implications: Firstly, they suggest that improving STM produces a generalised effect to sentence comprehension, and secondly, that these effects were specific to the type of training employed: Semantic STM training improved DS's performance in anomaly, but not repetition, while AK's repetition accuracy increased following phonological STM, but not semantic STM training.

Agrammatism and STM impairments

As discussed in the Introduction section of the current chapter, sentence comprehension difficulties have been traditionally attributed to mnemonic impairments in patients with intact speech processing and production, but deficient syntax processing in patients defined as agrammatic (Caramazza, Basili, Koller and Berndt, 1981). The current chapter was able to disentangle the contributions of STM and syntactic processing to sentence comprehension in an agrammatic patient, by selectively improving STM. The data showed that semantic STM training significantly facilitated performance on making sentence anomaly judgements. This indicates that semantic STM may have a role in sentence comprehension in an agrammatic case, aside from syntactic processing difficulties. However, the finding that phonological STM treatment did not improve repetition suggests that STM has little effect of DS's speech production difficulties.

Conclusions

Nonword and real word treatment approaches were used to selectively improve two forms of STM deficit, one for pSTM and one for sSTM. There was some generalisation to sentence comprehension, a pattern that was specific for each type of training and each diagnosed STM deficit. In an agrammatic patient with deficient semantic STM, difficulties in sentence anomaly judgements appeared due to semantic STM, but that there was little effect of improving phonological or semantic STM on repetition accuracy. There was specificity in the treatment effects observed: Normal semantic STM effects appeared after semantic STM training in a sSTM patient and phonological STM effects were produced following pSTM treatment in AK (pSTM), and this type-specificity was echoed in the generalised improvement to sentence comprehension. Taken together the data suggest that the best approach for a therapist treating deficits in STM and sentence comprehension is to ascertain the type of STM deficit at hand, and employ one type-specific STM treatment to produce directed gains in STM and sentence comprehension.

Chapter 7: Type-specific proactive interference in patients with semantic and phonological STM deficits

Abstract

Prior neuropsychological evidence suggests that semantic and phonological components of short-term memory (STM) are functionally and neurologically distinct. Proactive interference (PI) was examined in two STM-impaired patients, one with a profile consistent with a deficit in semantic STM, characterised by an absence of semantic effects in STM and another with deficient phonological STM. Experiment 1 assessed interference from real words and nonwords, by contrasting effects on verbal recall of phonological interference from nonwords, and semantic and phonological interference from real words. Experiment 2 tested susceptibility to phonological and semantic interference using an item recognition probe task. The findings indicated type-specific PI: there was heightened phonological PI for the semantic STM patient, and exaggerated effects of semantic PI in the phonological STM cases. The data are consistent with an account of rapid decay of type-specific representations in phonological and semantic STM.

Introduction

Recent work suggests that memory span tasks assess not only the ability to deal with current information but also the management of proactive interference (PI) from previous information, and that the ability to overcome this interference strongly predicts memory span performance (May, Hasher and Kane, 1999, Lustig, May and Hasher, 2001). From early studies it has been shown that PI disrupts both long and short term memory (Kane and Engle, 2000, Keppel and Underwood, 1962 respectively). More recently interest has been rekindled through studies demonstrating that different forms of PI can be distinguished. One distinction is between item non-specific PI, which occurs when different items are presented belonging to the same category and item specific PI, produced by repeated presentation of the same set of items across trials (Postle, Berger, Goldstein, Curtis and D'Esposito, 2001, Postle and Brush, 2004). The present chapter focuses on a different distinction, between PI based on the semantic representations of items and PI based on phonological representations. This distinction is derived from contrasting patterns of performance between patients with respectively impaired phonological and semantic short term memory (pSTM and sSTM).

The dissociation between pSTM and sSTM patients was first described by Martin and colleagues (Martin and Romani, 1994, Martin, Shelton and Yaffee, 1994). These patients can be distinguished in several ways including the absence of recency and phonological similarity effects in pSTM patients, and absent primacy and lexicality effects in sSTM patients. The distinction also emerges in probe recall tasks. Phonological STM patients are impaired when probed to recall which memorised items in a memory set *rhyme* with a probe word, while sSTM patients are impaired with category probes (e.g. 'which items in the list were animals?', Martin and He, 2004). As a memory component responsible for phonological representations may be separate from a component that deals with semantic representations, then recurring phonological information could selectively disrupt phonological encoding (phonological PI) and repeating items from the same semantic category may selectively affect semantic encoding (semantic PI).

It is possible that separate forms of PI, phonological and semantic, exist, but the data provide equivocal support. Barde, Schwartz, Chrisikou and Thompson-Schill (2010) conducted a patient group study reporting data from twenty aphasic patients, eleven with weak sSTM and nine with weak pSTM. The authors used an adaptation of a letter probe recognition task (Monsell, 1978), requiring the patients to remember three words presented consecutively on a computer screen. A probe item followed, and the patients were required to indicate whether the item did ('yes' trial) or did not ('no' trial) appear in the memory set. Interference was created by phonological and semantic similarity between words (manipulated separately). Barde et al. found that weak sSTM correlated with exaggerated semantic interference, and weak pSTM correlated with exaggerated phonological interference. Neither deficit correlated with the opposite type of interference. The authors concluded that there may be representational specificity in interference, and that sSTM and pSTM deficits are associated with different failures in preventing interference. Barde et al. propose a 'reactivation hypothesis' based on inhibition to accommodate the data, assuming that items are maintained by reactivation of their representations. They assert that a deficit in reactivating a particular form of STM representation will lead to a weak memory code which is susceptible to interference from other similar memory codes.

Hamilton and Martin (2007) investigated the possibility that phonological and semantic representations in STM may be selectively sensitive to, respectively, phonological and semantic PI in pSTM and sSTM patients. Using a similar probe task to the experiment used by Barde et al. (2010), patient ML (a sSTM patient) made more false positive errors and had slower RTs to any related-no trials than the controls. This finding is consistent with other studies reporting semantic STM-impaired patients with associated difficulties in resisting PI generally (e.g. Shimamura, Jurica, Mangels and Knight, 1995). Their second experiment showed that ML exhibited exaggerated PI effects with stimuli that were both semantically and phonologically related to previously presented items, relative to a control group. In contrast, their pSTM patient showed no exaggerated PI in any manipulation. This finding was against their expectations that ML would either show phonological but not semantic interference (*rapid decay* of semantic information), or semantic but not phonological interference (difficulties *inhibiting* semantic information). They conceptualise sSTM in terms of deficient inhibition, and pSTM in terms of rapid decay of phonological representations (Hamilton and Martin, 2002, 2007). They conclude that PI susceptibility is due to a deficit in the control processes involved in STM, e.g. in inhibiting both previously activated semantic and phonological representations. Patients with sSTM have deficient control processes and therefore present with a general susceptibility to PI across verbal tasks. Barde et al.'s findings, however, question the idea that semantic STM patients show general interference, from phonological and semantic features alike.

It is not clear within Barde et al.'s (2010) hypothesis why, for example, discrimination between phonologically related words should be weaker for a patient who cannot retain phonological information adequately than a participant with spared pSTM. It may be that these patients are less able to retain the subtle phonological differences between phonologically related items (e.g. the initial and final consonants of ‘bear-pear’ and ‘beat-beep’), but positing interference from, and therefore encoding of the core phonological information about the word that creates similarity is inconsistent with the absent phonological similarity effects and poor performance with rhyme probes that are characteristic of pSTM patients. The expectation for a pSTM patient should still be reduced interference from phonological similarity due to poor encoding of key phonological features. Although this depends to some extent on why phonological similarity is absent: For instance, if phonological similarity is absent and performance on pSTM tests is poor, increased interference from phonological information may not be inconsistent with absent phonological similarity effects, as poor performance with rhyme probes may be analogous with poor performance with phonologically dissimilar items.

Most of the verbal tasks used to explore proactive interference in STM-impaired patients have employed real word stimuli and manipulated phonological and semantic relatedness between the words to elicit PI (e.g. Barde et al. 2010, Hamilton and Martin, 2007). There may be problems with using real word stimuli to study type-specific PI as they possess both semantic and phonological features, and may be maintained by either form of representation. A role for semantic representations is suggested by findings from pSTM patients who show a normal lexicality effect (memory for words is better than for nonwords), despite failing to demonstrate phonological coding (e.g., where there is no evidence of a phonological similarity effect). Semantic STM patients, in contrast, can fail to demonstrate a lexical advantage and do show effects of phonological similarity, indicating reliance on phonological rather than lexical/semantic representations. In containing both semantic and phonological information, real words might allow patients to compensate for poor retention of, for example, phonological information by maintaining the semantic features of the items instead. This could go some way to explain the divergent findings found across existing studies (Barde, Schwartz, Thompson-Schill, 2006, Barde et al., 2010, Hamilton and Martin, 2007).

Another possibility is that different type-specific STM-patients are affected differently by PI as a function of the severity, or type of STM deficits that they present with. For instance, it is plausible that a patient with a representational STM deficit for a specific type of information might show abnormally *reduced* interference from the impaired type of information, owing to failures to encode this type of representation. On the other hand, a patient with a control impairment in STM might show exaggerated PI effects for the impaired type of information, because of an abnormal persistence of previously encoded similar information, and an impoverished ability to discriminate between these similar items.

To assess patients' performance where memory for another type of information cannot compensate for deficiencies in the other (where semantic information cannot be relied upon in a case of pSTM), interference from nonword stimuli was used. A probe task contrasting interference from real and non-word stimuli allows us to look at interference effects where both semantic and phonological information may be used in retention (real words), and where only phonological information can be employed (nonwords). Working under this assumption that closed sets of nonwords should elicit phonological interference, and closed sets of real words should produce both phonological and semantic PI, three hypotheses were identified: i) type-specific STM patients might show *exaggerated* interference effects from closed sets tapping the impaired information (nonwords = phonological, owing to reduced distinction between similar encoded items, as in Barde et al., 2010), ii) *reduced* interference from the impaired information (due to chronic failures to encode) and iii) increased interference from both types of information for the semantic STM patient, but no elevated effects for pSTM patients (because of impaired control processes in these patients, Hamilton and Martin, 2007).

A probe task paradigm was also employed, based on Monsell (1978) with real words to assess whether interference from phonological and semantic relatedness was type-specific in the patients (as in Barde et al.'s patient, 2006) or whether PI effects are general in semantic STM patients, but type-specific in phonological STM patients (Martin et al. 1994). For instance, is interference enhanced, as in Barde et al. 2006, 2010, or absent as one might predict from the lack of phonological similarity effects in these patients?

Current study

I assessed the performance of three neuropsychological patients with short-term memory deficits under conditions designed to elicit PI. In Experiment 1, a probe task was used, designed to look at the effect of item-specific interference from nonwords and words on recall. In Experiment 2 a probe-recognition test similar to that used by Barde et al. (2010) was employed, where the phonological and semantic relatedness of words were manipulated to produce, respectively, phonological and semantic PI.

Case Descriptions

Patient DS

At the time of testing DS was a 73-year-old male who suffered a Left Inferior Frontal CVA that resulted in non-fluent Broca's aphasia. DS's presentation was described in the previous chapter, and so was not reproduced here. The results from his MRI scan are provided in Figure 6.1). DS presented with a memory span of 2 items, and difficulties in short term memory that can broadly be characterised in terms of a semantic STM deficit (see Table 6.1 for a summary).

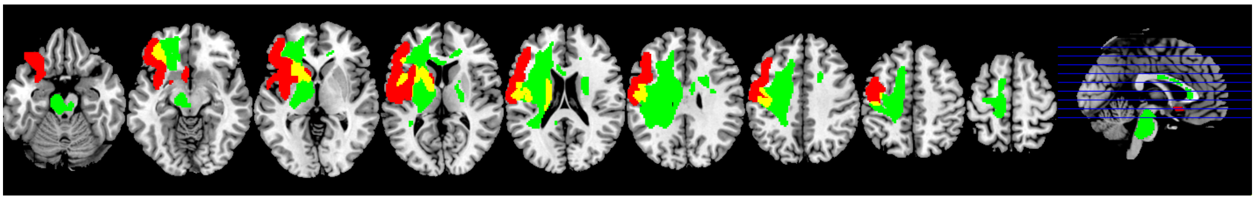


Figure 6.1. Brain images of patient DS: Corrected FWE p = 0.01 Extent threshold including only significant blobs containing ≥100 voxels. N.B. Grey matter lesion appears in red and white matter lesion in green. The lesion was created in SPM and added as an overlay onto a standard multi-slice template in MRIcron. The SPM analysis was a one sample t- test with the covariates healthy (140 brains aged 40+) vs. patient, age and gender.

Patient AK

Patient AK is a 74-year old male who was referred to the Behavioural Brain Sciences Centre at the University of Birmingham in 2008. He had previously worked as an accountant and had also studied law. At the time of referral, AK reported difficulties in memory and attention. His speech is fluent and grammatically correct. An fMRI scan revealed an extent of atrophy that was advanced for his age (Figure 6.2). Despite an otherwise preserved presentation, AK had a memory span of 3 items, and difficulties in phonological aspects of STM (Table 6.1).

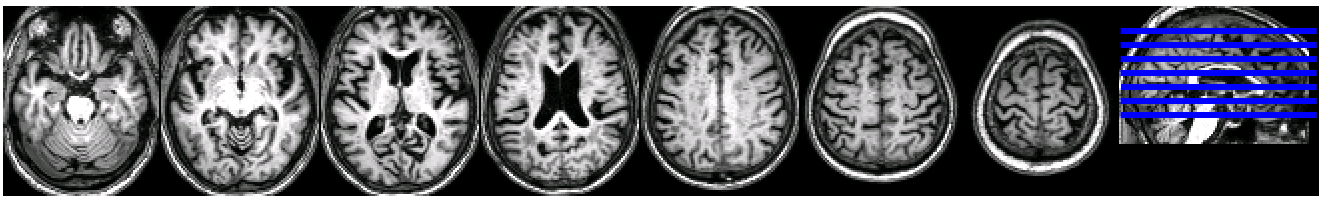


Figure 6.2. Brain images of patient AK: Corrected FWE p = 0.01 Extent threshold including only significant blobs containing ≥100 voxels. N.B. Grey matter lesion appears in dark grey and white matter lesion in light grey. The lesion was created in SPM and added as an overlay onto a standard multi-slice template in MRIcron. The SPM analysis was a one sample t- test with the covariates healthy (140 brains aged 40+) vs. patient, age and gender.

Patient MM

At the time of testing, patient MM was a 77-year old male who had previously been employed as a factory worker. He was referred to the centre in 2009 following a CVA. White matter damage was present within long association pathways bilaterally (most likely within the bilateral superior longitudinal, corona radiata, thalamic radiation, corpus callosum and the right internal capsule). Grey matter lesions were detected in the left and right thalamus, cerebellum and there was evidence suggesting small lesions within right hippocampus / parahippocampal gyrus (Figure 6.3). MM's speech is fluent in spontaneous production. His comprehension is good at the single word level (PALPA 47 39/40, Kay et al. 1996) but is significantly reduced with sentences (PALPA 55, auditory sentence-picture matching 32/60, Kay et al. 1996) and is slightly improved with written presentation (PALPA 56 41/60, Kay et al. 1996). This pattern of performance could be explicable by his memory problems: MM presents with chronically impaired memory span (of 2 items), and a pattern of performance on STM tasks consistent with a phonological STM deficit (Table 6.1).

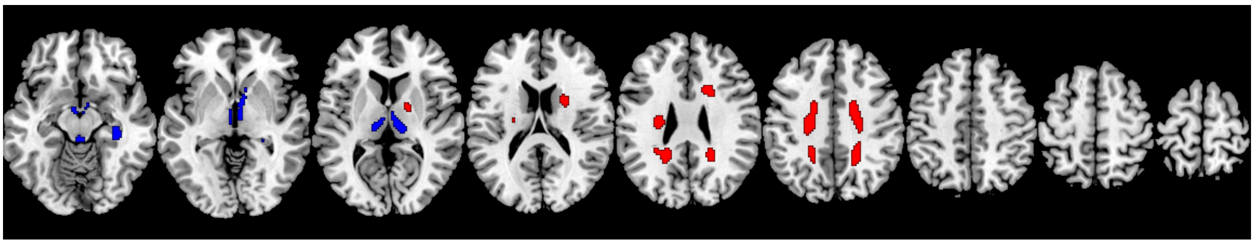


Figure 6.3. Brain images of patient MM: Corrected FWE $p = 0.01$ Extent threshold including only significant blobs containing ≥100 voxels. N.B. Grey matter lesion appears in blue and white matter lesion in red. The lesion was created in SPM and added as an overlay onto a standard multi-slice template in MRIcron. The SPM analysis was a one sample t- test with the covariates healthy (140 brains aged 40+) vs. patient, age and gender.

