

1. INTRODUCTION

Multiple sclerosis is a chronic inflammatory disease of the central nervous system in which the myelin sheaths encasing the neurons are attacked. Over the course of the disease, lesions are formed in the white matter of the cerebral hemispheres, cerebellum and brainstem (Friend et al., 1999). In addition to the well-understood inflammatory component, recent research has also focused on a possible neurodegenerative component (Lycklama à Nijeholt et al., 1998), linking MS to other neurodegenerative diseases such as Alzheimers, Parkinsons, and Motor Neurone disease.

MS has two main forms of initial presentation: relapsing-remitting and primary-progressive. Relapsing-remitting (RR) MS is expressed by recurring acute inflammatory episodes separated by periods of respite, whereas the primary-progressive (PP) form of the disease, affecting 15-40% of patients, is characterized by a subtle onset of symptoms which slowly advance over time without acute exacerbations or remissions (Thompson et al., 1997). The relapsing-remitting typology often develops into secondary progressive (SP) MS at a later stage, in which the remissions diminish and the symptoms become continuous (Thompson et al., 1997). The neurodegenerative component of MS may be more pronounced in the two progressive forms of the disease, especially in PP MS (Lycklama à Nijeholt et al., 1998).

The most noticeable symptoms of any form of MS are usually sensory and/or motor deficits (Heaton et al., 1985), with weakness, spasticity, ataxia, sensory loss, and urination difficulties troubling the majority of sufferers (Laakso et al., 2000). However, there is growing evidence that MS also affects cognitive function in at least a subset of cases. Memory deficits (Thornton & Raz 1997; Rao et al., 1989a), attention problems (Ron et al., 1991; Callanan et al., 1989) and cognitive slowing (Rao et al. 1989b) have all been reported.

Nevertheless, it remains true that only a minority of MS research is directed at the cognitive domain, and within that research there has been little focus on language deterioration, which will be the subject of the current study. This is partly because the lesions of MS are mainly found in subcortical areas, which do not correspond to the established language regions of the brain; language research has therefore not often been

considered a productive area of enquiry (Friend et al., 1999). However, it is a simplification to state that all language processing is carried out in the ‘language regions’ of the brain. Some researchers have stressed the importance of the subcortical white matter pathways for language, as they connect the cortical language regions to deep brain structures (eg. Alexander et al., 1987; Robin & Schienberg, 1990). Damasio et al. (1982) even observed language difficulties secondary to subcortical lesions in a different context, lending empirical evidence to the proposed relationship.

Given the presence of subcortical pathways supporting language, it seems logical that MS could cause linguistic disorders. There is, unfortunately, a paucity of comprehensive research in the field, which this study aims to address.

2. REVIEW OF THE RELEVANT LITERATURE

In the relatively few papers that have investigated the language abilities of people with MS, results have been somewhat contradictory. This may not in fact be particularly surprising as the disease course of MS is notoriously variable (Murdoch & Theodoros, 2000; Lucchinetti et al., 2000). Nonetheless, there are some patterns that bear further investigation, and there is a definite suggestion that language disorders can indeed be a symptom of MS in many cases.

2.1 LANGUAGE CAN BE AFFECTED IN MS

Though the most noticeable symptoms of MS are in the area of motor control, there are patients who complain of subjective language difficulties, primarily with word-finding (Laakso et al., 2000). For some, the problem is much more pronounced, and rare cases of full-blown aphasia have been documented: individual patients with motor aphasia (Olmos-Lau et al., 1977), global aphasia (Friedman et al., 1983), severe dysfluent aphasia (Achiron et al., 1992), and a language disorder with alexia and agraphia (Day et al., 1987) have all been reported. These extreme case studies give straightforward evidence that MS can affect language. However, these cases clearly represent atypical presentations of the disease, and as such cannot themselves influence a general hypothesis of language deficits in MS. We must therefore turn to group studies to get a clearer picture of the known language abilities of the majority of MS patients.

Many studies of cognitive abilities in MS have included a language component, which often consists of a confrontation naming test and/or a verbal fluency test embedded in a neuropsychological battery (Murdoch & Theodoras, 2000). In addition to this, some studies purely interested in language also use such tests as a source of data collection. As a result, naming and fluency tests are conducted with relative frequency with MS patients. However, neither type of test has given uncontroversial results. Some researchers have reported impairment with naming (Jambor, 1969; Beatty et al., 1988; Pozzilli et al., 1991), while others have found no such impairment (Ron et al. 1991; Callanan et al., 1989). Similarly, fluency tests have revealed both impairments (Pozzilli et al., 1991; Heaton et al., 1985) and intact performance (McAllister et al., 2005; Franklin et al., 1988).

Often there is variance even within a participant group. For example, Blackwood et al. (1991) found that the MS group as a whole performed within normal limits on a naming task, but on closer inspection four individuals were identified to have impaired performance (such impaired subsets are also found in some of the research discussed below in section 2.3; their existence contributes to the predictions for the current study). To account for the variable findings in naming and fluency tests, many authors (eg. Friend et al., 1999; Kujala et al., 1996) have blamed inconsistent participant recruitment criteria. The studies cited above vary greatly on control for disease subtype, exacerbation versus stable status and presence or extent of cognitive decline, all of which might be expected to have an impact on performance on language tests.

It is problematic that naming and fluency tests should account for so much of the literature on language in MS, as neither constitutes a complete or even terribly sensitive measure of linguistic ability. Naming requires the production of individual lexical items but places no strain on syntax, and very little emphasis on phonology. More worryingly, fluency tests, which require the participant to produce as many words as possible in one minute starting with a given letter or belonging to a given semantic category, are a notoriously slippery measure of linguistic prowess. Kennedy and Murdoch (1989) warn against using verbal fluency as the sole measure of language ability as there are too many cognitive components involved in successful completion of the task. Arnett et al. (2008) further caution that in timed tests requiring verbal responses, patients with MS are at a

disadvantage purely due to motor constraints. Analysis revealed that oromotor slowing accounted for a significant amount of participants' difficulty with neuropsychological batteries.