Table 6.1. Patients' performance across tasks designed to assess semantic and phonological STM. A '√' indicates the presence of the proposed pattern of performance reflecting either a semantic or phonological ST deficit.[5]

Suggests semantic or phonological STM deficit	***Patterns of performance***	***DS Semantic STM impairment***		***AK Phonological STM impairment***		***MM Phonological STM impairment***	
		X^2		X^2		X^2	
Semantic	i) Reduced / absent primacy effect	List length: 5 Primacy: 8/20 (40%) Medial positions: 44/80 (55%), $X^2(1) = 12.698$, $p < .001$	√	List length: 5 Primacy 16/20 (80%), Medial positions: 48/80 (60%), $X^2(1) = 2.778$, $p = .096$.	×	List length: 4 Primacy: 0/20 (0%) Medial positions: 12/40 (30%), $X^2(1) = 7.500$, $p = .006$.	√
Semantic	ii) Absent lexicality effect (words × nonwords)	Real words: 149/180 Nonwords: 96/180, $X^2(1) = 35.891$, $p < .001$	×	Real words: 242/280 Nonwords: 252/180, $X^2(1) = 1.718$, $p = .190$	√	Real words: 29/90 Nonwords: 30/90, $X^2(1) = .025$, $p = .874$	√
Semantic	iii) Absent frequency effect (high frequency × low frequency)	High frequency: 76/180 Low frequency: 72/180, $X^2(1) = .184$, $p = .668$	√	High frequency: 188/280 Low frequency: 156/280, $X^2(1) = 7.717$, $p = .005$	×	High frequency: 56/90 Low frequency: 21/90, $X^2(1) = 27.807$, $p < .001$	×
Semantic	iv) Conceptual span is worse than phonological span	Phonological: 22/40 Conceptual: 10/40, $X^2(1) = 7.500$, $p = .006$	√	Phonological 8/40; Conceptual 25/40	×	Phonological 4/40; Conceptual 26/40	×
Semantic	v) Worse in 'noun after' conditions in comprehension of adjective-noun structures	Noun after 5/20, Noun before 15/20, $X^2(1) = 9.800$, $p = .002$	√	Noun after 17/20, Noun before 18/20, $X^2(1) = .237$, $p = .626$	×	Noun after 7/20, Noun before 8/20, $X^2(1) = .107$, $p = .744$	×
Phonological	vi) Reduced / absent recency effect	List length: 5 Recency: 12/20 (60%) Medial positions: 44/80 (55%), $X^2(1) = 3.509$, $p = .061$.	×	List length: 5 Recency: 12/20 (60%) Medial positions: 48/80 (60%), $X^2(1) = .000$, $p = .1$	√	List length: 4 Recency: 15/20 (75%), Medial positions: 12/40 (30%), $X^2(1) = 10.909$, $p < .001$	×
Phonological	vii) Absent phonological similarity effect (phonologically similar × phonologically dissimilar)	Phonologically similar: 119/180 Phonologically dissimilar: 94/180, $X^2(1) = 7.186$, $p = .007$	×	Phonologically similar: 234/280 Phonologically dissimilar: 231/280, $X^2(1) = .114$, $p < .736$	√	Phonologically similar: 55/90 Phonologically dissimilar: 46/90, $X^2(1) = 1.827$, $p = .176$	√
Phonological	viii) Absent word length effect (short words × long words)	Short words: 117/180 Long words: 101/180, $X^2(1) = 2.977$, $p = .084$	√	Short words: 251/280 Long words: 223/280, $X^2(1) = 10.770$, $p < .001$	×	Short words: 26/90 Long words: 43/90, $X^2(1) = 6.792$, $p = .009$	×
Phonological	ix) Phonological span is worse than conceptual span	Phonological 22/40 Conceptual 10/40	×	Phonological 8/40 Conceptual 25/40 $X^2(1) = 14.907$, $p < .001$	√	Phonological 4/40 Conceptual 26/40 $X^2(1) = 25.813$, $p < .001$	√
Phonological	x) Better with written vs. spoken presentation	Spoken: 149/180 Written: 45/180, $X^2(1) = 1209.200$, $p < .001$	×	Spoken: 242/280 Written: 187/280, $X^2(1) = 30.143$, $p < .001$	×	Spoken: 29/90 Written: 45/90 $X^2(1) = 3.366$, $p = .067$	√

[5] √ denotes impairment (showing no significant effect of the manipulation), × denotes no impairment (showing a significant effect of the manipulation).

See Appendix E for methodological details for the tasks reported in Table 6.1.

Though every patient showed some effect consistent with the opposite STM type, all three patients showed more aspects of one type of deficit over the other: On this basis DS was characterised as presenting with impaired semantic STM, and AK and MM as having deficient phonological STM. Interestingly, the pattern of lexicality effects produced across the patients ran against the other data, in that patients with a phonological STM deficit did not show lexicality effects, whereas DS (semantic STM) did show a lexicality effect. Another unpredicted finding was the presence of a phonological similarity effect in the reverse direction (advantage for phonologically similar over dissimilar words) in DS (sSTM). However, this finding might be explicable by DS's difficulties in speech production, in that once key aspects of a phonological form are accessed and produced; access / production of phonologically related words might be easier. The effects of primacy, frequency, phonological similarity and word length, along with patterns of performance in phonological and semantic span tests, were all consistent with the patients' diagnoses.

Experiment 1: Probe tasks using real and non- words.

I assessed whether the pSTM patients showed reduced PI from repeating nonwords, as these stimuli contain phonological but not semantic representations. In contrast, the semantic STM patient should show an interference effect with nonwords, as encoding nonwords should elicit the use of preserved memory for phonological representations rather than semantic representations, which is impoverished in semantic STM patients. In real words, though, both patients should manifest PI as real words contain both semantic and phonological representations. To examine item-specific interference in these cases, performance was tested on a probe task with real and non- word stimuli, contrasting data from open (where items do not repeat) and closed (where items recur across trials) sets. Better performance with open relative to closed sets is taken to indicate an effect of proactive interference.

Method

Participants

Three STM patients (DS, AK and MM, described in the Case Description) participated in the study. Five male participants were used as controls for patients DS, AK and MM (mean age = 73, standard deviation = 3.38), and were paid for their participation.

Design

A 2 (Lexicality) × 2 (Set Type) × 2 (Probe Type) design was used. Performance on open and closed sets was compared, using real words with item probes in Experiment 1a and nonword stimuli in Experiment 1b.

Materials

The word stimuli were comprised of 480 five-letter one-syllable words in the open conditions, and the closed set was created from a sub-set of 15 of these words, which were matched to the open set on a range of lexical variables (see Appendix F for *M* ratings on lexical variables in each condition).

No words with the same initial letter or rhyme-sound appeared within each trial in either condition, and no repetitions of items occurred within trials in the closed set experiments.

The open set for the nonword stimuli contained 480 five-letter one-syllable words, with a closed set created from a sub-set of 15 of these words. As with the real word experiments, there were no repetitions of initial word letter within each trial in either condition, and no repetitions of items within trials in the closed set experiments.

Procedure

The experiments were created in E-prime and were displayed on a 1024 × 768 CTX computer screen. Each trial began with a fixation cross that appeared in the centre of the screen, and then the items were presented, one at a time for 1000 ms, with 500ms gap between items. As a 2-forced choice response method was adopted, accuracy was required to be above 60% in all open set conditions to ensure above-chance performance. For this reason, list length was altered for each patient according to individual ability: DS was tested using 4 items in both conditions. AK was tested on 5 word lists for both manipulations. MM was tested using 3 items for the word, and 2 items for the nonword condition. The controls were always tested using 5-item trials.

Items appeared in white against a black background. When all of the items had been displayed, a question mark appeared in the centre. Two recognition probes were then given, one on the left and one on the right side of the screen. For the open trials, the distracter probes were items that had never appeared before in the experiment. For the closed trials, distractors belonged to the same closed set as targets, but did not appear with the target list on that trial. Participants were instructed to press '1' if the item on the left was in the current list, and '2' if the item on the right was correct of the current list. Probes remained on the screen until a response was detected. List positions for the probed items were counterbalanced such that the probes were equally likely to assess memory for items at each position in all conditions. After each trial a slide displaying the word ‘continue?’ was presented and participants were invited to press the space bar when they were ready for the next trial. Figure 6.4 provides an example trial sequence.

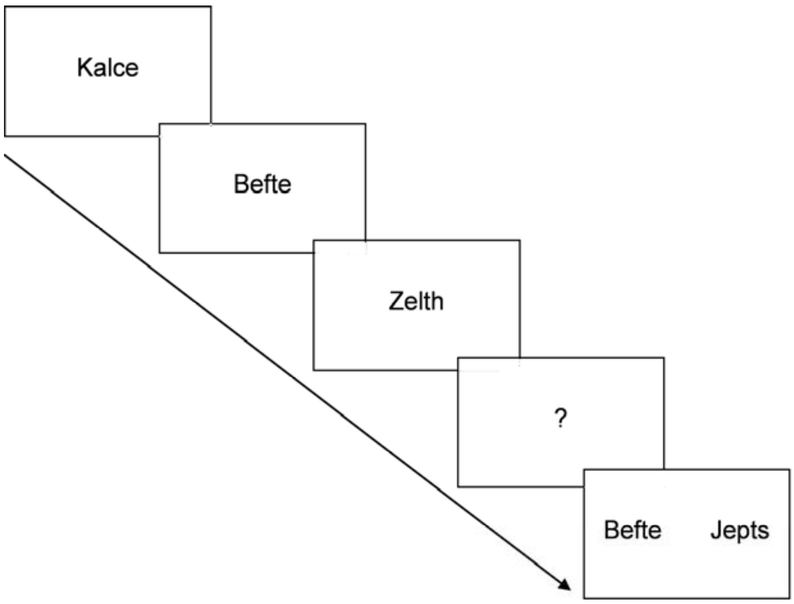


Figure 6.4. Example of a trial sequence for the nonword experiment (Experiment 1b, list length 3). In closed sets, the distracter probe was an item from a previous list. In open sets, the distracter probe was a new item.

Results

Experiment 1a: Word probe task

Accuracy data from patients and controls for Experiment 1 are provided in Figure 6.5.

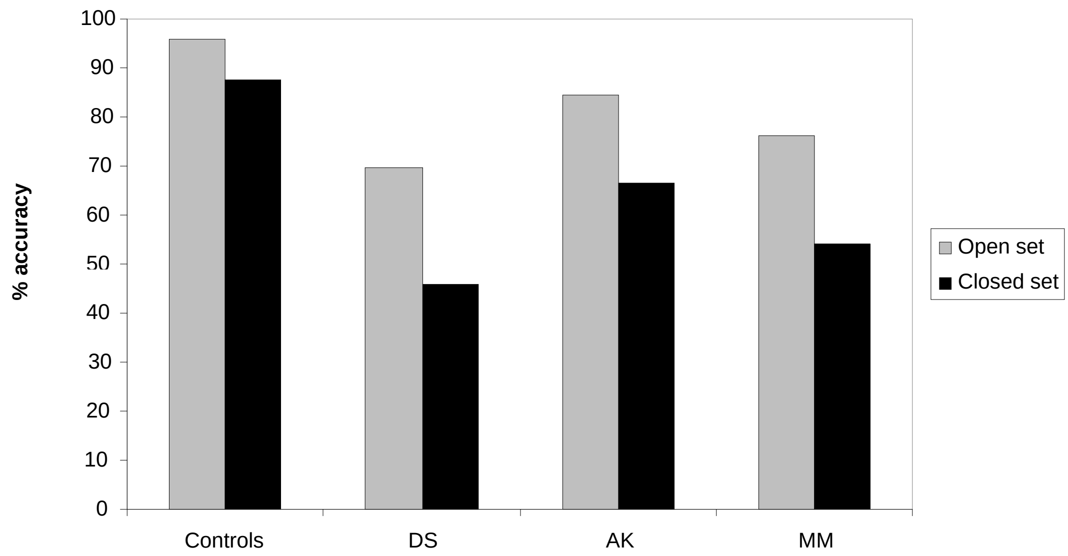


Figure 6.5. Percent accuracy data from Experiment 1a (real words and item probes, controls mean % accuracy). DS 4; AK 5; MM 3-word lists.

A loglinear analysis was performed on the patient accuracy data with the factors accuracy, patient type (semantic STM or phonological STM) and set type (closed vs. open). There were significant interactions between accuracy and set type (X^2 (1) = 29.695, $p < .0001$) and patient type (X^2 (1) = 9.279, $p = .002$), but no three way interaction between accuracy, set type and patient type (X^2 (1) = .005, $p = .942$).

Chi Square tests were conducted to examine accuracy scores in open and closed sets for each patient. All patients performed less accurately in the closed relative to open sets (DS, X^2 (1) = 11.297, $p < .001$; AK, X^2 (1) = 8.142, $p = .002$; MM, X^2 (1) = 10.110, $p < .001$). A significant advantage for the open set condition was found in the control data (X^2 (1) = 3.544, $p = .03$).

To compare the magnitude of the interference effect between the patients and controls, a univariate ANOVA was performed on the percent difference (open – closed) data from the control group and each patient (analysed separately). Tests of normality conducted on the data were not significant (Shapiro-Wilk test p's > .8). The interference effect for the controls was not significantly different to that of DS ($F(1, 4) = 1.816$, $p = .249$), AK ($F(1, 4) = 1.106$, $p = .352$) or MM ($F(1, 4) = 1.323$, $p = .314$).

Serial position data from Experiment 1a: Real word probe task.

DS **AK** **MM**

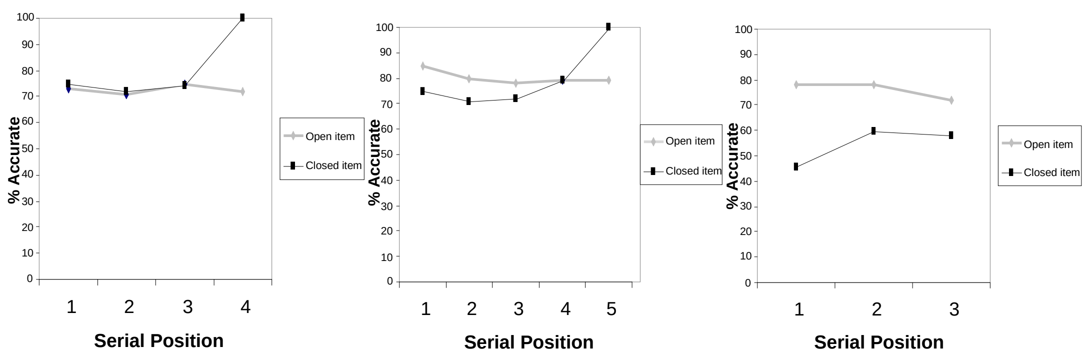


Figure 6.6. Percent accuracy data at each serial position in Experiment 1a for DS, AK and MM respectively.

Serial position effects were assessed by entering accuracy scores at medial positions, and initial or final positions into Chi Square analyses. No significant effects of primacy or recency emerged from analysis of the control data (all p's > .9).

DS showed no significant effect of primacy (initial 20/24 (83.33%), medial 35/48 (72.91%), $X^2(1) = .963$, $p = .327$) or recency (final 18/24 (71.88%), medial 35/48 (72.91%), $X^2(1) = .036$, $p = .850$, Figure 6.6, left panel) in the open set. In the closed set, there was no significant primacy effect (18/24 (75%), medial 35/48

(72.91%), $X^2(1) = 2.314$, $p = .128$), but a significant effect of recency (final 24/24 (100%), medial 35/48 (72.91%), $X^2(1) = 16.898$, $p < .0001$, Figure 6.6, left panel).

AK showed no significant primacy or recency effect in the open condition (initial 17/20 (85%) and medial 45/57 (78.94%), X^2 (1) = .169, $p = .681$, and final 19/20 (71.88%) and medial 45/57 (78.94%), X^2 (1) = .185, $p = .667$, respectively, Figure 6.6, centre panel). In the closed set, AK showed a significant recency effect (final 9/20 (100%), medial 40/54 (74.07%), X^2 (1) = 4.055, $p = .044$), but no effect of primacy (15/20 (75%), medial 40/54 (74.07%), X^2 (1) = .169, $p = .681$, Figure 6.6, centre panel).

MM showed no significant serial position effects (all p's > .1, Figure 6.6, right panel).

Experiment 1b: Nonword probe task.

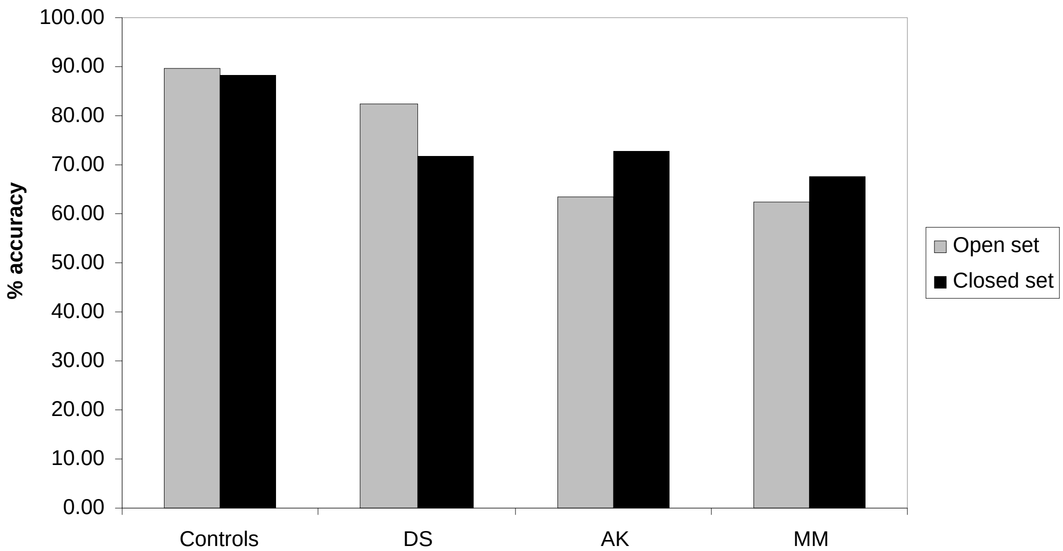


Figure 6.7. Percent accuracy data from Experiment 1b (nonwords and item probes, controls mean % accuracy). DS 4; AK 5; MM 2-word lists.