In clinical settings, health care professionals often assess language as though it were one indivisible unit, with only a naming or fluency test to see if it is 'intact' (Heaton et al., 1985). Linguistic theory, as shall be discussed in the next section, would suggest that each facet of language be considered separately in order to come to a more thorough diagnosis. Simple naming and fluency tests, though revealing, do not give a complete picture of language function.

2.2 THE LINGUISTIC PERSPECTIVE

It is generally assumed in the linguistic literature that different domains of language (eg. syntax, phonology) should be considered separately because they are seen to be interconnected but differentiable subsystems of language function. There is empirical evidence to support this view, arising from both the literature surrounding second language learning across the lifespan, and research on neurological damage.

Many researchers concerned with second language acquisition have made the assumption that the domains of language are separate (eg. Seliger, 1978; Eubank & Gregg, 1999; Long, 1990). Research in this field, especially as concerns the purported presence of a 'critical period' for learning language, is rife with disagreement, but nonetheless this simple assumption seems to find justification. The evidence comes from studying adult and child learners of second languages; in some domains, only child learners ever master the second language, while in others adults are at no disadvantage. It seems that the ability to master second language phonology is particularly sensitive to age: perfect acquisition of a native-like accent in a second language has been shown to be virtually nonexistent after puberty (eg. Flege et al., 1999). On the other hand, studies of second language syntactic abilities that probe very complex and difficult structures have revealed small but significant groups of adult learners who have attained near-native syntax (Montrul & Slabakova, 2000; Hopp, 2009; White & Genessee, 1996). Meanwhile, the ability to acquire semantic knowledge of a second language never seems to diminish, with the ability to acquire new lexical items preserved throughout adulthood (Snow &

Hoefnagel-Hohle, 1978). These findings, taken together, suggest different neurological underpinnings for the different domains of language, since some seem more constrained by age than others.

There are also cases in the literature in which an individual has missed the putative critical period in the acquisition of their first language, though these are thankfully rare. Often cited is the case of Genie, a girl denied language input from the age of 1.5 until she was 13. When Genie finally learned to speak, she was able to acquire about 250 words and understand their semantics, but she was never able to acquire syntactic competence (Curtiss, 1974). Of course, Genie is a problematic source of evidence due to her intense psychological trauma and other cognitive lacunae, but it has been pointed out (Eubank & Gregg, 1999) that there is no reason these factors should affect her syntax more than her semantics; Genie's differential achievements do support a pluralistic conceptualisation of linguistic abilities. As a single case study, however, we can only draw limited conclusions from her experience.

Support with greater statistical validity comes from the study of acquired language disorders, in which previously acquired language abilities are lost during adulthood due to brain injury or disease. There are many disorders that can affect one domain of language while leaving others intact, which strongly suggests that in the current study of language in MS it would be relevant to consider multiple facets of language. Examples of individually impaired phonology include foreign accent syndrome, in which the patient develops abnormal pronunciation patterns in their first language after trauma or in the early stages of organic brain disease (Luzzi et al., 2008). Semantics can be singled out in Gogi aphasia, which is characterized by preserved fluent speech and correct use of function words, but difficulty finding concrete words. This condition intriguingly shows selective reading deficits in Japanese speakers who know both Kanji and Kana scripts: the holistic, non-phonological Kanji script becomes impaired while phonological Kana is preserved (Sasanuma & Monoi, 1975). Gogi aphasia is thought to correspond roughly to the Western concept of 'semantic dementia' (Mummery et al., 1999). Broca's aphasia is the classic example of selectively impaired syntax, and though there have been many differing opinions on the exact characterization of this condition, it is true that the production of lexical words is preserved compared to

that of function words and syntactic structure, and that language comprehension remains largely unimpaired (Zurif et al., 1993), suggesting a deficit that is more syntactic than semantic.

Such dissociations have also been observed on the neurological level in unimpaired participants with fMRI studies; mounting evidence suggests that morphological, syntactic and semantic information processing take place in different centres in the brain (Depreto & Bookheimer, 1999; Moro et al., 2001). It is important to take such evidence into account when studying ‘language’, and remember that we are not speaking of a single ability, but rather a cluster of interconnected but dissociable skills.

2.3 BEYOND NAMING AND FLUENCY

Given the above arguments in favour of a composite view of language, it seems unfortunate that the majority of the research on language in MS has focused on only two simple tests, naming and verbal fluency, which do not represent the full array of linguistic domains. However, a few studies have investigated different areas of language use, and together they hint at wide-ranging potential deficits for MS patients.

2.3.1 SYNTAX

a. RECEPTIVE

Grossman et al. (1995) looked at the receptive syntactic abilities of people with MS using a picture-matching task designed by the authors. The participants heard a sentence, and then were asked to point to whichever of four drawings represented the meaning of the sentence. The stimuli were manipulated for grammatical voice and presence and location of a relative clause. The three foil pictures represented a subject-object reversal, incorrect agent error, and incorrect verb error. As another test, participants were asked to answer comprehension questions about sentences similar to those in the first task, but with additional manipulation of semantic reversibility. It was found that in these comprehension tasks, MS patients as a group performed comparably to controls, though with a significant predominance of subject-object reversal errors, unlike controls. A subgroup was identified having significant impairment on subordinate

phrases, passives, and semantically reversible sentences – the more demanding syntactic procedures, which cannot be compensated for with semantic knowledge.

b. EXPRESSIVE

When it comes to expressive measures of syntax, there is mixed evidence for impairment in MS. Wallace & Holmes (1993) tested four participants with the Arizona Battery for Communication Disorders and found significant problems with generative writing. King (2009), however, performed an analysis on naturalistically collected language data produced by people with MS and found no discernible difference between the mean length of utterance, number of clauses, or syntactic correctness compared with controls.

2.3.2 SEMANTICS

a. RECEPTIVE

Receptive semantics was tested in a study by Laatu et al. (1999), in which words were given to the participants and they were asked to call up or recognize various traits and associations to do with those words. This was done with both concrete and abstract stimuli, the participants being expected to describe the concrete words (for instance, *beaver*) and define the abstract ones (such as *comedy*). The participants then ranked the relative importance of various traits to a given word, and organized hierarchically related words (eg. activities, sports, basketball) according to the relationships between them. It was found that the participants with MS were impaired on all tasks compared with controls.

b. EXPRESSIVE

Expressive measures of semantics include naming tests, where the participants are asked to produce target lexical items. These tests, as described in section 2.1, are performed fairly regularly with MS patients, with mixed results.