Percent accuracy scores for the patients and controls in each condition are provided in Figure 6.7. No significant difference was found in accuracy between closed and open sets in the control data for Experiment 1b X^2 (1) = .817, $p = .817$. A loglinear analysis on the patient accuracy data revealed a significant three way interaction between accuracy, set type and patient type (X^2 (1) = 5.135, p = .023).

A significant interference effect was only found with patient DS (X^2 (1) = 2.948, p < .05). Both pSTM patients showed better performance with closed sets, but this pattern failed to reach significance in both cases (AK, X^2 (1) = 1.946, p = .081; MM, X^2 (1) = .573, p = .224).

The % difference (open – closed) data between the control group and each patient (analysed separately). The interference effect for the controls was not significantly different to that of DS (F(1, 4) = 1.915, p = .239), AK (F(1, 4) = 3.861, p = .121) or MM (F(1, 4) = 26.772, p = .490).

Serial Position data for Experiment 1b: Nonword probe task

Semantic STM **Phonological STM** **Phonological STM**

DS **AK** **MM**

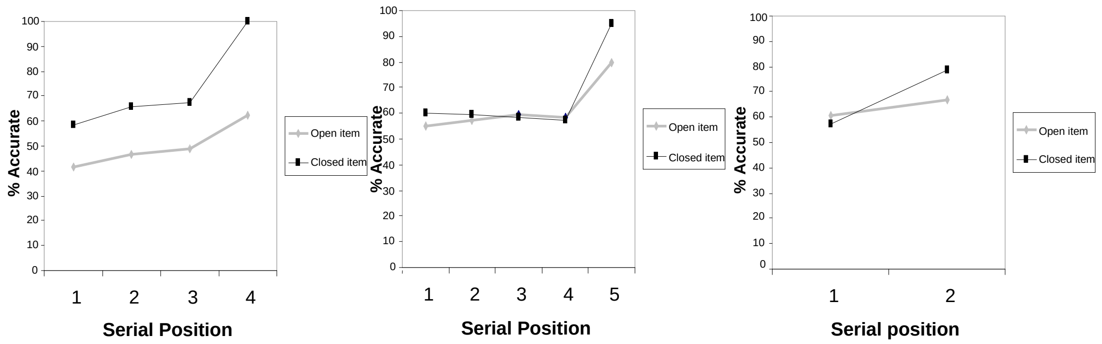


Figure 6.8. Percent accuracy data at each serial position in Experiment 1b for DS, AK and MM respectively.

Figure 6.8 shows percent accuracy scores across serial position for each patient. Chi Square tests found no significant associations between accuracy and position for DS in the open set, Primacy: initial 10/24 (41.66%) and medial 23/48 (47.91%), X^2 (1) =

.570, p = .450; Recency: final 15/24 (62.5%), medial 23/48 (47.91%), X^2 (1) = .890, p = .346. In the closed set, DS showed a significant effect of recency (final 24/24 (100%), medial 32/48 (66.66%), X^2 (1) = 9.879, p = .002), but not primacy (initial 14/24 (58.33%), medial 32/48 (66.66%), X^2 (1) = .667, p = .414, Figure 6.8, left panel).

In the open set condition, AK showed no significant effect of serial position: Primacy: initial 11/20 (55%) and medial 35/60 (58.33%), X^2 (1) = .083, p = .773; Recency: final 16/20 (80%) and medial 35/60 (58.33%), X^2 (1) = 1.805, p = .179. In the closed set, there was no significant primacy effect (initial 12/20 (60%), medial 35/60 (58.33%), X^2 (1) = .699, p = .403), but there was a significant recency effect (final 19/20 (95%), medial 35/60 (58.33%), X^2 (1) = 5.094, p = .024, Figure 6.8, central panel).

Chi Square comparisons with the medial position data were not possible for MM as he was tested using 2 items in each list. However, when comparing these initial and final positions there were no significant differences between accuracy and position for MM in the open set data (X^2 (1) = .389, p = .533), but there was a significant advantage for list-final positions in the closed set data (X^2 (1) = 6.558, p = .010, Figure 6.8, right panel).

Overall, these results reflect a stronger recency effect than primacy effect in closed sets for all patients, indicating the use of phonological information in encoding.

Discussion

The pSTM patients (AK and MM) did not show interference from repeating nonwords, whereas the sSTM patient (DS) showed an elevated interference effect from closed sets of nonwords relative to controls (Figure 6.7). These contrasts support the distinction between the different STM deficits,

With real word stimuli, all patients showed an effect of interference that was larger than the control group (Figure 6.5). The accuracy data for the control group was not always at ceiling, and showed some signs of interference (Figures 6.5 and 6.7). The greater magnitude of PI effects produced by the sSTM patient compared with controls in both studies, suggests that interference is maximised when phonological and semantic STM patients rely disproportionately on preserved retention of semantic / phonological information.

Interestingly, there was a non-significant trend for the pSTM patients to report closed set nonwords better than nonwords in the open set. This may reflect the difficulty such patients have in establishing stable phonological representations to support their STM performance. For example, in their description of a patient with a proposed impairment at the phonological loop (patient PV), Baddeley, Papagno and Vallar (1988) reported difficulties in learning new words despite preserved abilities in remembering familiar words. With a closed set of items, the patient may be able to build up phonological representations across trials, supporting their performance over and above the fact that items are repeating. In contrast, in participants with stable phonological representations there will be little benefit from item repetition and disadvantages due to proactive interference.

It has been suggested that recency and primacy effects in serial recall are dependent on phonological and semantic information respectively (e.g. Martin and Lesch, 1996). In Martin and Lesch (1996), a sSTM patient showed exaggerated recency effects in relation to primacy (reflecting dependency on phonological representations), whereas primacy effects were more pronounced in the pSTM patients (reflecting dependency on semantic representations). The serial position analyses for the data with nonwords (Experiment 1b) reinforce this. The data demonstrate that under conditions of phonologically-based interference, all patients showed an exaggerated recency effect compared to primacy effects. This is consistent with reliance on phonological information. However, this pattern was not replicated with real words (Experiment 1a), where again, all patients showed only recency effects.

Experiment 2: Item recognition with semantically and phonologically-related probes

Experiment 1b found that closed sets of nonwords failed to elicit interference in the pSTM patient (AK), but produced exaggerated PI effects in the sSTM patient (DS). On the other hand, real words produced interference effects in both patients Experiment 1a). This pattern of results can be attributed to nonwords containing phonological, but not semantic information, and real words, in containing both, lead the performance of STM-patients to proceed in a compensatory manner, encoding words based on the preserved type of information, resulting in exaggerated PI effects in these patients for the preserved type of information in STM. To further explore this idea, Experiment 2 used an item-recognition test based on Monsell (1978) to contrast semantic and phonological information, manipulated through phonological and semantic relatedness between items, in a real word context. One possibility is that the STM patients will show reduced interference when relatedness on the impaired dimension is manipulated. This pattern could be explicable by an impoverished ability to encode items on the basis of this type of information, leading to reduced sensitivity to interference in these conditions. This outcome would be consistent with results from Experiment 1. Experiment 2 also sought to explore in more detail the idea that PI effects may act in a compensatory manner, through contrasting the magnitude of interference effects for PI from the preserved information with the interference effect elicited for each type of interference condition in controls.

Method

Participants

Participants were the same as described in Experiment 1.

Design

Participants received a list of words followed by a probe. The probe was present in the list on half of the trials and absent on the other half. On both present and absent trials, the probe could be semantically related, phonologically related or unrelated to items in the list. On related trials, the probe was related to a stimulus in the current list, or 1 or 2 lists back, and the same held for the related probe trials. 'Yes' trials were those trials where the probe appeared in the current list, and 'no' trials were those where the probe did not appear in the current list. Relatedness was manipulated within both no and yes trials, yes-related trials were those where the probe had appeared in the current list, but was also related to an item on a previous trial.

Materials

A modified item-recognition task was used, similar to Barde et al. (2010), to create phonological and semantic interference (Figure 6.9). Phonological interference was produced when probes rhymed with one of the words in the list, and semantic interference was produced when probes matched the category of one of the list words. Materials consisted of 4-5 letter words. As in Barde et al. (2010), the inclusion of orthographically similar items was avoided, but where this was not possible, they occurred equally across related phonological and semantic conditions. The phonological and semantic manipulations were conducted separately, creating two sub-experiments. On each trial, three consecutive words were presented, followed by a probe word. There were four conditions: 1). Related-yes, where the probe word did appear in the current list, and was also related to another word viewed 1, 2, or 3 trials ago. 2). Related-no, where the probe word did not appear in the current list, but was related to another word that appeared 1, 2, or 3 trials back. 3). Unrelated-yes, where the probe word appeared in the current list, but was unrelated to any previously viewed word in the preceding 3 trials. 4). Unrelated-no, where the probe word did not appear in the current list, and was unrelated to previously viewed words in the preceding 3 trials. There were 36 trials in each condition (Related-no, Unrelated-no, Related-yes, Unrelated-yes), for both semantic and phonological relatedness. Two filler trials began each of the three randomised blocks. There were 150 trials in each experiment. The numbers of semantic and phonologically related pairs were roughly equivalent, and there were similar numbers of exemplars from each category. Probes, and test items that were related to the probes, were closely matched on a range of lexical and phonological variables across relatedness conditions, and across the phonological and semantic interference manipulations (see Appendix G for *M* lexical variables in each condition).

Procedure

Before each experiment, participants were advised that they would see lists of three words on the screen. They were instructed to memorise each word and then respond as quickly as possible as to whether the fourth probe word had appeared in the list, pressing '1' for a yes and '2' for a no response. Participants were asked to respond as quickly as possible, and were not informed that there would be any phonological and semantic relatedness between items. Participants performed the phonological and semantic versions of the experiment on separate occasions, and at least 4 days intervened between testing sessions. Controls completed each condition once, but the conditions were conducted with each patient three times. See Figure 6.9 for an illustrated example of a trial sequence.

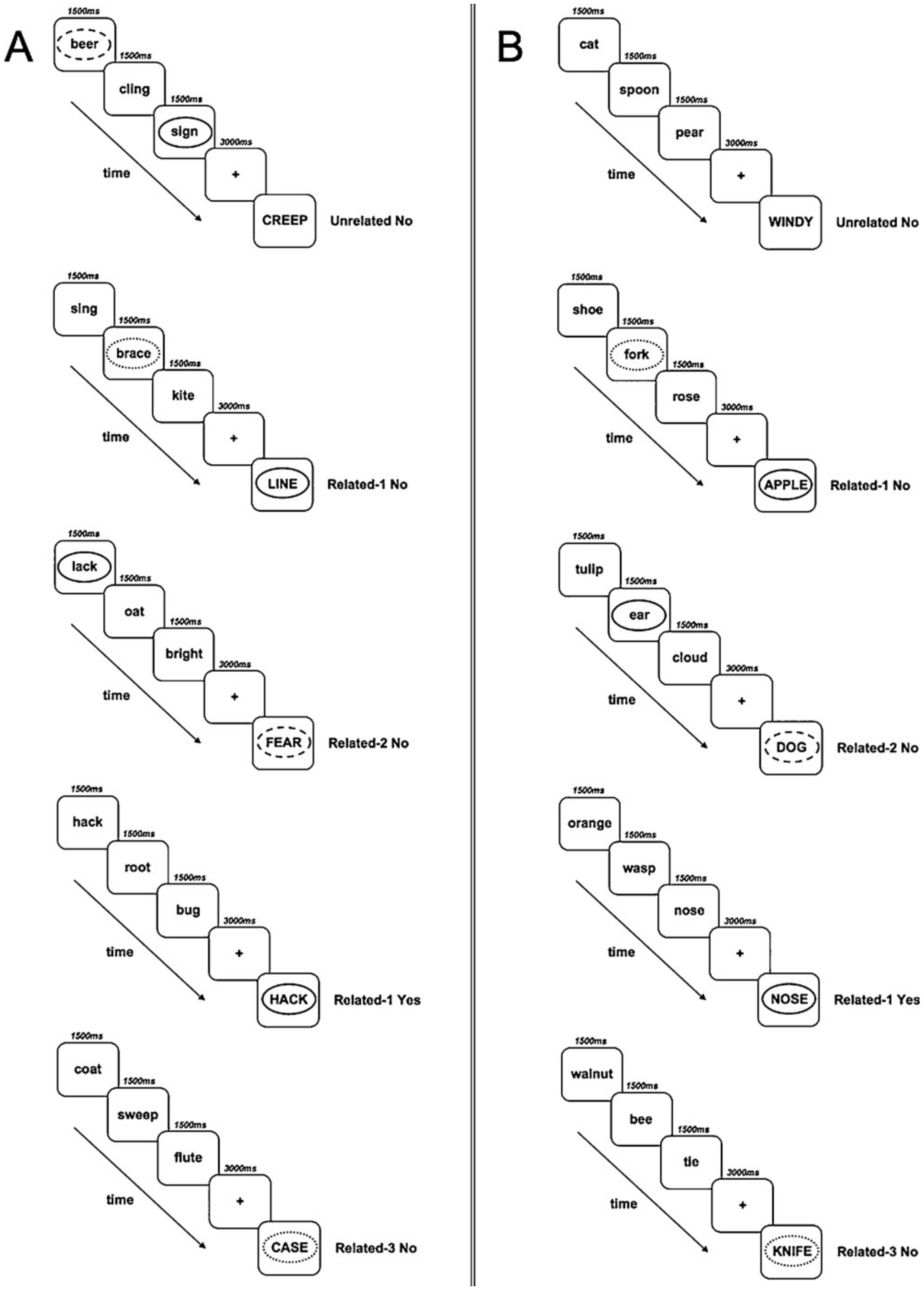


Figure 6.9: Examples of trials for each condition for Experiment 2. Panel A = phonological version; panel B = semantic version. Probe items are either phonologically (panel A) or semantically (panel B) related to previously viewed items (e.g. in panel A, the probe word case is related to 'brace', encountered 3 trials ago). Oval shapes denote relatedness of items. Figure taken from Barde et al. 2010.

Results

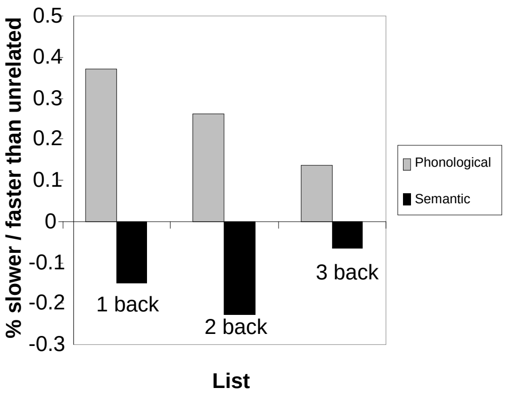
Table 6.2 shows the accuracy scores for each patient and for the control group. The experiment used a two-forced choice response and a score of above 60% in each condition was required before an analysis was conducted on the RT data. Patients DS and AK and controls performed near ceiling in most conditions, with the exception of DS's performance in the phonological Related-3 no condition. MM scored at chance level in all conditions and so his data were excluded from the analysis.

Table 6.2: Accuracy rates in each condition for patients and controls.

	Relatedness	**Unrelated no**	**Related-1 no**	**Related-2 no**	**Related-3 no**
DS	Phonological	98/108 (91%)	31/36 (86%)	29/36 (81%)	26/36 (72%)
	Semantic	104/108 (96%)	34/36 (94%)	33/36 (91.66)	34/36 (94%)
AK	Phonological	100/108 (93%)	34/36 (94%)	32/36 (88%)	33/36 (92%)
	Semantic	99/108 (92%)	36/36 (100%)	36/36 (100%)	34/36 (94%)
MM	Phonological	52/108 (48.1%)	22/36 (61%)	16/36 (44%)	13/36 (36%)
	Semantic	56/108 (52%)	13/36 (36%)	15/36 (42%)	17/36 (47%)
Controls	Phonological	511/520 (98%)	172/180 (96%)	180/180 (100%)	180/180 (100%)
	Semantic	516/520 (99%)	176/180 (98%)	170/180 (94%)	175/180 (97%)

The interference scores that appear in Figure 6.10 were calculated by subtracting mean RT for each related 'no' condition (Related-1, 2, or 3) from the mean RT for the unrelated 'no', and dividing the result by the unrelated 'no' mean RT, to normalise the data for differences in overall response speed ((related M – unrelated M) / unrelated M.). Scores on the present trials for the patients and controls were roughly equivalent to those scored on the absent trials (see Appendix H for accuracy on present trials).

Patient DS **Patient AK**



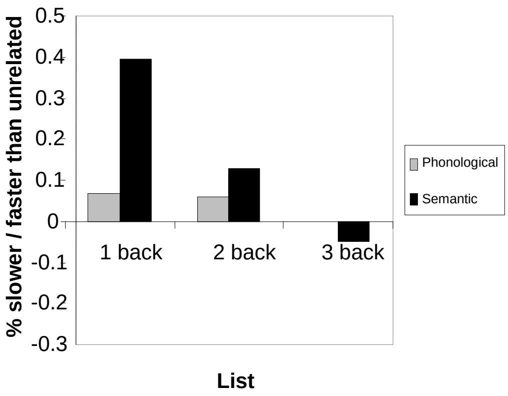


Figure 6.10: Interference scores on 'no' trials across phonological and semantic conditions for DS (sSTM), and AK (pSTM) respectively. Positive scores indicate interference (slowed scores in the related condition).

Only reaction times (RTs) from correct trials were considered. Any RTs that were more than 3 standard deviations away from each condition mean were removed. This led to the removal of 5 trials in DS's data set (3 from the Phonological, and 2 from the Semantic sub-experiment), and eight trials from AK's data set (4 from the Phonological and 4 from the Semantic sub-experiment). Normality tests conducted on these cleansed data were not significant (Shapiro-Wilk test p's > .7)

The normalised RT data were entered into a univariate ANOVA with the factors relatedness (Unrelated vs. Related-1, 2, or 3 back), patient (DS or AK), and type of relatedness/interference (semantic or phonological). The analysis returned a highly significant three-way interaction between patient, relatedness and type ($F(3, 560) = 6.959$, $p < .0001$). There were also significant two-way interactions between patient and type ($F(1, 560) = 29.046$, $p < .0001$), and relatedness and type ($F(3, 560) = 6.959$, $p = .003$).

One-way ANOVAs were performed on the data, with relatedness (Unrelated vs. Related-1, 2, or 3 back) entered as a factor. In the phonological interference data, there were significant differences between DS's (a sSTM patient) RTs in the unrelated and related trials where the related items appeared in the previous trial (Related-1, $F(1,98) = 11.665$, $p = .001$), and 2 trials back ($F(1,98) = 8.652$, $p = .004$). The difference approached significance for items 3 trials back ($F(1,98) = 3.121$, $p = .080$, Figure 6.10, left panel). For the experiment manipulating semantic relatedness, DS did not show interference effects. He actually scored significantly better in the Unrelated vs. Related-2 contrast (Related-1, $F(1,98) = 3.091$, $p > .05$; Related -2, $F(1,98) = 6.711$, $p = .011$, and Related-3, $F(1,98) = .393$, $p = .532$, Figure 6.10, left panel).

Patient AK (a pSTM patient) showed the opposite pattern of results to DS. No differences existed between the unrelated and related conditions in the phonological interference data (Related-1, $F(1,98) = 1.543$, $p = .217$, Related-2 $F(1,98) = 1.273$, $p = .262$, Related-3, $F(1,98) = .003$, $p = .960$, Figure 6.10, right panel). When manipulating semantic relatedness, AK showed a highly significant effect of interference for the related-1 contrast ($F(1,98) = 11.460$, $p = .001$), and an effect of relatedness that approached significance for related-2 trials ($F(1,98) = 3.537$, $p = .063$). No significant difference was found when comparing unrelated and Related-3 reaction times ($F(1,98) = .329$, $p = .329$, Figure 6.10, right panel).

Controls

The control group made very few errors in the experiment (Table 6.2). No data points were outside of 2 standard deviations away from the condition mean, and so no trials were removed from the control data. There was very little difference between the unrelated and related conditions for the control group in both experiments (interference scores ranged from -0.03 - 0.04 across semantic and phonological manipulations, and patterns were not systematic). No significant evidence for either semantic or phonological interference was found when RTs from the control group were entered into two-factor ANOVAs (all p's > .4, Figure 6.11).

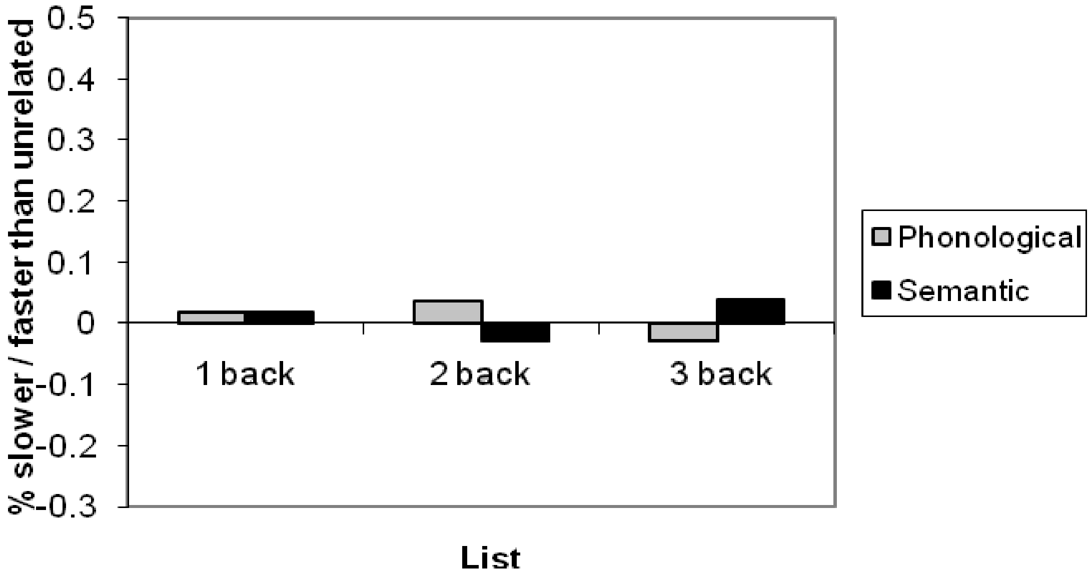


Figure 6.11: Interference scores on 'no' trials across phonological and semantic conditions for the control group. Positive scores indicate interference (slowed scores in the related condition).

Discussion

Both DS and AK showed reliable interference effects from related probes, while no interference effects were apparent in the controls. Task difficulty (in accuracy and RTs) was not matched across the patients and controls, so any conclusions should be cautious. Nevertheless, the results point to the STM patients being vulnerable to proactive interference. For DS there was reliable interference from phonological probes, which he was relatively slow to reject on no trials compared with the unrelated baseline. In contrast, there was no effect of semantic interference in DS's RT data. For AK, there was interference from semantic probes and no effect of interference from phonological probes. The double dissociation between the effects of phonological and semantic probes on DS and AK provides strong evidence for the two patients relying on different forms of STM representation – phonological in the case of DS and semantic in the case of AK. As a consequence of this, DS showed reliable effects of phonological interference while AK showed effects of semantic interference. In each case, the patients were relatively slow to reject a probe that was similar to a representation of an earlier item, held respectively in phonological and semantic STM (for DS and AK).

General Discussion

The current study found increased, and differential forms of interference effects in STM in patients with impaired phonological and semantic STM. The sSTM patient (DS) showed strong effects of phonological interference, but no semantic interference between items (indeed there was evidence for facilitated rejection of semantic distractors in Experiment 2). The opposite pattern of results was observed in the pSTM patient (AK). AK showed semantic but not phonological interference effects. DS also showed item-specific interference effects for real word stimuli and interference, comparing open and closed sets, of nonwords. The data suggest that patients with deficits in semantic STM show increased phonological interference when having to temporarily maintain items for letter recall / recognition. In contrast, patients with impaired phonological STM have increased semantic interference. The results are consistent with the existence of distinct phonological and semantic STMs, and with damage to one form of representation making patients vulnerable to interference based on the residual form of representation used to support STM performance.

The item recognition data in Experiment 1 are inconsistent with previous investigations of type-specific PI in semantic and phonological STM patients. The finding that the phonological STM patients did not show interference from repeating phonological representations (nonwords) suggests that the impairment in encoding phonological information in pSTM patients leads to a reduced likelihood of phonological information interfering with recall, which is at odds with the increased interference effect from phonological information in pSTM patients reported by Barde et al. (2006, 2010), and the finding of normal interference effects from both phonological and semantic information in Hamilton and Martin's pSTM patient (2002, 2007).

Similarly, the item recognition data when manipulating semantic and phonological relatedness (Experiment 2) were different from the findings reported in Barde et al. (2010), with the pSTM patient showing reduced interference from phonological relatedness and the sSTM patient demonstrating an absent interference effect from semantic relatedness. The divergent results produced between ours and Barde et al.'s (2010) study could be due to differences in the severity of the patients' deficits. The patients reported by Barde et al. (2010) were described as experiencing mild-to-moderate aphasia. In terms of the extent to their STM impairments, the pSTM patients achieved phonological similarity effects ranging from 0.4 - 2.0 (*M* = 1.2, *SD* = 0.5, smaller differences indicating weaker phonological similarity effects, and hence, impaired pSTM). The semantic STM-impaired group achieved lexicality effects ranging from 0.2 – 1.6 (*M* = 0.8, *SD* = 0.5, where smaller differences indicates weaker lexicality effects, or impaired sSTM). When calculated in the same way, the phonological similarity effects produced by the pSTM patients here were comparatively much weaker (AK = 0.01, MM = 0.1). The sSTM patient (DS) did show a substantial lexicality effect (DS = 0.29). However, performance on other semantic STM tests indicated that DS's deficit was in semantic STM, and indeed the deficit appeared to be stronger than those described by Barde et al. (2010). Therefore, a chronic deficit in, say, phonological STM could result in a failure to encode phonological information in the first place, producing a performance in STM based instead, perhaps, primarily on the semantic properties of the to-be-remembered words. This could explain the lack of interference effects evinced in the patients, along with the findings of exaggerated performance in the preserved information in STM. A weaker deficit in phonological STM may allow for some encoding of phonological information, but encoding is inefficient, leading to impoverished distinctions between phonological information in the current and previous lists.

Proactive interference from semantically related lists might be produced in sSTM patients because the 'probed' item in the stimuli list scores highly on lexical variables, eliciting a semantic encoding strategy. At retrieval, the probe item activates a high number of semantic neighbours. In impaired semantic retention, this combination leads to difficulty in selecting the appropriate competitor. This idea may also apply to Barde's finding of heightened phonological PI in pSTM cases. These items, chosen for the possibility of matching a probe phonologically, may be more phonologically frequent (affecting encoding), and contain more phonological neighbours (leading to problems at the retrieval stage).

Less easy to explain in terms of lexical and phonological biases in stimuli is the absence of a phonological interference effect in sSTM, and of semantic interference effect in pSTM. Impaired memory for phonological information may lead to a focus on semantic information in encoding that is more helpful than in controls. However, if an item encountered by a pSTM patient is lexically frequent, semantic information may be salient at encoding. The semantic probe may activate many semantic neighbours, but in this case, semantic STM is preserved, and so selection is easier. This account also applies to the absent phonological interference effect in semantic STM.

The fact that the present stimuli comprised words matched on lexical frequency, number of items within semantic category, and phonological neighbourhoods may have effectively removed biases at encoding. This in turn may reveal the direct effect of previously viewed phonologically or semantically-related neighbours of the current probe. This idea is consistent with the fact that the control group here did not show semantic or phonological interference effects.

When inspecting the interference scores provided in Barde et al. (2010), 16/20 patients showed a degree of interference for the preserved information type that is larger than those observed in their control group. Although the magnitude of these interference effects were not as large as those introduced by the impaired information type, the fact that some interference effect existed for the preserved type of information indicates some degree of persistence that may be larger than effects present in controls.

Whether patients show *abnormal* effects of interference is an important consideration for this topic. Here, as well as in Barde et al.'s (2010) and Hamilton and Martin's (2002, 2007) investigations, patients showed effects of PI of greater magnitude than those observed in age-matched controls. This consistency appears despite the suggestion that older participants are more susceptible to PI made elsewhere (May et al. 1999, Lustig et al. 2001). However, cross-condition ceiling performance in controls may diminish any existing interference effect. Experiment 1 in the current study reported accuracy rates in the control data that were not at ceiling.

These conditions perhaps provide the most meaningful comparison of interference effects in the patient data. The emphasis here was on relative effects of semantic and phonological interference, but future studies might seek to increase task difficulty in control experiments to investigate more closely the extent to which interference effects in patients are abnormal.

Hamilton and Martin (2002, 2005, 2007) report elevated effects of both phonological and semantic interference in their sSTM patient (ML), and no abnormal effects of interference in their patient with pSTM (EA). This finding is consistent with other findings that suggest that frontal damage leads to difficulties with semantic information, and interference effects (Shimamura et al. 1995, Thompson-Schill et al. 1997, Jonides et al. 1998, and Nelson et al. 2009). Their inhibition deficit hypothesis suggests that semantic buffer cases have an underlying failure to inhibit verbal information, rather than abnormal decay of semantic information. Hoffman et al. (2011) propose related explanations for the performance of semantic STM patients; couched in terms of impaired cognitive control mechanisms, rather than a deficient semantic STM buffer. However, Hoffman et al. (2011) argue that the patterns of performance demonstrated by their three semantic STM cases are unlikely to be due to inhibitory failure alone, as their patients did not make intrusions of items from previous trials. However, as Hoffman et al. (2011) only used tasks manipulating semantic information, they did not address whether phonological proactive interference was abnormally high in these patients.

Recent categorisations of semantic deficits might explain some of the inconsistencies in the nature of interference effects in sSTM patients. Jefferies and Lambon Ralph (2006) investigated the nature of semantic memory impairments in semantic dementia and aphasia. They found that where the performance of semantic dementia patients indicated degraded amodal semantic representations, the aphasic patients performance differed across (i) input/output modalities, (ii) with phonemic cueing, and (iii) as a function of how easily semantic associations between stimuli can be made, and distracters discarded. Therefore, patterns of interference in sSTM might be influenced by the strength of the semantic relationship between related test stimuli and a given probe, as well as by the type of semantic processing required to identify the relatedness between items. Also crucial to any interpretation of these data is whether patients in these studies present with additional executive problems in coordinating representations in different forms of STM. In terms of the patients from the current study, DS has frontal damage and some degree of general executive loss. AK and MM do not present with frontal damage and this could protect them from more general interference as in Hamilton and Martin (2002, 2007).

Hamilton and Martin (2007) found some important differences in types of errors made by the patients, reporting frequent intrusion errors in ML (sSTM), and very few in EA (pSTM). Both the current and Barde et al's (2010) analyses used RT as a dependent variable in the item-recognition task manipulating phonological and semantic relatedness. Accuracy rates in Barde et al.'s patient data are not reported. Although AK performed at ceiling, and MM performed at chance level in the current investigations, the pattern of accuracy produced by the sSTM patient DS indicates a greater deleterious effect of phonological than semantic relatedness. Future investigations may seek to increase task difficulty to produce below-ceiling performance, to investigate type-specific PI in semantic and phonological STM patients.

Conclusion

The current chapter sought to investigate types of PI in STM patients. The results support type-specificity in PI. Notably, phonological STM patients did not show interference from phonological information, but did from semantically related stimuli. In contrast, a semantic STM patient did not show semantic interference, but did show phonological interference. It may be that a type-specific deficit in STM leads to a reliance on the preserved type of information, manifested in abnormal PI for, for example, phonological information in the case of a semantic deficit. In addition, the STM patient showed improved rejection of semantic distractors that were from the same category as targets which could reflect semantic inhibition. Future studies may therefore investigate whether the magnitude of a PI effect was in the normal range in these patients, relative to controls under task-difficulty matched conditions.

Chapter 8: General Discussion

The present thesis described four rehabilitation studies of neuropsychological patients presenting with cognitive deficits, and two experimental investigations that examined the nature of patients' deficits in further detail. A key aim throughout was to investigate the use of the rehabilitation method as a theoretically informative tool. This section is separated into the two underlying themes of the work: 1. The nature of neighbourhood effects (*N* effects) in language and how they might be used to direct generalised improvement following rehabilitation (Chapters 2, 3, 4 and 5) and 2. Diagnosis and rehabilitation: testing associations and dissociations in the cognitive architecture (Chapters 6 and 7). The implications of the findings for rehabilitation and cognitive theory are discussed.

The nature of neighbourhood effects (N effects) in language and how they might be used to direct generalised improvement following rehabilitation

The number of items that can be feasibly treated directly during an intervention is limited. Therefore, it is important to produce optimal generalisation of training one set of items to untreated stimuli. A key distinction apparent in studies of the orthographic *N* effect in reading is that words with high numbers of neighbours benefit under conditions of parallel processing when lexical frequency is low (Andrews, 1989, 1992; Arguin, Bub, and Bowers, 1998; Sears, Hino, and Lupker, 1995), but are at a disadvantage when letters are processed in series (e.g. when words are presented sequentially in letter fragments: Snodgrass and Mintzer, 1993, or where the presentation of a letter distinguishing it from a neighbour is delayed by 100ms: Pugh, Rexer, Peter and Katz, 1994). There are other 'neighbourhoods' or 'families' of words besides orthographic neighbourhoods. For instance, in past tense verb morphology, some irregular past tense verbs can be grouped on the basis of shared phonological (rhyming) properties (e.g. sleep-slept, weep-wept), and some cognitive models of past tense verb processing assert that past-tense processing is facilitated by this neighbourly support from similar items.

Chapter 2 explored generalised improvement to untreated irregular verbs that take similar versus different phonological transformations to treated items. Generalised improvement was found to be mediated by the similarity of these transformations in a nonfluent aphasic patient; i.e. there was improvement that generalised to irregular past tense verbs whose transformations rhymed with treated items, but no generalised improvement was found to untreated irregular verbs with different phonological transformations. There were no error types post-therapy that suggested interference from treated neighbours, and no significant reduction in accuracy was observed on the untreated different neighbour set. The data suggested that irregular verb rules are coded in a generalisable form such that a gradual learning algorithm might rank constraints on associations between verb stems and their past tenses are based on probabilities, so that these free-variation constraints are kept or rejected based on how well they perform (e.g. Albright and Hayes, 2003). The results also demonstrated that this feature can be harnessed in a rehabilitation design in order to direct patterns of generalised improvement. Further, results from Chapter 3, which investigated the factors mediating past tense verb production in a patient with a semantic deficit, suggested that properties of words may operate in a compensatory manner, in that when support from word meaning knowledge / lexical frequency was low, phonological attributes (e.g. neighbourhood support) modulated irregular past tense processing. This pattern is analogous with findings from normal participants, where facilitatory *N* effects arise only with low frequency words (Andrews, 1989, 1992; Arguin, Bub, and Bowers, 1998; Sears, Hino, and Lupker, 1995)

There were findings reported in the current thesis that suggested that, under some conditions, learning treated words is detrimental to their (untreated) neighbourhoods. Chapter 4 reported a rehabilitation study with a type A graphemic buffer (GB) patient to assess the nature of the orthographic neighbourhood effect in spelling. Type A GB patients present with a bow shaped accuracy curve across letter positions (where medial sections are more erroneous). Following one rehabilitation study that showed a neighbourhood-mediated pattern of generalised improvement following rehabilitation of a GB deficit (Sage and Ellis, 2006), the aim was to test the nature of the *N* effect in a similar patient (JF). By contrasting performance on neighbours of treated items with shared (clock-block) versus changed (clock-cloak) middle sections, the current work suggested that generalised improvement was mediated by intact middle section neighbours rather than by orthographic neighbourhoods more generally.