2.3.3 PHONOLOGY

a. RECEPTIVE

Receptive phonological skills can be measured using auditory discrimination tasks. Ivnik (1978) performed such a task as part of a larger cognitive battery. He had participants listen to nonsense syllables, then underline on a sheet which written equivalent most closely corresponded to the sound they had heard. The findings indicated that there was no impairment in the MS group on this task.

b. EXPRESSIVE

A study on expressive phonology was carried out by Kujala (1996). Participants were given a non-word production task, with written stimuli. They read a list of pseudo-words aloud. The authors found that MS patients with cognitive decline performed significantly worse than cognitively preserved MS patients on this task, indicating a phonological deficit that is cognitive, rather than perceptual, in nature.

2.3.4 WRITTEN LANGUAGE

a. RECEPTIVE

Compared to the previously described domains, a relatively high number of studies have measured written language ability in MS. Receptive skills (reading) have been measured variously with: timed list-reading (Jennekens-Schinkel et al., 1990), in which participants were able to read fewer words than controls; the Schonell Graded Word reading test (Jambor, 1969), in which participants showed a slight impairment reading single words aloud in an untimed task; and paragraph-reading followed by comprehension questions (Kujala et al., 1996), in which participants were significantly impaired.

b. EXPRESSIVE

Expressive skills (writing) have been investigated using analysis of errors in writing samples (Jennekens-Schinkel et al., 1990), which revealed no impairment, and with spelling tests, which showed a small but significant impairment in the study of Kujala et al. (1996), but not in that of Jambor (1969).

2.4 DISEASE COURSE

There is one further issue on which the literature is inconclusive, and this is the effect of disease course on cognitive and language symptoms. This issue is relevant to the current study because the relapsing-remitting and primary-progressive subtypes differ on many counts, including age of onset, prognosis with regard to disability and mortality, epidemiology, and response to immunosuppression (Olerup et al., 1989). Indeed, the literature tends to report a discrepancy in language deficits, with greater problems attributed to chronic-progressive MS than to relapsing-remitting MS (eg. Heaton, 1985; Friend et al., 1999). Unfortunately, it is not always clear where the secondary progressive patients fit in this dichotomy. However, some studies have found no effect of disease course (eg. Wishart & Sharpe, 1997) and a great many studies do not even consider it as a factor, sometimes neglecting to identify the type of MS suffered by participants (eg. Callanan et al., 1989; Ivnik, 1987; Peyser et al., 1980). Henry & Beatty (2006) found that factoring out age, duration of illness and neurological disability attenuated the observed advantage of relapsing-remitting over chronic-progressive, implicating confounding factors in the discrepancy.

3. PURPOSE OF THIS STUDY

Overall, the studies described in 2.1 and 2.3 paint an intriguing picture of deficits in multiple areas of linguistic function in MS. However, the findings are fragmented, with each study looking at one or a few aspects of language. Since many of the studies had different recruitment criteria, it is impossible to draw conclusions between them. For instance, Ivnik (1978) found no impairment with receptive phonology, but Kujala et al. (1996) observed problems with expressive phonology *in cognitively impaired patients*. The discrepancy could be due to the receptive/expressive dichotomy, or to the presence of a group with cognitive impairment – without performing both tests on the same group of patients, it is impossible to tell. Because of such confounding variables, we cannot yet say whether there is any connection between the various observed language deficits. It might be that some of the problems are associated. Alternately, they could all be individual effects of the patients' disease progression. As yet, there is no study directly comparing MS patients on all the measures listed above.

The present study will therefore attempt a comprehensive overview of multiple facets of language in order to ascertain whether there are patterns of language ability and disability that mark MS, or whether subsets of patients with specific language disorders can be identified. Furthermore, given the lack of consensus in the literature concerning the influence of disease course on language deterioration, this variable will be carefully documented in an attempt to find correlations between disease course and language ability. Studies vary on whether they collapse secondary progressive and primary progressive MS together into a ‘chronic progressive’ group, so this study will analyse them separately to see whether the combination is justified. The components of language that will be tested are phonology, syntax, semantics and written language; each of which will be further broken down into tests of both receptive and expressive language abilities.

4. METHODS

4.1 MATERIALS

Tests were chosen with consideration to the abilities and challenges of people with MS, as well as to address some of the possible confounds present in the literature cited in 2.3. None of the tests chosen are timed, and the physical strain was kept to a minimum, with limited written and oral responses (many of the tasks involve only pointing). All computer-based tests were presented using a Toshiba laptop computer with a 10-inch touchscreen.

4.1.1 SYNTAX

a. RECEPTIVE

For receptive syntax, the Test of Reception of Grammar (TROG) (Bishop, 1983) was used. This test is similar to one of those employed by Grossman et al. (1995) in that participants hear an orally presented sentence, then point to one of four pictures which illustrates the content of the sentence. The presence of relative clauses and the grammatical voice of the sentence are manipulated, and most sentences are semantically reversible. Moreover the distractor pictures represent subject-object reversal, amongst other options, which was the major error made by patients in the study of Grossman et al. The TROG has the added condition, however, of including trials in which coloured

shapes are the arguments, and spatial relations (eg. to be inside) are the verbs, which reduces the efficacy of semantically supported comprehension strategies, forcing participants to rely purely on syntax. Additionally, this test has the practical advantage of being widely used and standardized in both child and, increasingly, brain damaged populations. In order to decrease the time taken, only the second 40 items of the tests were used. As the TROG increases in difficulty, these are the most syntactically demanding sentences. The pictures were displayed on the computer screen, while the sentences were read aloud in a naturalistic manner.

b. EXPRESSIVE

Expressive syntax was measured with the Boston Cookie Theft description test (Goodglass & Kaplan, 1983), in which participants are asked to describe the events shown in a very dynamic illustrated scene. This test elicits natural language samples, which should be comparable to those analysed by King (2009). The participants were asked to write their descriptions down rather than saying them aloud in order to promote grammatical production: spoken language is generally less complete and syntactically clean than written (Pinker, 1994). These language samples were then analysed for Mean Length of Utterance in morphemes (MLU), and given Complexity Index (CI) scores.