In the normal literature on reading, neighbourhood effects vary across conditions of serial and parallel processing (e.g. Andrews, 1989, 1992; Arguin, Bub, and Bowers, 1998; Sears, Hino, and Lupker, 1995; Snodgrass and Mintzer, 1993 and Pugh, Rexer, Peter and Katz, 1994). E.g. Snodgras and Mintzer (1993) showed that when words are presented sequentially in letter fragments the *N* effect was deleterious in normal participants. Another way of assessing the relations between neighbourhood effects and serial/parallel processing is to assess neuropsychological patients who process letters in series. Letter-by-letter reading has been described as a perceptual deficit that manifests itself in an inability to process words in parallel (Geschwind, 1965a and b). However, recent work has indicated that these patients are able to process words in parallel when perceptual demands are low (e.g. with words comprised of low-confusability letters, Arguin 2005). Chapter 5 reported two letter-by-letter patients showing effects of letter confusability. At initial baseline, there were some key differences between the patients: DM was a 'pure' alexic case, who presented with accurate, but show reading, detrimentally affected by length. On the other hand, MAH was a patient experiencing expressive and receptive aphasia. His accuracy in reading was poor, but was better in some conditions (high frequency and imageability), and showed a reliable length effect. Reading times and accuracy scores for both patients showed an advantage for low relative to high confusability words. The patients exhibited interesting effects of neighbourhood size: low confusability words showed facilitative effects of high *N* whereas high confusability words showed a deleterious effect of high *N*.

These data are consistent with the idea that neighbourhood effects are facilitative when words are processed in parallel. However, when words are processed in a serial fashion, high *N* words produce greater competition. Presumably this competitive effect would be particularly detrimental when neighbours differ by a letter at the terminal letter positions. Reading performance appeared to be mediated to some extent by letter confusability in the letter-by-letter patients, therefore the amount of generalised improvement from a therapy approach designed to improve parallel processing with another promoting letter identification and discrimination. While improvement after parallel processing (word-level) therapy generalised only to low confusability words, letter-level treatment improved both low and high confusability words. This pattern of results further suggested that letter confusability mediates parallel processing in these patients, and that a therapy designed to overcome letter confusability may promote parallel processing. The two patients showed different post-therapy results: Effects of orthographic neighbourhood on high confusability words became facilitative in MAH's accuracy scores after letter, but not word-level therapy, and after word therapy in DM's RT data. These results could be due to the better pre-therapy reading abilities of DM compared to MAH, which may have led DM to benefit more from word therapy, resulting in an advantage for high neighbourhood words in both confusability sets in this case.

Section summary

As in the literature with normal participants, there was evidence for both facilitative and deleterious effects of neighbourhood in the neuropsychological and rehabilitation studies. Throughout, assertions have been made about how neighbourhood effects may behave across types of patients and in different cognitive tasks. Firstly, in the investigations of past tense verb production there were only facilitative effects of *N*, both in terms of *N* size at pre-therapy baseline (Chapters 2 and 3), and in the effect of generalisation to other neighbours post-therapy (Chapter 2). However, while no deleterious effects of neighbourhood were produced on real word items in these studies, there were some incorrect irregular-style past tense inflections on nonwords produced by a semantic patient in Chapter 3, and post-therapy in a nonfluent aphasic patient in Chapter 2. Therefore, the increased competition generated by treated words to other words is a major implication for neighbourhood-directed rehabilitation. As in studies of the *N* effect in normal participants, there was some evidence that interference from competing neighbours may be particularly detrimental when the task itself is serial (Chapter 4) or in reading when a patient's presentation includes the processing of letters in series (Chapter 5).

Taken together, this section of work suggested that generalised improvement to untreated items following the direct treatment of a given set may be guided via language neighbourhoods. There is an issue of potentially increasing neighbourhood competition following treatment, though some ways in which rehabilitation studies can offset this have been identified here. Verb 'neighbourhoods' in the English irregular past tense were defined in the investigations as verb pairings that shared phonological transformations (e.g. sleep-slept, weep-wept). To some extent the type of transformation each verb takes must be specified by the verb stem. Due to the limited number of English verbs, there are far fewer past-tense competitors to present tense verbs (i.e. verbs whose stems share phonological features with a given present-past verb pairing, but take a different transformation, e.g. beep-beeped is a regular competitor for sleep-slept). This may allow effective generalisation to large groups of irregulars taking similar transformations. Therefore, due to the limited nature of the English irregular past tense, issues pertaining to neighbourhood competition may not be as problematic in the context of rehabilitation of past tense verb production, relative to rehabilitation of spelling or reading of general words.

Chapter 4 described the neighbourhood-mediated generalisation effects following direct word treatment of a graphemic buffer deficit. The type A graphemic buffer patient described committed more errors at the middle letter positions relative to the beginning and end positions. Given this, and evidence that improvement generalised only to orthographic neighbours of treated items whose middle sections remained intact (i.e. took a change at the initial or final letter positions), it was concluded that selecting words with neighbourhoods that are i) large and ii) share consistent letter combinations at the problematic middle sections of words might be the most effective strategy for therapists treating type A GB patients. It was also speculated that Type B GB patients, who show a linear decline in accuracy across the word, might show greater generalised improvements to untreated neighbours whose terminal letter positions remain intact (take a letter change at the initial or medial sections).

Our investigation of two letter-by-letter patients (Chapter 5) indicated that orthographic neighbourhoods may exert different types of effects as a function of letter confusability (mediated by whether perceptual demands allow parallel processing). This finding was repeated in the rehabilitation data, where *N* effects on high confusability word sets became facilitative in both patients after the therapies. Therefore, *N*-mediated treatment may not be beneficial in cases of serial processing, but therapy approaches designed to either promote parallel processing (DM) or improve the efficacy of a serial strategy (MAH) may bring about facilitative effects of *N*, and the current investigations suggest that where patients show facilitative *N* effects at baseline, they might be used therapeutically in a rehabilitation setting, directing generalised improvements to untreated orthographic neighbours of treated items.

Diagnosis and rehabilitation: associations and dissociations in the cognitive architecture

As discussed in the General Introduction, one of the benefits of a rehabilitation approach in testing cognitive theories is the capacity to address the problem of specificity in cases with multiple deficits, exploring the effects of directly treating one type of ability / stimuli on other behaviours and stimuli sets. Therefore, rehabilitation studies can assess associations between cognitive behaviours and the nature of the associations between them, i.e. *how*, and under what conditions, does one cognitive skill recruit another. Chapters 6 and 7 investigated type-specific (phonological and semantic) deficits in STM. Chapter 6 tested whether STM and sentence comprehension abilities in two STM patients were linked, and the extent to which they fitted phonological and semantic STM profiles. Two rehabilitation procedures were devised, one aiming to improve phonological STM (using nonword recall) and another that aimed to promote semantic STM (real word recall). The two treatments were applied in an ABACA design (A = assessment, B = phonological STM treatment, C = semantic STM treatment), and aimed to selectively improve types of STM deficits in one pSTM and one sSTM patient (patients AK and DS, respectively). Phonological STM treatment remediated a profile of pSTM, and a semantic STM therapy improved a semantic STM deficit, bringing about, respectively, effects of phonological and semantic properties in STM that were absent pre-therapy. Overall, the first (pSTM) treatment led to gains in STM performance in both patients, though there were no significant differences in overall accuracy between post-treatment baseline performances in either patient.

In addition to improving STM, there was some generalisation to sentence comprehension, and this generalisation was mediated to some extent by the type of training and the diagnosed STM deficit. One patient described here (DS, sSTM) presented with agrammatic speech production. In this case, difficulties in sentence anomaly judgements were due to semantic STM, but that there was little effect of improving phonological or semantic STM on repetition accuracy, suggesting that floor performance in repetition was largely due to deficient language production. However, other aspects of DS's performance on STM tasks appeared due to problems in STM independent of his difficulties with language, based on the patterns of improvement in STM tasks observed over baseline phases. AK showed generalised improvement in sentence repetition following phonological STM training.

Overall, type-specificity in STM appears to be an important consideration for STM rehabilitation. The current data suggest that it might be advantageous to run a baseline assessing the extent of a phonological or semantic STM deficit, prior to treatment specifically designed to improve the deficient function, might be the most effective method for treating deficits in STM and sentence comprehension. It was also speculated that treating STM may lead to more generalisable improvements in sentence comprehension compared with a direct treatment of a set of sentences, as improvement from these techniques can often be restricted to the types of sentences treated, or to the treated items themselves (see Mitchum, Greenwald and Berndt, 2000 for a review).

A further factor assessed in the current study of STM was the ability to resist proactive interference (May, Hasher and Kane, 1999, Lustig, May and Hasher, 2001). Given the divergent findings in work on the nature of the PI effect in patients showing type-specific STM deficits, the aim was to test patterns of PI in three type-specific STM patients in Chapter 7. Two probe recognition tasks were employed, one using either words or nonwords that belonged to either open or closed sets (where the closed set comprised the interference condition). The other recognition task us real words, that manipulated interference from phonological (phonological relatedness between words) and semantic (semantic relatedness between words) PI. The results suggested type-specificity in PI: Firstly, the phonological STM patients showed interference from closed sets of real words (an advantage for open over closed sets), but not for nonwords, where there was no significant interference effect. In contrast, the semantic STM patient showed interference for both real words and nonwords. This pattern of results was attributed to a deficient encoding of phonological information in pSTM patients that prevents items with only phonological features (nonwords) from persisting in memory. Thus, just as pSTM patients fail to show effects of phonological similarity, so they may be insensitive to recurring items when these items possess phonological but not semantic properties. It was tentatively also suggested that interference in the real word condition demonstrated by the semantic STM patient may have been due to encoding using phonological over semantic properties of the words. Similarly, results from the probe task manipulating phonological and semantic similarity revealed that the phonological STM patient did not show interference from phonological information, and the semantic STM patient did not show semantic interference.

Taken together, the results are consistent with type specific proactive interference in the opposite direction to reports elsewhere (Barde et al., 2010). Here, a deficient sensitivity to phonological features of words (e.g. phonological similarity) at encoding appeared to produce an absent PI effect from phonological information in the pSTM patients, and this same pattern was observed with interference from semantic information in the sSTM patient. In terms of the contribution of PI to STM deficits in the neuropsychological patients, it was surmised that type-specific PI may not contribute to type-specific STM deficits, as both the PI effects for the deficient type of information in the patients, as well as the patterns of phonological and semantic effects in STM suggested degraded phonological / semantic representations, rather than abnormal persistence of this type of information. However, it may be that type-specific STM impairments lead to abnormal interference from the preserved type of information (e.g. enhanced semantic PI effects in pSTM patients) though this was not directly assessed in the PI and STM work.

Section summary

The work on types of STM deficits included in the current thesis has valid implications for cognitive theory: Firstly, the investigations reported in Chapter 6 offer support to previous studies that correlate absent effects of semantic variables on STM performance in sSTM patients, and a lack of phonological effects on STM in pSTM patients. This support is derived not only from the consistency in this pattern across the patients at pre-therapy baseline, but also from the appearance of semantic effects on STM following sSTM training in the sSTM patient, a finding that was mirrored in phonological STM effects for the pSTM patient after pSTM rehabilitation. Secondly, this pattern of type-specific consistency extended to sentence comprehension, adding weight to the notion that comprehension uses types of STM selectively, dependent on the comprehension condition. Thirdly, the studies into the nature of PI in these patients suggested that type-specific STM deficits are due to problematic encoding, rather than abnormal persistence, of the impaired type of information in STM. Finally, the rehabilitation method adopted in Chapter 6 allowed us to study the contributions of an STM deficit aside from language processing problems in an agrammatic patient, and showed that poor performance on these tasks was due, at least in part, to deficient STM.

Overall Conclusions

The present thesis investigated the efficacy of language neighbourhoods in directing generalised improvement after rehabilitation, and the use of the rehabilitation method in bringing about specificity of cognitive function in patients who present with an array of cognitive deficits, in order to test associations between behaviours.

Our work into language neighbourhoods and rehabilitation indicated that the effects of exploiting neighbourhood factors in the therapeutic setting was not always clear cut, due to increased competition for items from non-trained neighbourhoods after the treatment of a subset. However, the current thesis presents some factors that might offset this issue: 1. In sets of items where there are fewer competitors, *N*-directed rehabilitation may be largely facilitative (e.g. in the irregular past tense, Chapter 2), 2. In patients who show disproportionate accuracy rates across letter position, using items with large neighbourhoods, where letters are intact over the defective letter position is beneficial (Chapter 4), and *N* effects may be particularly detrimental in patients forced to process letters in series (e.g. letter by letter readers under conditions of high letter confusability, Chapter 5). This can be offset by reducing perceptual demands either experimentally (using low confusability words) or therapeutically (through reading treatments, Chapter 5).

In studies using rehabilitation and theoretically-motivated diagnosis to explore associations and dissociations in the cognitive architecture (Chapters 6 and 7) it was suggested that therapy approaches may be a useful tool in asserting direct cause and effect relations between cognitive functions in patients with a range of cognitive problems. Chapter 6 described an agrammatic patient who also showed difficulties in semantic STM. A treatment designed to specifically improve STM allowed us to make some distinctions about impaired behaviours due to impaired STM (those tasks where performance improved as a function of therapy, e.g. conceptual span), and those due to deficient language processing / production (tasks where performance showed negligible change following STM treatment, e.g. sentence repetition). This work was promising from a rehabilitation perspective, as improving the deficient component of STM appeared to be the most beneficial treatment strategy in the patients, and that type-specific generalised improvements were found in STM tasks and (to some degree) in sentence comprehension. These findings are also interesting from a theoretical standpoint, adding to the literature suggesting that the absence of phonological / semantic effects in STM are driven by deficient encoding of this type of information, observed in the patterns of improvement evinced across baseline phases. Findings also suggested a link between STM and sentence comprehension, and that this association may be mediated by type-specific STM deficits, and the type of STM employed in a given comprehension task.

References

Abel, S., Schultz, A., Radermacher, I., Willmes, K., & Huber, W. (2005). Decreasing and increasing cues in naming therapy for aphasia. Aphasiology, 19(9), 831–848.

Albright, A. and Hayes, B. (2001). Rules vs. analogy in English past tenses: A computational/ experimental study. Ms., Department of Linguistics, UCLA. Cited in Albright, A. and Hayes, B. (2002). *Modeling English Past Tense Intuitions with Minimal Generalization. Annual Meeting of the ACL archive Proceedings of the ACL-02 workshop on morphological and phonological learning - Volume 6*.

Andrews, S. (1989). Frequency and neighbourhood effects on lexical access: Activation or search? *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 15, 802–814.

Andrews, S. (1992). Frequency and neighbourhood effects on lexical access: Lexical similarity or orthographic redundancy? *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 18, 234–254.

Annoni, J. M., Lemay, A. M., de Mattos Pimenta, M. A. and Roch Lecours, A. (1998). The contribution of attentional mechanisms to an irregularity effect at the graphemic level. *Brain and Language*, 63, 64–78.

Arguin, M. and Bub, D. (1994). Pure Alexia: Attempted Rehabilitation and its implications for Interpretation of the Deficit. *Brain and Language*, 47, 233-268.

Arguin, M., and Bub, D. (1997). Lexical constraints on reading accuracy in neglect dyslexia. *Cognitive Neuropsychology,* 14, 765–800.

Arguin, M., and Bub, D. (2005). Parallel processing blocked by letter similarity in letter by letter dyslexia: A replication. *Cognitive Neuropsychology,* 22, 589-602.

Arguin, M., Bub, D. and Bowers, J. (1998). Extent and limits of covert lexical activation in letter-by-letter reading. *Cognitive Neuropsychology,* 15, 53–92.

Arguin, M., Fiset, S. and Bub, D. (2002). Sequential and parallel letter processing in letter-by-letter dyslexia. *Cognitive Neuropsychology*, 19, 535–555.

Baayen, R. H., Dijkstra, T. and Schreuder, R. (1997). Singulars and plurals in Dutch: Evidence for a parallel dual-route model. *Journal of Memory and Language*, 37, 94–117.

Baddeley, A. D. (1986). *Working memory*. New York: Oxford University Press.

Baddeley, A. D. and Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (8, 47-89). San Diego, CA: Academic Press.

Baddeley, A. D. and Wilson, B. A. (1994).When implicit learning fails: Amnesia and the problem of error elimination. *Neuropsychologia*, 32, 53–68.

Baddeley, A. D., Vallar, G. and Wilson, B. (1987). Sentence comprehension and phonological memory: Some neuropsychological evidence. In M. Coltheart (Ed.). *Attention and performance XII: The psychology of reading.* Hove, UK: Lawrence Elbaum Associates Ltd.