MLU is a measurement commonly used in syntactic research with children, but is a useful measure with adult samples as well King, (2009). It is calculated by dividing a given language sample into utterances – in this case written sentences finished with a period – and counting the number of morphemes in each utterance. The total number of morphemes is then divided by the number of utterances to find the mean. Chapman's (1981) guidelines for counting morphemes were adapted for use on written language samples by adults. The guidelines used are as follows:

- All compound words (two or more free morphemes), proper names and ritual reduplications count as single words
- All irregular pasts of the verb (*got, did, went, saw*) count as one morpheme
- All auxiliaries (*is, have, will, can, must, would*) count as separate morphemes. [...] All inflections, for example, possessive (*s*), plural (*s*), third person singular (*s*), regular past (*ed*) and progressive (*ing*), count as separate morphemes

CI is a measurement used to determine the extent to which an individual makes use of dependent clauses. This was calculated according to the protocol laid out in the Edmonton Narrative Norms Instrument (Schneider et al., 2005). In brief, the language sample must first be divided into C-units, which are often but not always congruous with the sentences in the text. C-units are limited to only one main clause; therefore, coordinated sentences (eg. I lost my wallet and my sister gave me money) must be split into two C-units. The number of dependent clauses is then counted, added to the number of main clauses, and the total is divided by the number of C-units.

4.1.2 SEMANTICS

a. RECEPTIVE

The test employed for receptive semantics was a shortened version of the Pyramids and Palm Trees test (PPT) (Howard & Patterson, 1992), the Kissing and Dancing test (KDT) (Bak & Hodges, 2003), and the Tomatoes and Tuna fish test (TTT) (van der Hulst et al., 2010). The PPT is concerned with the relationships between nouns, the KDT deals with verbs, and the TTT, sequential events and cause-effect scenarios. This set of tests involves making decisions about sets of pictures. Participants must choose which of two alternatives is best associated with a third image. For instance, palm trees are associated with pyramids according to geography, but the other option, pine trees, has no such connection. These tests require only touching the correct image on the computer screen and are simple and straightforward to perform. They were chosen in preference to more complex tasks like those employed by Laatu et al. (1999), which make additional cognitive and linguistic demands. Laatu et al. found impairment in all their tests; it will be interesting to see if there is a fundamental semantic impairment on this basic level.

b. EXPRESSIVE

Expressive semantics was explored with a naming test, as in much of the literature. The Graded Naming Test (GNT) (McKenna & Warrington, 1983) was chosen due to the range of difficulty it provides. Given the variability other researchers have observed in naming tests, it was important to prepare for the possibility of different levels

of impairment. The test consists of black and white line drawings, and participants simply provide the names of the items out loud.

4.1.3 PHONOLOGY

a. RECEPTIVE

Receptive phonology was measured using a minimal pairs discrimination test. Words that differ by only one phoneme are spoken aloud by recorded voices in a powerpoint slideshow developed by P.J. Rewaj (personal communication, May 2011). There are two voices, one male and one female – from the East and West of Scotland, respectively – but both words of a given minimal pair were spoken by the same speaker. There was a one-second delay between the words, and participants were asked to give their judgement aloud immediately after hearing the words. The test included two conditions; in one, the pairs were real words of one or two syllables (eg. puddle, puzzle), in the other they were non-words of comparable length and syllabic complexity, never violating the rules of English phonology (eg. tusset, tuzzet). The phoneme that differed between the two words was always a consonant, and would vary only by one feature (place, manner or voicing). The target phonemes were also varied according to their position in the word (initial, middle, final or metathetic – switching the first and last consonant). An advantage of using a minimal-pairs task is that it rules out the need for any reading, as in the study of Ivnik (1978), which may be confounded by visual difficulties, or problems specific to written language. In the minimal pairs test, on the other hand, participants need only say aloud if the words they hear are the same or different.

b. EXPRESSIVE

Expressive phonology was tested with a non-word repetition task, again with oral presentation by the recorded Scottish voices to avoid visual confounds. This is an improvement over Kujala et al.'s (1996) non-word reading test, not only because the reading component is eliminated, but because the test is untimed. As Arnett et al. (2008) showed, using timed tests with MS patients disproportionately decreases their performance.

There were two conditions for the patients: immediate and delayed repetition. In the second condition, the test included a delay of 15 seconds between the time they heard the non-word and the time they repeated it out loud. The justification for this second condition comes from Lucchelli & Papagno (2005), a study concerned with patients who have anarthria, a loss of motor control over the speech apparatus. The authors attributed spelling deficits in these patients to impaired subvocal rehearsal, which they suggested was linked to their loss of overt motor control. In an extension of that idea, this delayed repetition test aims to tax the abilities of patients with MS, who also have motor control issues. Even in patients who do not exhibit any noticeable anarthria, there could be some incipient decline of function in this area, which might impair their ability to retain a word via subvocal rehearsal over a 15 second pause.

If this is the case, it would be interesting to see if a similar decline could be invoked in controls by disrupting their subvocal rehearsal: for this reason, a third condition was used for controls in which the 15 second delay was still present, but the participants were asked to repeat ‘the, the, the...’ subvocally during that time. If patients have poor results on the second condition, the results of this third condition would help with interpretation. This test is experimental, and is currently also being tested in a study of patients with motor neurone disease (P.J. Rewaj, personal communication, May 2011).

All the pronunciations were encoded by the examiner and transcribed to IPA for analysis. Rather than simply marking answers right or wrong, a more in-depth analysis was carried out, counting substitutions, deletions, insertions, transpositions (eg. [mɒmb**l**ɒmənt] instead of [mɒmb**o**lmənt]), and transpositions with substitution (eg. [p**a**rslət] instead of [p**r**æslət]) (see Kan et al., 2006, for a longer explanation of this method as applied to spelling errors). The total number of errors was tallied for comparison.

4.1.4 WRITTEN LANGUAGE

a. RECEPTIVE

When testing written language it is of course impossible to eliminate the need for reading. However, the test of reading was chosen to be as accessible as possible. The National Adult Reading Test (NART) was used, which requires the reading of single

words printed in large letters, with one word presented per page. The test consists entirely of irregular words. If participants know the words and can comprehend them as they read, they should be able to provide the correct, irregular pronunciation. If they are not able to access the semantics of the word through reading, they should by default produce a regularized pronunciation.

It should be noted that this test normally serves a very different function in studies of patient populations: it is often used as a test of pre-morbid IQ. This is because in many diseases such as dementia, which affect cognition but not language specifically, the ability to read irregular words is preserved (Crawford et al., 2001). The NART contains some very rare, often antiquated words, and the number a person is able to read correctly will tend to correlate very well with the IQ they demonstrated before the onset of their illness – though it is debatable whether it truly measures IQ, or whether it measures education level, which is highly related but not identical. However, in the current study the goal is indeed to find a specific language deficit, so the NART will be used to test for reading deficits, with education level being held constant across the controls and patients.

b. EXPRESSIVE

Expressive written language was analyzed with a spelling test developed by P.J. Rewaj (personal communication, May 2011), with words said aloud by the same Scottish-accented voices as in the phonological tasks. As mentioned in 2.3, most studies to date have found little or no impairment in spelling, so again an extra condition was added to the test in which participants had to observe a delay of 15 seconds between hearing the word and beginning to write. In adding difficulty to the task (and taxing the powers of subvocal rehearsal, as discussed above), it may be possible to amplify small effects. Again, a third condition was added for controls in which they were asked to repeat ‘the, the, the...’ subvocally during the 15 second delay. In the spelling test points were simply assigned for words spelled completely correctly, since there were very few errors.

4.2 PARTICIPANTS

The MS patients recruited for this study all had clinically definite diagnoses of MS according to the 2005 McDonald criteria (Polman et al., 2005). They had no other conditions that might have had a bearing on the tests, including other neuropsychological conditions. There were 20 patients in all, 13 females and 7 males. The mean age of patients was 51.0 (min = 28.3, max = 67.3) and the mean number of years spent in full-time education was 12.7 (min = 9, max = 19). English was the first language or language of primary usage of all the patients. They were drawn from the subject pool of a larger study, in which they were screened using a history and physical examination and classified into subgroups according to disease course. Addenbrooke's Cognitive Examination – Revised (ACE-R) (Mioshi et al., 2006) was also administered, and the results are displayed below in Table 1. Two patients did not complete the ACE-R due to fatigue. One patient did not complete the NART, and another did not complete the PPT, KDT and TTT, due to time constraints. One patient was precluded from completing any of the phonological tests due to a diagnosed hearing impairment of medium severity. Other hearing problems and spelling difficulties were self-reported, and the results of participants with these conditions are taken into consideration in the discussion section. The patients were also asked if they had experienced any subjective language difficulties arising secondary to MS.

There was also a control group of 20 participants, 11 females, 9 males, none of whom reported hearing impairments or dyslexia (a possible bias which will be considered in the discussion). The ACE-R was administered to screen for incipient dementia, and none of the controls scored below 88, which is the conservative cut-off level. The mean age for the control population was 52.5 (min = 24.7, max = 70.8) and the mean number of years of full-time education was 12.9 (min = 9, max = 20).

Table 1. Demographics of MS patient population

Patient	Sex	MS subtype	Years education	Age	ACE-R (/100)	Hearing complaints	Spelling complaints	Subjective language difficulty
1.01	F	SP	12	67.3	92	YES: Skipped tests	No	Yes: Word finding, articulation
1.02	M	SP	13	48.2	96	Yes: Recent infection	No	Yes: Word finding
1.03	F	SP	11	56.1	97	Yes: Childhood problems	No	Yes: Reading aloud, word finding
1.04	F	SP	11	55	89	No	No	No
1.05	M	SP	16	52.2	97	No	No	Yes: Word finding
1.06	F	PP	10	49.5	90	No	No	No
1.07	F	PP	13	64.3	85	No	Yes: Always been bad at spelling	No
1.08	F	SP	15	56.8	100	No	No	Yes: Word finding, articulation
1.09	M	RR	12	54.5	na	No	Yes: dyslexia	No
1.10	M	PP	15	48.8	92	No	No	No
1.11	M	SP	9	51.8	na	No	No	No
1.12	F	RR	9	40.6	95	Yes: identified in army testing, age 15	Yes: Held back in primary school due to spelling	Yes: Spelling out loud, word finding, articulation, mishears
1.13	F	RR	12	48.2	87	No	No	Yes: articulation
1.14	F	RR	10	28.3	89	No	No	Yes: Word finding, articulation
1.15	F	SP	12	51	88	No	No	Yes: Word finding, articulation
1.16	F	SP	17	63.3	88	No	No	No
1.17	M	RR	19	42.1	98	No	No	No
1.18	F	PP	9	42.5	92	No	No	No
1.19	F	RR	12	48.4	96	No	No	Yes: Word finding, articulation
1.20	M	SP	16	52	93	No	No	Yes: Name & place finding

One further participant (age 28.8, with 10 years of full-time education) was initially tested as a control, but it became evident that she was scoring well outside the normal range on several tests. This participant was the sibling of a patient, and, according to the physician who grouped the patients into MS subtypes, there was a query of MS in her medical history, though she does not qualify for diagnosis according to McDonald's 2005 criteria. This participant's data was analysed individually and is treated as a special case in the results and discussion.

5. EXPECTATIONS

Firstly, this study aims to add evidence to the under-researched field of language disorders in MS. Previous research has generally suggested the presence of linguistic problems, but in any given domain of language the support is scanty or mixed. It is expected in this study that language effects will be observed in at least some of the tests performed, probably in only a subset of the participants.