Barde, L. H. F., Schwartz, M. F. and Thompson-Schill, S. (2006). The role of left inferior frontal gyrus (LIFG) in semantic short-term memory: A comparison of two case studies. *Brain and Language*, 99, 8–219.

Barde, L. H. F., Schwartz, M. F., Chrysikou, E. G. and Thompson-Schill, S. L. (2010). Reduced short-term memory span in aphasia and susceptibility to interference: Contribution of material-specific maintenance deficits. *Neuropsychologia*, 48, 4, 909-920.

Baus, C., Costa, A. and Carreiras, M. (2008). Neighbourhood density and frequency effects in speech production: A case for interactivity. *Language and Cognitive Processes*, 23, 6, 866-888.

Beeson, P. M. (1998). *Treatment for letter-by-letter reading: A case study*, in: Clinical decision making in aphasia treatment. N. Helm-Estabrooks and A.L. Holland (Eds), Singlular Press, 153-177.

Behrmann, M., Nelson, J. and Sekuler, E. B. (1998). Visual complexity in letter-by-letter reading: Pure alexia is not pure. *Neuropsychologia*. 36 (11), 1115-1132.

Behrmann, M., Plaut, D. C. and Nelson, J. (1998). A literature review and new data supporting an interactive account of letter-by-letter reading. *Cognitive Neuropsychology,* 15*,* 7-51.

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

Bertram, R., Schreuder, R. and Baayen, R. H. (2000). The balance of storage and computation in morphological processing: The role of word formation type, affixal homonymy and productivity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 489–511.

Bird, H., Lambon Ralph, M. A., Seidenberg, M. S., McClelland, J. L. and Patterson, K. (2003). Deficits in phonology and past-tense morphology: What’s the connection? *Journal of Memory and Language,* 48, 502-526.

Bishop, D. (1989). *Test for the Reception of Grammar (2nd Ed).* University of Manchester: Age and Cognitive Performance Research Centre.

Bormann, T., Kulke, F. and Blanken, G. (2008). The influence of word frequency on semantic word substitutions in aphasic naming. *Aphasiology*, 22 (12), 1313-1320.

Bormann, T., Wallesch, C. W., Blanken, G. (2008). "Fragment errors" in deep dysgraphia: further support for a lexical hypothesis. *Cognitive Neuropsychology*, 25 (5), 745-64.

Bowers, J. S., Davis, C. J., and Hanley, D. A. (2005). Automatic semantic activation of embedded words: Is there a “hat” in “that”? *Journal of Memory and Language*, 52, 131–143.

Bub, D. and Kertesz, A. (1982). Deep agraphia. *Brain and Language*, 17, 146–165.

Bub, D., Black, S. E., and Howell, J. (1989). Word recognition and orthographic context effects in a letter-by-letter reader. *Brain and Language*, 36, 357-376.

Buchwald, A. and Rapp, B. (2004). Rethinking the Graphemic Buffer? *Brain and Language*, 91, 100-101.

Butterworth, B., Campbell, R., and Howard, D. (1986). The uses of short-term memory: A case study. *Quarterly Journal of Experimental Psychology*, 38A, 705-737.

Buxbaum, L. J. and Coslett, H. B. (1996). Deep Dyslexic Phenomena in a Letter-by-Letter Reader. *Brain and Language*, 54, 1, 136-167.

Bybee, J. L., and Slobin, D. I. (1982). Rules and schemas in the development and use of the English past tense. *Language*, 58, 265–289.

Bybee, J., and Moder, C. (1983). Morphological classes as natural categories. *Language*, 59, 251–270.

Byng, S. (1988). Sentence Processing Deficits: Theory and Therapy, *Cognitive Neuropsychology,* 5, 629–676.

Byng, S. and Black, M. (1999). *Reversible Sentence Comprehension Test*. Oxford: Winslow Press.

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

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

Caplan, D. and Hildebrandt, N. (1988). *Disorders of Syntactic Comprehension.* Cambridge, Mass: MIT Press.

Caplan, D. and Waters, G. (1999). Verbal working memory and sentence comprehension. *Behavioural and Brain Sciences*, 22, 77-94.

Caplan, D. and Waters, G. S. (1995). Aphasic disorders of syntactic comprehension and working memory capacity. *Cognitive Neuropsychology*, 12, 637-649.

Caramazza, A and Miceli, G. (1990). The structure of graphemic representations, *Cognition*, 37 (3), 243-297.

Caramazza, A. (1989). Cognitive neuropsychology and rehabilitation: An unfulfilled promise? In X. Seron and G. Deloche (Eds.). *Cognitive approaches in neuropsychological rehabilitation*. Hillsdale, NJ: Erlbaum, 383-398.

Caramazza, A., and Hillis, A. E. (1990). Levels of representation, coordinate frames, and unilateral neglect. *Cognitive Neuropsychology*, 7, 391-445.

Caramazza, A., Baili, A. G., Koller, J. J. and Berndt, R. S. (1981). An investigation of repetition and language processing in a case of conduction aphasia. *Brain and Language*, 14, 235-271.

Caramazza, A., Miceli, G. Villa, G., and Romani, C. (1987). The role of the graphemic buffer in spelling: Evidence from a case of acquired dysgraphia. *Cognition*, 26, 59-85.

Chertkow, H. and Bub D. (1990). Semantic memory loss in dementia of Alzheimer's type. *Brain*, 113, 397-417.

Chialant, D. and Caramazza, A. (1998). Perceptual and lexical factors in a case of letter-by-letter reading. *Cognitive Neuropsychology*, 15 (1/2), 167-201.

Chomsky, N. (1981). Principles and parameters in syntactic theory. In N. Hornstein and D. Lightfoot (Eds.). *Explanation in Linguistics* (123-146). London: Longman.

Cicerone, K. (2002) Remediation of 'working attention' in mild traumatic brain injury. *Brain Injury*, 16: 185-195.

Cipolotti, L., Bird, C., Glasspool, D. W. and Shallice, T. S. (2004). The impact of deep dysgraphia on graphemic output buffer disorders. *Neurocase,* 10 (6), 405-419.

Clare, L., Wilson, B. A., Breen, E. K., and Hodges, J. R. (1999). Errorless learning of face–name associations in early Alzheimer's disease. *Neurocase*, 5, 37-46.

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

Coltheart, M. and Byng, S. (1989). A treatment for surface dyslexia. In Seron, X. and Deloche, G. (Eds). *Cognitive approaches in neuropsychological rehabilitation*. (159-174). Hillsdale, N.J.: Lawrence Erlbaum Associates Ltd.

Coltheart, M., and Rastle, K. (1994). Serial processing in reading aloud: Evidence for dual-route models of reading. *Journal of Experimental Psychology: Human Perception and Performance*, 20, 1197-1211.

Coltheart, M., Bates, A. and Castles, A. (1994). Cognitive Neuropsychology and Rehabilitation. In Riddoch, M. J. and Humphreys, G. W. (Eds). *Cognitive Neuropsychology and Cognitive Rehabilitation*. Lawrence Erlbaum Associates Ltd. Hove, UK.

Coltheart, M., Curtis, B., Atkins, P. and Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review*, 100, 589-608.

Coltheart, M., Curtis, B., Atkins, P. and Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review*, 100, 589-608.

Coltheart, M., Rastle, K., Perry, C., Langdon, R. and Ziegler, J. (2001). The DRC model: A model of visual word recognition and reading aloud. *Psychological Review*, 108, 204-258.

Conroy, P., Sage, K. and Lambon Ralph, M.A. (2008). The effects of decreasing and increasing cue therapy on improving naming speed and accuracy for verbs and nouns in aphasia. *Aphasiology*, 23: 6, 707-730

Coslett, H. B. and Saffran, E. M. (1989) Preserved reading in pure alexia. *Brain,* 112**,** 327-359.

Coslett, H. B., Saffran, E. M., Greenbaum, S. and Schwartz, H. (1993). Reading in pure alexia: The effect of strategy. *Brain*, 116, 21-37.

Crisp, J. and Lambon Ralph, M.A. (2006). Unlocking the nature of the phonological-deep dyslexial continuum: the keys to reading aloud are in phonology and semantics. *Journal of Cognitive Neuroscience*, 18, 348-362.

Daneman, M. and Carpenter, P. A. (1980). Individual differences in working memory capacity-more evidence for a general capacity theory. *Memory*, 6, 122-149.

de Partz, M. P. (1995). Deficit of the graphemic buffer: Effects of a written lexical segmentation strategy. *Neuropsychological Rehabilitation*, 5, 129-147.

Dell, G. (1986). A spreading activation theory of retrieval in sentence production. *Psychological Review*, 93, 283-321.

Dell, G. S. and O'Seaghdha, P. G. (1992). Stages of lexical access in language production. *Cognition, 42*, 287-314. Reprinted (1993) in W. J. M. Levelt (Ed.), *Lexical access in speech production* (287-314). Cambridge, MA: Blackwell.

Deloche, G., Ferrand, I., Metz-Lutz, M - N., Dordain, M., Kremin, H., Mannequin, D., Perrier, D., Pichard, B; Quint, S., Larroque, C; Cardebat, D.; Evelyne Naud, E.; Bergego, C; Pradat-diehl, P; Tessier, C. (1992). Confrontation naming rehabilitation in aphasics: A computerised written technique. *Neuropsychological Rehabilitation*, 2, 117 – 124.

Desai, R., Conant, L. L., Waldron, E. and Binder, J. R. (2006). fMRI of Past Tense Processing: The Effects of Phonological Complexity and Task Difficulty. *Journal of Cognitive Neuroscience* 18, 2.

Domahs, F., Bartha, L. and Delazer, M. (2003) Rehabilitation of arithmetic abilities: Different intervention strategies for multiplication. *Brain and Language*, 87 165–166.

Ellis, A. and Young, A. (1988). *Human Cognitive Psychology*. Hove and London. Lawrence Erlbaum Associates Ltd.

Ellis, A. W. (1982). Spelling and writing (and reading and speaking). In A.W. Ellis (Ed.), *Normality and pathology in cognitive functions* (113-145). New York: Academic Press.

Ellis, A. W. and Franklin, S. (1994). Cognitive Neuropsychology and the Remediation of Disorders of Spoken Language. In Riddoch, M. J. and Humphreys, G. W. (Eds). *Cognitive Neuropsychology and Cognitive Rehabilitation*. Lawrence Erlbaum Associates Ltd. Hove, UK.

Ellis, A. W. and Young, A. W. (1988). *Human Cognitive Neuropsychology*. London: Lawrence Erlbaum Associates Ltd.

Engle, R. W., Carullo, J. J., and Collins, K. W. (1991). Individual differences in working memory for comprehension and following directions. *Journal of Educational Research*, 84, 253-262.

Evans, J. J., Wilson, B. A., Schuri, U., Andrade, J., Baddeley, A.D., Bruna,O., Canavan, T. , Della Sala, S., Green, R., Laaksonen, R., Lorenzi, L., and Taussik, I. (2000). A comparison of “Errorless” and “Trial-and error” learning methods for teaching individuals with acquired memory deficits. *Neuropsychological Rehabilitation*, 10, 67-101.

Faroqi-Shah, Y. (2007). Are regular and irregular verbs dissociated in non-fluent aphasia?: A meta-analysis. *Brain Research Bulletin*, 74, 1-13,

Fedorenko, E., Gibson, E. and Rohde, D. (2006). The nature of working memory capacity in sentence comprehension: Evidence against domain-specific working memory resources, *Journal of Memory and Language*, 54 (4), 541–553.

Fillingham, J. K., Hodgson, C., Sage, K., & Lambon Ralph, M. A. (2003). The application of errorless learning to aphasic disorders: A review of theory and practice. Neuropsychological Rehabilitation, 13(3), 337–363.

Fillingham, J. K., Sage, K., & Lambon Ralph, M. A. (2005a). The treatment of anomia using errorless vs. errrorful learning: Are frontal executive skills and feedback important? International Journal of Language and Communication Disorders, 40(4), 505–524.

Fillingham, J. K., Sage, K., & Lambon Ralph, M. A. (2005b). Further explorations and an overview of errorless and errorful therapy for anomia: The number of naming attempts during therapy effects outcome. Aphasiology, 19(7), 597–614.

Fillingham, J. K., Sage, K., & Lambon Ralph, M. A. (2006). The treatment of anomia using errorless learning. Neuropsychological Rehabilitation, 16, 129–154.

Fiset, D., Arguin, M., and McCabe, E. (2006). The breakdown of parallel processing in letter-by-letter dyslexia. *Cognitive Neuropsychology*, 23 (2), 240-260.

Fiset, D., Arguin, M., Bub, D., Humphreys, G. W. and Riddoch, M. J. (2005). How to make the word length effect disappear in letter-by-letter dyslexia: Implications for an account of the disorder. *Psychological Science*, 16, 535-541.

Forster, K. I. (1979). Levels of processing and the structure of the language processor. In Wales, R.J. and Walker, E.C. (Eds), *New approaches to language mechanisms*. Hillsdale, N.J.: Lawrence Erlbaum Associates Ltd.

Francis, D., Clark, N., Humphreys, G. (2003). The treatment of an auditory working memory deficit and the implications for sentence comprehension abilities in mild "receptive" aphasia. A*phasiology*, 17 (8), 723-750.

Friedmann, N. (2002). Question Production in Agrammatism: The Tree Pruning Hypothesis. *Brain and Language*, 80, 2, 160-187.

Friedmann, N. and Gvion, A. (2003). Sentence comprehension and working memory limitation in aphasia: A dissociation between semantic-syntactic and phonological reactivation. *Brain and Language*, 86, 23-39.

Funnell, E. (1987). Morphological errors in acquired dyslexia: A case of mistaken identity. *Quarterly Journal of Experimental Psychology*, 39A, 497-538.

Garrett, M. (1975). The analysis of sentence production. In G. Bower (Ed.). *The psychology of learning and motivation*. (133-177). New York: Academic Press.

Garrett, M. (1980). Levels of processing in sentence production. In Butterworth, B. (Ed.). *Language Production* (Volume 1), 177-220. New York: Academic Press.

Geschwind N. (1965a). Disconnexion syndromes in animals and man. I. *Brain*; 88, 237-94.

Geschwind N. (1965b) Disconnexion syndromes in animals and man. II. *Brain*; 88, 585-644.

Gibson, E. (2000). *Dependency locality theory: A distance-based theory of linguistic complexity*. Image, language, brain: Papers from the first Mind Articulation Project Symposium, Marantz, A., Miyashita, Y., and O'Neil, W. (Eds), 95-126. Cambridge, MA: MIT Press.

Gilmore, G. C., Hersh, H., Caramazza, A., and Griffin, J. (1979). Multidimensional letter similarity derived from recognition errors. *Perception and Psychophysics*, 25, 425-431.

Girelli, L and Seron, X. (2001) Rehabilitation of number processing and calculation skills. *Aphasiology*, 15 (7), 695-712.

Glasspool, D. W. and Houghton, G. (2005) Serial order and consonant-vowel structure in a graphemic output buffer model. *Brain and Language,* 94 (3), 304-330.

Glasspool, D. W., Shallice, T. and Cipolotti, L. (1999). Neuropsychologically plausible sequence generation in a multi-layer network model of spelling. In D. Heinke, G. W. Humphreys and A. Olson (Eds.) *Connectionist Models in Cognitive Neuroscience*. London: Springer-Verlag.

Glasspool, D. W., Shallice, T. and Cipolotti, L. (2006). Towards a unified process model for graphemic buffer disorder and deep dysgraphia, *Cognitive Neuropsychology*, 23: 3, 479 - 512.

Glisky, E. L. (1995). Acquisition and transfer of word processing skill by an amnesic patient. *Neuropsychological Rehabilitation*, 5, 299–318.

Goldblum, M. C., Gomez, C. M., Dalla Barba, G., Boller, F., Deweer, B., Hahn, V and Dubois, B. (1998). The influence of semantic and perceptual encoding on recognition memory in Alzheimer's disease. *Neuropsychologia*, 36, 8, 717-7

Gordon, P . C., Hendrick, R., and Levine, W. H. (2002). Memory-load interference in syntactic processing. *Psychological Science*, 13, 425-430.

Grodzinsky, Y. (1986). Language deficits and the theory of syntax. *Brain and Language*, 27, 135-159.

Grodzinsky, Y. (1995). Trace deletion, *theta*-roles, and cognitive strategies. *Brain and Language*, 51, 467-497.

Hamilton, A. C. and Martin, R. C. (2002). Inhibition and proactive interference effects in impaired semantic short-term memory of aphasic patients. *Brain and Language*, 83, 9-224.

Hamilton, A. C. and Martin, R. C. (2007). Proactive interference in a semantic short-term memory deficit: Role of semantic and phonological relatedness. *Cortex*, 43, 112-123.

Hanten, G. and Martin, R. (2000). Contributions of phonological and semantic short-term memory to sentence processing: Evidence from two cases of closed head injury in children. *Journal of Memory and Language*, 43, 335-361

Haslam, C., Kay, J., Tree, J. J. and Baron, R. (2009). Dysgraphia in dementia: A systematic investigation of graphemic buffer features in a case series, *Neurocase*, 15 (4), 338-351

Hillis, A. and Caramazza, A. (1989). The graphemic buffer and attentional mechanisms. *Brain and Language*, 36, 208-235.

Hitch, G. J., and Baddeley, A. D. (1976). Verbal reasoning and working memory. *Quarterly Journal of Experimental Psychology*, 28, 603-621.

Hoffman, P., Jefferies, E., and Lambon Ralph, M.A. (2011). Explaining semantic short-term memory deficits: Evidence for the critical role of semantic control. *Neuropsychologia*, 49, 368-381.