A key question in this study is whether or not the language effects in MS follow some kind of reliable pattern of degeneration. If this were the case, we would expect the same problems to appear in all the patients. However, if (which seems more likely given the well-known variability of MS presentation) the symptoms observed in the literature are due to individual expressions of the disease, it should be possible to view a double dissociation between impairments in any of the different domains, and to find patients with no language impairments at all.

Additionally, it is expected that disease course will have an effect, possibly with progressive types (primary and secondary) engendering more language disintegration than the relapsing-remitting type, because this pattern was observed in multiple studies. It is tentatively hypothesized that each test will reveal a subset of impaired participants, rendering the overall MS score lower for patients than controls, but that these subsets will each contain at least partially different members, with the worse performance observed in progressive-type patients.

6. RESULTS

First, a group analysis was carried out comparing the patients to controls on each language test in an effort to pinpoint tests that are sensitive to MS. Next, to answer the question of whether the course taken by the disease has an effect, the patients were separated into three subgroups (RR, PP and SP) and compared against one another. Following this, an individual analysis was carried out on the patients in order to identify any greatly impaired individuals, using the controls as a standard. This was in order to ascertain whether any observed language deficits were due to a general marginal impairment in the MS patients, or whether there were subgroups with affected language skills. Finally, the results of the one individual – initially tested as a control, then excluded due to potential MS-like traits and markedly poor performance – were examined separately as a case study to investigate the possibility that language tests may be early detectors of MS.

6.1 GROUP ANALYSIS

The TROG scores were analysed with a nonparametric Mann-Whitney U test because the data was not normally distributed. This was due to ceiling effects among the controls, and great variance among the patients, causing heterogenous variance, $F(1,38) = 8.12$, $p < .05$. The patient scores (Mdn = 37.5) were significantly lower ($U = 104$, $p < .05$) than those of the controls (Mdn = 39). The results are shown in Figure 1.

The GNT results also had heterogenous variance, $F(1,38) = 14.4$, $p < .05$, and the Mann-Whitney U test was again used. The patient scores (Mdn = 23) were significantly lower than controls scores (Mdn = 26), $U = 117$, $p < .05$. The results can be seen in Figure 2.

The NART scores as well had unequal variances, $F(1,37) = 6.41$, $p < .05$, requiring a Mann-Whitney U test. The controls (Mdn = 7) made significantly fewer errors than the patients (Mdn = 14), $U = 72.5$, $p < .05$. The results are displayed in Figure 3.

Figure 1. TROG scores of patients and controls

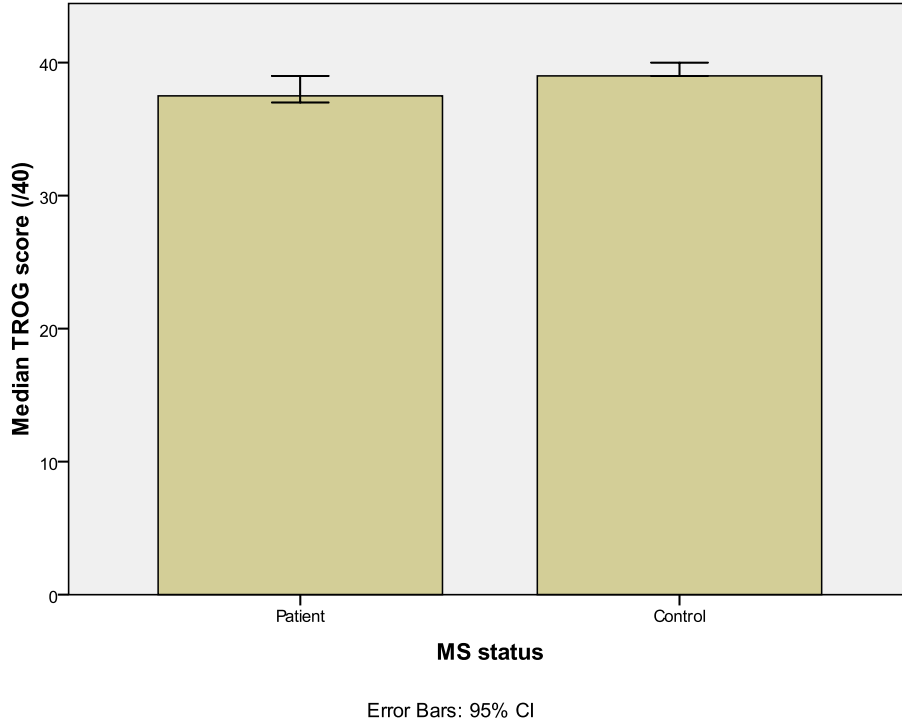


Figure 2. GNT scores of patients and controls

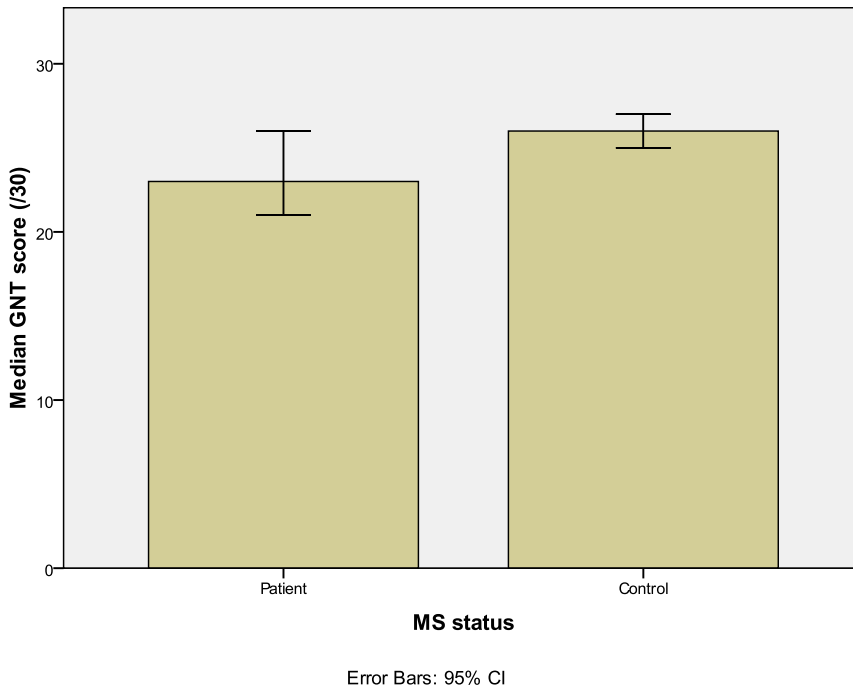
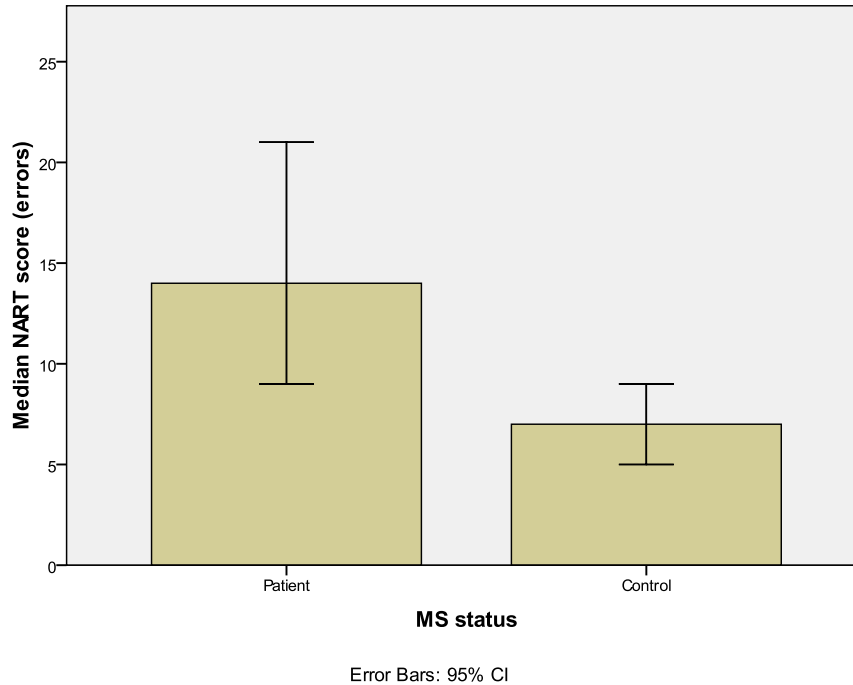
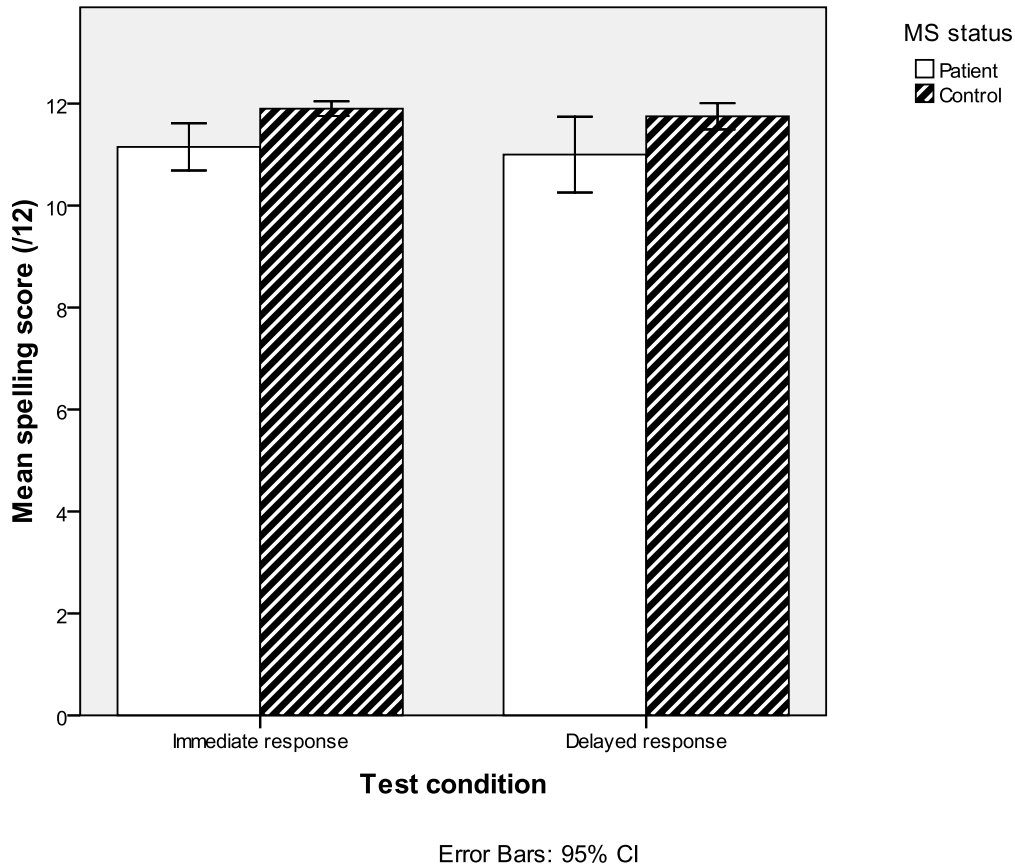


Figure 3. NART error scores of patients and controls



The results from the spelling test were analysed using a factorial mixed ANOVA. The between-subject factor was MS status (patient or control), and the within-subject factor was test condition (immediate response or delay). There was no significant main effect of test condition, $F(1,38) = .922$, *ns*, indicating that overall, the 15 second delay did not cause people to spell more words incorrectly. There was also no significant interaction between MS status and test condition, $F(1,38) < .001$, *ns*. The variance was heterogeneous for both the immediate response condition, $F(1,38) = 32.5$, $p < .05$, and the delay condition, $F(1,38) = 7.07$, $p < .05$, compromising the accuracy of the *F* test for MS status, so the data from the two testing conditions were averaged and a Mann-Whitney U test was carried out on these combined spelling results. The difference between patient and control spelling was significant, $U = 119$, $p < .05$. Figure 4 illustrates the spelling results before they were averaged.