Holland, R. and Lambon Ralph, M.A. (2010). The anterior temporal semantic hub is a part of the language neural network: Selective disruption of irregular past tense verbs by rTMS. *Cerebral Cortex*, 20, 2771-2775.

Houghton, G. (1990). The problem of serial order: A neural network model of sequence learning and recall. In R. Dale, C. Mellish, and M. Zock (Eds.), *Current research in natural language generation* (287–319). London: Academic Press.

Houghton, G., Glasspool, D. and Shallice, T. (1994). Spelling and serial recall: Insights from a competitive queuing model. In G. D. A. Brown and N. C. Ellis (Eds.), Handbook of spelling: Theory, process, and intervention. Chichester, UK: John Wiley and Sons.

Howard, D. and Patterson, K. (1989). Models for therapy. In Seron, X. and Deloche, G. (Eds.), *Cognitive approaches in neuropsychological rehabilitation*. Hillsdale, NJ: Erlbaum.

Howard, D. and Patterson, K. (1992). *The Pyramids and Palm Trees Test: A test of semantic access from words and pictures*. Thames Valley Test Company, Bury St. Edmunds, Suffolk, England.

Humphreys, G. W., Evett, L. J. and Quinlan, P. T. (1990). Early orthographic processing in visual word recognition. *Cognitive Psychology*, 22. 517-560.

Ischebeck, A., Zamarian, L., Siedentopf, C., Koppelstaetter, F., Benke, T., Felber, S. and Delazer, M. (2006). How specifically do we learn? Imaging the learning of multiplication and subtraction. *Neuroimage*, 30, 1365-1375.

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

Jefferies E., Hoffman P., Jones R. and Lambon Ralph M. A. (2008). The impact of semantic impairment on verbal short-term memory in stroke aphasia and semantic dementia: A comparative study. *Journal of Memory and Language*. 58 (1), 66-87.

Jefferies, E. and Lambon Ralph, M. A. (2006). Semantic impairment in stroke aphasia versus semantic dementia: a case-series comparison. *Brain*, 129, 2132-2147.

Jensen, A. R. (1962) Spelling errors and the serial position effect. *Journal of Educational Psychology,* 53, 106-109.

Jensen, L. R. (2000). Canonical structure without access to verbs? *Aphasiology*, 14, 827-850.

Jescheniak, J., and Levelt, W. J. M. (1994). Word frequency effects in production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 824-843.

Joanisse, M. and Seidenberg, M. S. (1999). Impairments in verb morphology after brain injury: A connectionist model. *Proceedings of the National Academy of Sciences (USA)* 96, 7592-7597.

Jonides, J., Smith, E. E., Marshuetz, C., Koeppe, R. A., and Reuter-Lorenz, P. A. (1998). Inhibition in verbal working memory revealed by brain activation. *Proceedings of National Academy of Sciences*, USA, 95, 8410-8413.

Jonsdottir, M. K., Shallice, T. and Wise, R. (1996). Phonological mediation and the graphemic buffer disorder in spelling: Cross language differences? *Cognition*, 59, 169–197.

Just, M. A., and Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, 98, 122–149.

Katz, R. B. (1991). Limited retention of information in the graphemic buffer. *Cortex* 27, 11. 1-119.

Kay, J., and Hanley, R. (1991). Simultaneous form perception and serial letter recognition in a case of letter-by-letter reading. *Cognitive Neuropsychology*, 8, 249-273.

Kay, J., Lesser, R and Coltheart, M. (1992). *PALPA: Psycholinguistic Assessments of Language Processing in Aphasia*. Lawrence Erlbaum Associates, Hove, England.

Kinsbourne M., Warrington E. K. (1962). A disorder of simultaneous form perception. *Brain*, 85, 461–486.

Kohn, S. E., Smith, K. L. and Arsenault, J. K. (1990). The remediation of conduction aphasia via sentence repetition: a case study. *British Journal of Disorders of Communication*, 24, 45-60.

Kurowski, K., Hazen, E. and Blumstein, S. E. (2003). The nature of speech production impairments in anterior aphasics: An acoustic analysis of voicing in fricative consonants. *Brain and Language*, 84, 353–371.

Lambon Ralph, M. A., Braber, N., McClelland, J. L. and Patterson, K. (2005). What underlies the neuropsychological pattern of irregular > regular past-tense verb production? *Brain and Language*, 93, 109-119.

Landis, T., Graves, R. and Goodglass, H. (1982). Aphasic reading and writing: possible evidence for right hemisphere participation. *Cortex*, 18, 105-112.

Lavidor, M., and Ellis, A. W. (2002). Word length and orthographic neighbourhood size effects in the left and right cerebral hemispheres. *Brain and Language*, 80, 45-62.

Lecours, A. R. and Lhermitte, F. (1969). Phonemic paraphasias: Linguistic structures and tentative hypotheses. *Cortex*, 5, 193-228.

Levine, D. N. and Calvano, R. (1978) A study of the visual defect in verbal alexia simultanagnosia. *Brain*, 101, 65–81.

Longworth, C. E., Tyler, L. K. and Marslen-Wilson, W. D. (2003). Language deficits and basal ganglia lesions: The past tense. *Brain and Language*, 87, 7-8.

Lustig, C., May, C. and Hasher, L. (2001). Working Memory Span and the Role of Proactive Interference. *Journal of Experimental Psychology*, 130 (2), 199-207.

Marcus, G. F., Pinker, S., Ullman, M., Hollander, M., Rosen, T. J. and Xu, F. (1992). Over regularization in language acquisition. *Monographs of the Society for Research in Child Development,* 57, 228.

Marshall, J., Chiat, S. and Pring, T. (1997). An impairment in processing verbs' thematic roles: A therapy study. *Aphasiology,* 11, 855-876.

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word recognition, *Cognition*, 25, 71-102.

Marslen-Wilson, W. D. and Tyler, L. K. (2007). Morphology, language and the brain: The decompositional substrate for language comprehension. *Philosophical Transactions of the Royal Society*, 362, 823-836.

Marslen-Wilson, W. D., Tyler, L. K. (1997). Dissociating types of mental computation, *Nature*, 387, 592–594.

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

Martin, R. and Romani, C. (1994). Verbal working memory and sentence comprehension: A multiple components view. *Neuropsychology* 8, 506–523.

Martin, R. C. (1987). Articulatory and phonological deficits in short-term memory and their relation to syntactic processing. *Brain and Language*, 32, 159-192.

Martin, R. C. (1990a). Neuropsychological evidence on the role of short-term memory in sentence processing. In G. Vallar and T. Shallice (Eds.). *Neuropsychological impairments of short-term memory* (390-427). New York: Cambridge University Press.

Martin, R. C. (1990b). The consequences of reduced memory span for the comprehension of semantic versus syntactic information. *Brain and Language*, 38, 1-20.

Martin, R. C. (2006). The neuropsychology of sentence processing: Where do we stand? *Cognitive Neuropsychology*, 23, 74-95.

Martin, R. C. and Freedman, M. L. (2001). Short-term retention of lexical-semantic representations: Implications for speech production. *Memory*, 9 (4), 261-280.

Martin, R. C. and He, T. (2004). Semantic short-term memory and its role in sentence processing: A replication. *Brain and Language*, 89, 1, 76-82.

Martin, R. C. and Lesch, M. Associations and dissociations between language impairment and list recall: Implications for models of STM. In Gathercole, S.E. (1996) *Models of Short Term Memory* Psychology Press, Hove.

Martin, R. C. and Romani, C. (1994). Verbal working memory and sentence comprehension: A multiple components view. *Neuropsychology,* 8, 506–523.

Martin, R. C., Shelton, J. R., and Yaffee, L. S., (1994). Language processing and working memory: Neuropsychological evidence for separate phonological and semantic capacities. *Journal of Memory and Language,* 33, 83–111.

May, C. P., Hasher, L. and Kane, M. J. (1999). The role of interference in memory span. *Memory and Cognition*, 27, 759-767.

McCarthy, R. A., Warrington, E. K. (1987). Understanding: A function of short-term memory? *Brain*, 110, 1565-1578.

McClelland, J. L. and Patterson, K. (2002a). 'Words *Or* Rules' cannot exploit the regularity in exceptions: Reply to Pinker and Ullman. *Trends in Cognitive Science*, 6, 464-465.

McClelland, J. L. and Patterson, K. (2002b). Rules or connections in past-tense inflections: What does the evidence rule out? *Trends in Cognitive Sciences,* 6, 465-472.

McClelland, J. L., Rumelhart, D. E., and Hinton, G. E. (1986). *The Appeal of Parallel Distributed Processing*. Vol I, Ch 1.

McNeil, M.R. and Prescott, T.E. (1978). *The Revised Token Test*. Baltimore, M.D: University Park Press.

Menn, L. (1999). Discussion at symposium: 'Filler Syllables from Developmental and Functional Perspectives'. *8th International Congress for the Study of Child Language*. San Sebastian, Spain.

Miceli, G, Capasso, R, Ivella, A and Caramazza, A. (1997). Acquired dysgraphia in alphabetic and stenographic handwriting. *Cortex*, 33, 355–67.

Mitchum, C. C. and Berndt, R. S. (1994). Verb retrieval and sentence construction: Effects of targeted intervention. In G. W. Humphreys and M. J. Riddoch (Eds.). *Cognitive Neuropsychology and Cognitive Rehabilitation* (317-348). Hove, UK: Lawrence Erlbaum Associates.

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

Moody, S. (1988). The Moyer reading technique re-evaluated. *Cortex*, 24, 474-476.

Moyer, S.B. (1979). Rehabilitation of acquired alexia: a case study. *Cortex,* 15, 139-144.

Mozer, M. (1987). Early parallel processing in reading: A connectionist approach. In M. Coltheart (Ed.), *Attention and performance, Vol. XII: The psychology of reading* (83-104). Hove, UK: Lawrence Erlbaum Associates Ltd.

Nelson, J. K., Reuter-Lorenz, P. A., Persson, J., Sylvester, C - Y. C. and Jonides, J. (2009). Mapping interference resolution across task domains: A shared control process in left inferior frontal gyrus. *Brain Research*, 1256, 92-100

Osswald, K., Humphreys, G. W. and Olson, A. (2002). Words are more than the sum of their parts: Evidence for detrimental effects of word-level information in alexia, *Cognitive Neuropsychology*, 19, 8, 675-695.

Page, M., Wilson, B. A., Norris, D., Shiel, A. and Carter, G. (2001). Familiarity and recollection in errorless learning. *Journal of the International Neuropsychological Society*, 7, 413 (abstract).

Patterson, K. (1994) Reading, writing and rehabilitation: A reckoning. In M. J. Riddoch and G. W. Humphreys (Eds.), *Cognitive Neuropsychology and Cognitive Rehabilitation* (425-447). Hove: Lawrence Erlbaum Associates.

Patterson, K. and Kay, J. (1982) Letter-by-letter reading: Psychological descriptions of a neurological syndrome. *Quarterly Journal of Experimental Psychology*, 34a, 411-441.

Patterson, K. and Shewell, C. (1987). Speak and spell: Dissociations and word-class effects. In Coltheart, M. Sartori, G. and Job, R. (Eds). *The cognitive neuropsychology of language*. (273-292). London: Lawrence Erlbaum Associates Ltd.

Patterson, K., Lambon Ralph, M. A., Hodges, J. R. and McClelland J. L. (2001). Deficits in irregular past-tense verb morphology associated with degraded semantic knowledge *Neuropsychologia*, 39, 709–724.

Penke, M. and Westermann, G. (2006). Broca's Area and Inflectional Morphology: Evidence from Broca's Aphasia and Computer Modeling. *Cortex*, 42, 4, 563-576.

Pinker, S. (1999). *Words and rules*. New York, NY: Basic Books.

Pinker, S. and Prince, A. (1988). On language and connectionism: analysis of a parallel distributed processing model of language acquisition. *Cognition*, 28, 73-194.

Pinker, S. and Ullman, M. (2002). The past and future of the past tense. *Trends in Cognitive Sciences*, 6, 456-462.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S. and Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review,* 103*,* 56-115.

Postararo, L., Zinelli, P., and Mazzuchi, A. (1998). Selective impairment of the graphemic buffer in acquired dysgraphia: A case study. *Brain and Language*, 35, 274-286.

Postle, B. R. and Brush, L.N. (2004). The neural bases of the effects of item-nonspecific proactive interference in working memory. *Cognitive, Affective and Behavioral Neuroscience*, 4, 379-392.

Postle, B. R., Berger, J. S., Goldstein, J. H., Curtis, C. E., and D'Esposito, M. (2001). Behavioural and neurophysiological correlates of episodic coding, proactive interference, and list length effects in a running span verbal working memory task. *Cognitive, Affective, and Behavioural Neuroscience*, 1, 10-21.

Prasada, S. and Pinker, S. (1993). Generalization of regular and irregular morphological patterns. *Language and Cognitive Processes* 8, 1–56.

Price, C. J. and Humphreys, G. W. (1992). Letter-by-letter reading? Functional deficits and compensatory strategies. *Cognitive Neuropsychology*, 9, 427–457.

Pugh, K., Rexer, K., Peter, M. and Katz, L. (1994). Neighborhood effects in visual word recognition: Effects of letter delay and nonword context difficulty. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 639-648.

Rapcsak, S. (1997). Disorders of writing. In L. J. G. Rothi and K. M. Heilman (Eds), Apraxia: The neuropsychology of action (149-172). Hove, UK: Psychology Press.

Rapp, B. and Kong, D. (2002). Revealing the component functions of the graphemic buffer. *Brain and Language*, 83, 112-114.

Rastle, K., Harrington, J. and Coltheart, M. (2002). 358,534 nonwords: The ARC Nonword Database. Quarterly Journal of Experimental Psychology, 55A, 1339-1362.

Raymer, A., Cudworth, C. and Haley, M. (2003). Spelling treatment for an individual with dysgraphia: Analysis of generalisation to untrained words. *Aphasiology*, 17 (6-7), 607-624.

Reuter-Lorenz, P. A., and Brunn, J. L. (1990). A prelexical basis for letter-by-letter reading: A case study. *Cognitive Neuropsychology*, 7, l-20.

Rochon, E. and Reichman, S. (2004). A modular treatment for sentence processing impairments: Sentence comprehension. *Journal of Speech-Language Pathology and Audiology*, 28, 25-33.

Rochon, E., Laird, L., Bose, A. and Scofield, J. (2005). Mapping therapy for sentence production impairments in nonfluent aphasia. *Neuropsychological Rehabilitation*, 15, 1-36.

Rosazza, C., Appollonio, I., Isella, V. and Shallice, T. (2007). Qualitatively different forms of pure alexia. *Cognitive Neuropsychology*, 24, 393–418.

Saffran, E. M. and Martin, O. S. M. (1975). Immediate memory for word lists and sentences in a patient with deficient auditory working memory. *Brain and Language*, 2, 420-433.

Sage K. and Ellis. A. (2004). Lexical influences in graphemic buffer disorder. *Cognitive Neuropsychology*, 21, 2-4, 381-400.

Sage, K and Ellis, A. W. (2006). Using orthographic neighbours to treat a case of graphemic buffer disorder *Aphasiology*, 20, 851-870.

Sage, K., Hesketh, A. and Ralph, M. A. Lambon (2005). Using errorless learning to treat letter-by-letter reading: Contrasting word versus letter-based therapy. *Neuropsychological Rehabilitation*, 15, 5, 619-642.

Schiller, N. O., Greenhall, J. A., Shelton, J. R., and Caramazza, A. (2001). Serial order effects in spelling errors: Evidence from two dysgraphic patients. *Neurocase*, 7, 1-14.

Schreuder, R., and Baayen, R. H. (1995). Modeling morphological processing. In F. LaurieBeth (Ed.), *Morphological aspects of language processing* (131-154). Hillsdale, NJ, England: Lawrence Erlbaum Associates.

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

Sears, C. R., Hino, Y., and Lupker, S. J. (1995). Neighborhood size and neighborhood frequency effects in word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 876–900.

Seidenberg, M. S. and McClelland, J. L (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, 96, 523-568.

Seidenberg, M.S. (1988). Cognitive neuropsychology and language: The state of the art. *Cognitive Neuropsychology*, 5, 403-426.

Seidenberg, M.S. and Arnoldussen, A. (2003). The brain makes a distinction between hard and easy stimuli: Comments on Beretta et al. *Brain and Language* 85, 527–530.

Shallice, T. and Saffran, E. (1986). Lexical processing in the absence of explicit word identification: Evidence from a letter-by-letter reader. *Cognitive Neuropsychology*, 3, 429-458.

Shallice, T., Glasspool, D. W. and Houghton, G. (1995). Can neuropsychological evidence inform connectionist modelling? Analyses of spelling. *Language and Cognitive Processes*, 10, 3 and 4, 195-225.

Shimamura, A. P., Jurica, P. J., Mangels, J. A. and Knight, R. T. (1995). Susceptibility to Memory Interference Effects following Frontal Lobe Damage: Findings from Tests of Paired-Associate Learning. *Journal of Cognitive Neuroscience*, 7, 2, 144-152

Snodgrass, J. G., and Mintzer, M. (1993). Neighborhood effects in visual word recognition: Facilitatory or inhibitory? *Memory and Cognition,* 21**,** 247–266.

Squires, E. J., Hunkin, N. M., and Parkin, A. J. (1996). Memory notebook training in a case of severe amnesia: Generalising from paired associate learning to real life. *Neuropsychological Rehabilitation,* 6, 55–65.

Stemberger, J. P. (1993). Vowel dominance in overregularizations. *Journal of Child Language*, 20, 503-521.

Stemberger, J. P. (2004a). Phonological priming and irregular past. *Journal of Memory and Language*, 50, 82–95.

Stemberger, J. P. (2004b). Neighbourhood effects on error rates in speech production. *Brain and Language*, 90, 413–422.

Stemberger, J. P., and MacWhinney, B. (1986). Form-oriented errors in inflectional processing. *Cognitive Psychology*, 18, 329–354.

Swinburn, K., Porter, G. and Howard, D. (2004). *Comprehensive Aphasia Test*. Psychology Press; UK

Tainturier, J. and Caramazza, A. (1996). The status of double letters in graphemic representations, *Journal of Memory and Language*, 35, 53–73.

Thompson, C. K., Shapiro, L. P., Tait, M. E., Jacobs, B. J. and Shneider, S. L. (1996). Training Wh-Question production in agrammatic aphasia: Analysis of argument and adjunct movement. *Brain and Language,* 52, 175–228.

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K. and Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A re-evaluation. Proceedings of National Academy of Sciences, USA. *Neurobiology*, 94, 14792 - 14797.

Tuomainen, J. and Laine, M. (1991). Multiple oral rereading technique in rehabilitation of pure alexia. *Aphasiology* 5, 401–409.

Turner, M. L. and Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, 28, 127-154.

Turner, M. S., Cipolotti, L.,Yousry, T. and Shallice, T. (2007). Qualitatively different memory impairments across frontal lobe subgroups. *Neuropsychologia*, 45, 1540-1552.

Tyler, L. K., Randall, B. and Marslen-Wilson, W. D. (2002). Phonology and neuropsychology of the English past tense. *Neuropsychologia,* 40, 1154-1166.

Tyler, L. K., Stamatakis, E. and Marslen-Wilson, W. D. (2005a). Differentiating the neural language system: Contrasting the regular and irregular past tense. *Brain and Language,* 95, 172-173.

Tyler, L. K., Stamatakis, E., Post, B., Randall, B. and Marslen-Wilson, W. D. (2005b). Temporal and frontal systems in speech comprehension: An fMRI study of past tense processing. *Neuropsychologia,* 43, 1963-1974.

Ullman, M., Corkin, S., Coppola, M., Hickok, G., Growdon, J. H., Koroshetz, W. J. and Pinker, S. (1997). A neural dissociation within language: Evidence that the mental dictionary is part of declarative memory, and that grammatical rules are processed by the procedural system. *Journal of Cognitive Neuroscience,* 9, 266-276.

Vallar, G. and Baddeley, A.D. (1984a). Fractionation of working memory. Neuropsychological evidence for a phonological short-term store. *Journal of Verbal Learning and Verbal Behaviour*, 23, 151-162.

Vallar, G. and Baddeley, A.D. (1984b). Phonological short-term store, phonological processing and sentence comprehension: A neuropsychological case study. *Cognitive Neuropsychology*, I, 121-141.

Vallar, G. and Baddeley, A.D. (1987). Phonological short-term store and sentence processing. *Cognitive Neuropsychology*, 4, 417-438.

Vallat, C., Azouvi, P., Hardisson, H., Meffert, R., Tessier, C., and Pradat-Diehl, P. (2005) Rehabilitation of verbal working memory after left hemisphere stroke. *Brain Injury*, 19 (13), 1157-1164.

Ward, J. and Romani, C. (1998). Lexical activation and serial position effects in spelling: Evidence from a single case study. *Neurocase*, 4, 189-206.

Warrington, E. K. (1984). *Recognition Memory Test: Manual*. Berkshire, UK: NFER-Nelson.

Warrington, E.K. and Shallice, T. (1980). Wordform dyslexia. *Brain*, 103, 99-112.

Waters, G. and Caplan, D. (1996). The capacity theory of sentence comprehension: Critique of Just and Carpenter (1992), *Psychological Review* 103, 761-772.

Waters, G., Caplan, D. and Hilderbrandt, N. (1991). On the structure of verbal short-term memory and its functional role in sentence comprehension: Evidence from neuropsychology. *Cognitive Neuropsychology*, 8, 81-126.

Weinrich, M., Boser, K. I, McCall, D. (1999). Representation of linguistic rules in the brain: evidence from training an aphasic patient to produce past tense verb morphology. *Brain and Language*, 70:144-58.

Wilson, B. A., and Evans, J. J. (1996). Error-free learning in the rehabilitation of people with memory impairments. *Journal of Head Trauma Rehabilitation*, 11 (2), 54-64.

Wilson, B. A., Baddeley, A. D., Evans, J. J. and Shiel, A. (1994). Errorless learning in the rehabilitation of memory impaired people. *Neuropsychological Rehabilitation*, 4, 307-326.

Wing, A. M., and Baddeley, A. D. (1980). Spelling errors in handwriting: A corpus and distributional analysis. In: U. Frith (Ed.), *Cognitive processes in spelling* (251-285). London: Academic Press.

Appendix

Appendix A. Guidelines for the scoring of syntax (Chapter 2)

Sentence group	**Group description**	**Canonical / non-canonical**	**Inflection type**	**1 point criteria**
1 'the man is adjusting the mirror'	Transitive, accusative	Canonical	where verb has no inflection, or the +s inflection	'being' doesn't appear before verb 'by' doesn't appear after verb
2 'the girl hides near the dog'	Unaccusative, uses locative relation	Canonical	where verb has no inflection, or the +s inflection	'near' or 'next to' appears after verb 'being' doesn't appear between first noun class and verb
3 'the bird watches out for the dog'	Unaccusative, 'in attribution to', using 'for'	Canonical	where verb has no inflection, or the +s inflection	'for' appears after verb 'being' doesn't appear between first noun class and verb
4 'the girl shows off to the woman'	Unaccusative, expressing direction towards something using 'to' or 'at' or 'on'	Canonical	where verb has no inflection, or the +s inflection	'to' or 'at' appears after verb 'being' doesn't appear between first noun class and verb
5 'the train leaves the station'	accusative, where verb alone is sufficient to achieve a relation between noun classes	Canonical	where verb has no inflection, or the +s inflection	'being' doesn't appear between first noun class and verb
6 'the mirror is being adjusted by the man'	Transitive, accusative	Non-canonical	where verb has no inflection, or the +s inflection	'by' appears after verb
7 'the dog is near the girl who hides'	Unaccusative, uses locative relation	Non-canonical	where verb has no inflection, or the +s inflection	'near' or 'next to' or 'in front of' appears after verb 'who' appears before agent
8 'the dog is watched out for by the bird'	Unaccusative, 'in attribution to', using 'for'	Non-canonical	where verb has no inflection, or the +s inflection	'for by' appears after verb
9 'the woman is shown off to by the girl'	Unaccusative, expressing direction towards something using 'to' or 'at' or 'on'	Non-canonical	where verb has no inflection, or the +s inflection	'to by' or 'at by' or 'on by' appears after verb
10 'the woman is taken over by the man'	unaccusative, where verb alone is sufficient to achieve a relation between noun classes	Non-canonical	where verb has no inflection, or the +s inflection	'by' appears after verb

<u>Appendix B. Irregular past tense untreated verb sets: allocated into groups according to McClelland and Patterson (2002), Chapter 2</u>

Untreated verbs with *treated* quasi-regular affixes				Untreated verbs with *untreated* quasi-regular affixes				
Group 2	Group 4	Group 9	Group 8	Group 1	Group 3	Group 5	Group 6	Group 7
Keep-kept	Bleed-bled	Tear-tore	Bring-brought	Make-made	Cast-cast	Smell-Smelt	Deal-dealt	Build-built
Say-said	Breed-bred	Sing-sang	Teach-taught	Have-had	Hit-hit	Spell-spelt	Mean-meant	Bend-bent
Do-did	Read-read	Get-got	Think-thought		Let-let	Burn-burnt	Dream-dreamt	Lend-lent
Sleep-slept	Speed-sped	Shine-shone			Put-put	Dwell-dwelt	Feel-felt	Send-sent
	Ride-rode	Draw-drew			Quit-quit	Spill-spilt	Kneel-knelt	Spend-spent
	Slide-slid	Grow-grew			Rid-rid	Learn-learnt		
	Fight-fought	Sew-sew			Set-set			
		Throw-threw			Shed-shed			
					Slit-slit			
Treated items from each group								
Hear-heard	*Hide-hid*	*Drink-drunk*	*Catch-caught*					
Lose-lost	*Lead-Led*	*Bind-bound*	*Seek-sought*					
Flee-fled	*Feed-Fed*	*Sink-sunk*	*Fight-fought*					

Appendix C. Scores across sessions including McNemar tests for the significance of changes (Chapter 2)

Sessions compared	Past tense inflection scores (reverse items)	McNemar significance value	Present tense inflection scores (linear items)	McNemar significance value
Session 1 and session 2	10/38 vs. 16/38	$p = .286$	38/38 vs. 25/38	$p < .001$
Session 1 and session 2b	10/38 vs. 2/38	$p = .039$	38/38 vs. 22/38	$p < .001$
Session 1 and session 3	10/38 vs. 11/38	$p = 1.000$	38/38 vs. 38/38	N.S.
Session 1 and session 3b	10/38 vs. 31/38	$p < .001$	38/38 vs. 16/38	$p < .001$
Session 1 and session 4	10/38 vs. 27/38	$p = .001$	38/38 vs. 15/38	$p < .001$
Session 1 and session 4b	10/38 vs. 38/38	$p < .001$	38/38 vs. 38/38	N.S.
Session 2 and session 2b	16/38 vs. 2/38	$p = .001$	25/38 vs. 22/38	$p = .581$(b)
Session 2 and session 3	16/38 vs. 11/38	$p =.227$	25/38 vs. 38/38	$p < .001$
Session 2 and session 3b	16/38 vs. 31/38	$p = .001$	25/38 vs. 16/38	$p = .022$
Session 2 and session 4	16/38 vs. 27/38	$p = .013$	25/38 vs. 15/38	$p = .041$
Session 2 and session 4b	16/38 vs. 38/38	$p < .001$	25/38 vs. 38/38	$p < .001$
Session 2b and session 3	2/38 vs. 11/38	$p = .012$	22/38 vs. 38/38	$p < .001$
Session 2band session 3b	2/38 vs. 31/38	$p < .001$	22/38 vs. 16/38	$p = .210$
Session 2b and session 4	2/38 vs. 27/38	$p < .001$	22/38 vs. 15/38	$p = .143$
Session 2b and session 4b	2/38 vs. 38/38	$p < .001$	22/38 vs. 38/38	$p < .001$
Session 3 and session 3b	11/38 vs. 31/38	$p < .001$	38/38 vs. 16/38	$p < .001$
Session 3 and session 4	11/38 vs. 27/38	$p = .001$	38/38 vs. 15/38	$p < .001$
Session 3 and session 4b	11/38 vs. 38/38	$p < .001$	38/38 vs. 38/38	N.S.
Session 3b and session 4	31/38 vs. 27/38	$p =.388$	16/38 vs. 15/38	$p = 1.000$
Session 3b and session 4b	31/38 vs. 38/38	$p = .388$	16/38 vs. 38/38	$p < .001$
Session 4 and session 4b	27/38 vs. 38/38	$p = .016$	15/38 vs. 38/38	$p < .001$

Appendix D. Methodological details for tasks comprising the Birmingham Cognitive Screen (BCoS) assessment battery (Chapter 4)

Task	Description
LANGUAGE	
Picture Naming	The participant is presented with 14 black and white picture items (objects, fruit, vegetables and animals) and asked to verbally produce the name of each, under a time limit of 15 seconds per picture.
Sentence Construction	The participant is shows a photograph of a person pantomiming an action (e.g. a woman putting a book in her bag).
Instruction Comprehension	Examiner's judgement of comprehension ability.
Reading	Participants are required to read two sentences and six nonwords.
Writing	Participants are asked to write down 4 real words and a nonword.
PRAXIS / CONTROL AND PLANNING OF ACTION	
Visuo-Constructional Abilities (Complex Figure Copy)	Participants are asked to make an immediate reproduction of an abstract Figure.
Multi-Step Object Use	Participants are presented with several target (torch and batteries) and distracter (visual – glue stick, semantic – matches) objects, and are instructed to make the torch work.
Gesture Recognition	The examiner produces 3 gestures (e.g. goodbye) and pantomimes use of 3 objects (e.g. lighter). Participants are required to select the target gesture from 4 alternatives.
Gesture Production	The participant is asked to produce 3 gestures (e.g. hitch-hiking) and to pantomime the use of 3 objects (e.g. a hammer).
Meaningless Gesture Imitation	The examiner produces 4 meaningless gestures and the participant is required to produce the same gesture once the examiner is finished.
LONG-TERM MEMORY	
Orientation	Patients are asked questions about personal information (address, birthdate) and time and space (e.g. current year)
Episodic Memory (Newly Acquired Knowledge)	Patients are read a story and are asked to recall as many details of it as possible in both a free recall, and recognition condition. The recall and recognition components are tested both immediately afterwards (immediate recall, recognition), and after a delay of 5 intervening tasks (delayed recall, recognition).
ATTENTION AND EXECUTIVE FUNCTIONS	
Spatial Attention	*Visual extinction*: The examiner positions his/her index fingers approximately 20 cm either side of his/her head. The participant is instructed to observe the examiner making finger movements either on the left, right, or both sides and report where the movements occurred. *Tactile extinction*: The participant places his/her hands palm-down on the table and is asked to close his/her eyes. The examiner lightly taps the participant's right or left hand or both hands concurrently, and the patient reports which side was tapped. *Key cancellation*: An A4 page with landscape set up is positioned in front of the participant at participant's midline. Participants are instructed to cross out any pictures of keys on the page, amongst distracter pictures.
Controlled Attention	*Auditory attention*: The participant is played a recording of a man saying some target and distracter words, and is asked to tap on a table whenever he/she hears one of three target words. *Rule-finding and switching*: The participant follows a dot moving around a grid from page to page according to rules that change unpredictably. Their job is to anticipate where the dot will move next.
MATHEMATICAL /NUMBER ABILITIES	
Number Reading	The participant reads nine numerical items (numbers, prices and clock times).
Number Writing	The participant writes numbers and prices to dictation.
Calculation	The participant completes 4 calculations.

Appendix E. Methodological details for stimuli designed to assess phonological and semantic STM impairments (Chapters 6 and 7)

The data reported from numbers i-iii, vi-viii and x used memory span tasks, where the examiner read out lists of items and asked patients to recall as many items as possible. For each patient, the data are taken from accuracy rates at list lengths one item above verbal-verbal word span unless reported otherwise. Consistent with previous work on the subject (Jefferies, Hoffman, Jones and Ralph, 2008), primacy and recency effects in i and vi were calculated from the number of items correctly recalled in the medial positions and the initial and final positions respectively. For patients DS and AK, a list length of 5 was used, while MM was tested with lists of 4 items.

The iv and ix data are computed by contrasting results of a phonological probe task (e.g. name the items in the current list that rhymed with cat) with those from a semantic probe task (e.g. name the items in the current list that were items of furniture). For every patient, each trial randomly presented 6 items from three semantic / phonological categories, and probed one of these pairs: For example, in the semantic case, the items chair, lion, chisel, tiger, table, hammer were presented followed by the probe ‘name the animals’, the correct response being lion and tiger. The v data contrasted comprehension of sentences where the noun appeared after versus before adjectives, and a poorer performance on after conditions indicated impaired semantic STM (Martin and Freedman, 2001).

Appendix F. Lexical variable information for Experiment 1a, Chapter 7

	N	**Celex frequency**	**Log frequency**	**Familiarity**	**Letter length**	**Imageability**
Open	480	57.991	1.18	353.5	5	414.6292
Closed	15	67.706	1.34	360.266	5	442.0667

Appendix G. Lexical variable information for Experiment 2, Chapter 7

Set		*N*	Celex frequency	Log frequency	Familiarity	Letter length	Imageability
Phonological NO	Related	153	65.021	1.15	345.732	4.93	400.11
	Unrelated	153	64.233	1.24	365.037	4.95	420.97
Phonological YES	Related	153	48.19	1.10	339.29	4.93	397.32
	Unrelated	153	47.21	1.12	326.44	4.916	409.18
Semantic NO	Related	153	64.822	1.29	371.10	5.03	408.58
	Unrelated	153	62.84	1.30	368.18	4.93	455.3509
Semantic YES	Related	153	47.87	1.13	336.97	4.97	435.93
	Unrelated	153	37.08	1.12	314.22	4.98	418.85

Appendix H. Accuracy scores across present trials for Experiment 2 (Chapter 7)

	Relatedness	**Unrelated yes**	**Related-1 yes**	**Related-2 yes**	**Related-3 yes**
DS	Phonological	100/108 (92.5%)	32/36 (88.8%)	30/36 (83.3%)	32/36 (88.8%)
	Semantic	105/108 (97.2%)	33/36 (91.6%)	29/36 (80.55%)	35/36 (97.2%)
AK	Phonological	98/108 (90.74%)	32/36 (88.8%)	31/36 (86.1%)	36/36 (100%)
	Semantic	100/108 (92.5%)	35/36 (97.2%)	34/36 (94.4%)	33/36 (91.66%)
MM	Phonological	54/108 (50%)	24/36 (50%)	18/36 (50%)	19/36 (52.7%)
	Semantic	60/108 (55.5%)	15/36 (41.6%)	14/36 (38.8%)	20/36 (55.5%)
Controls	Phonological	510/520 (98.07%)	175/180 (97.22%)	179/180 (99.4%)	179/180 (99.4%)
	Semantic	518/520 (99.61%)	179/180 (99.4%)	169/180 (93.8%)	176/180 (97.7%)