Figure 4. Spelling results for patients and controls

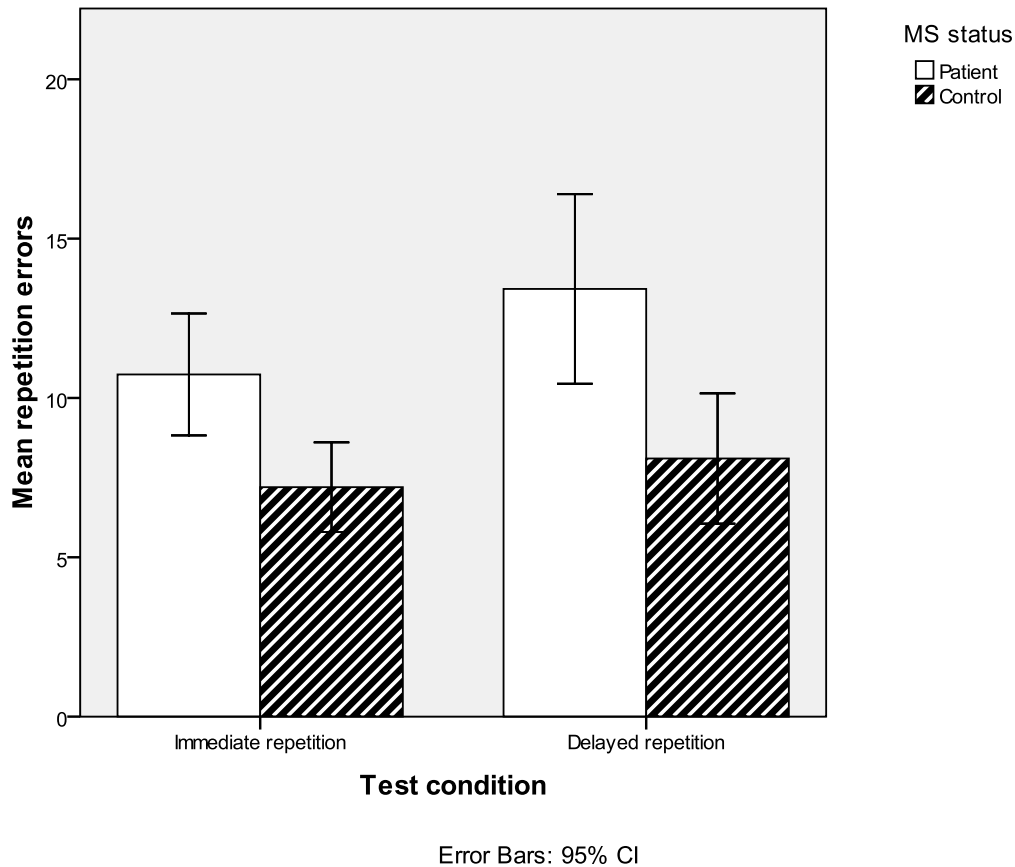


Though the condition with the 15 second delay seemed to add no difficulty for either group, the third condition – the phonological loop suppression used only with the controls – nonetheless deserved investigation. A one-way repeated measures ANOVA was performed on all the spelling data from the controls, with the three testing conditions (immediate response, delayed response, and delay with suppression) as the within-subject factor. The test revealed that the control spelling test scores were not significantly affected by delay or suppression, $F(1,19) = 1.87, ns$.

The non-word repetition test, like the spelling test, was first analysed using a factorial mixed ANOVA with MS status as the between-subjects variable, and test condition (immediate response or delay) as the within-subjects variable. In this case, there was a significant main effect of condition, $F(1,37) = 6.85, p < .05$, indicating that the 15 second delay had an effect on people's ability to repeat a non-word. There was

also a significant main effect of MS status, $F(1,37) = 12.2$, $p < .05$, which indicates that controls and patients performed differently overall. However, there was no significant interaction between MS status and test condition, $F(1,37) = 1.70$, ns , indicating that the delay did not affect the patients differently from the controls. Figure 5 illustrates the data, showing that in both conditions MS patients made more errors than controls.

Figure 5. Mean errors in non-word repetition test for patients and controls



Again, the suppression condition, applied only to the controls, was investigated. As there was a known effect of immediate versus delayed response in the above ANOVA, the area of special interest was whether or not the suppression condition was more difficult for controls than the delay condition. A paired-samples t-test was therefore performed: the number of errors made in the delay condition ($M = 8.32$, $SD = 4.37$) and the suppression condition ($M = 7.74$, $SD = 3.07$) were not significantly different, $t(18) = .561$, ns .

The test of minimal pair discrimination was also analysed with a factorial mixed ANOVA, with MS status as the between-subjects factor and test condition (word or non-word) as the within-subjects factor. There was no main effect of test condition, $F(1,37) = .187$, *ns*, which indicates that the difficulty level of the task did not change according to whether the stimuli were real words or not. There was also no interaction effect between MS status and test condition, $F(1,37) = .050$, *ns*. The variance was significantly different for both the non-word condition, $F(1,37) = 7.06$, $p < .05$, and the word condition, $F(1,37) = 8.35$, $p < .05$, compromising the between-subjects comparison. Therefore, the results were averaged for the two tests, and a Mann-Whitney U test was carried out on these new mean scores. There was no significant difference between the groups on minimal pair discrimination, $U = 128$, *ns*.

In a study that is admittedly rather heavy on statistical analysis, it was deemed prudent to combine the analyses of the rest of the tests as much as possible, to reduce the likelihood of Type I error. The two measures of the Boston Cookie Theft language sample were analysed together in one MANOVA. There was one patient who did not complete the KDT, PPT, and TTT, so these tests were analysed separately in another MANOVA to avoid having this patient's Boston Cookie Theft results omitted. The results of these two tests are displayed in Table 2 below.

Table 2. Analysis of control vs. patient performance on five tests

Test	<i>df</i>	<i>F</i>	<i>p</i>
BCT-MLU	1,38	.031	<i>ns</i>
BCT-CI	1,38	.266	<i>ns</i>
PPT	1,37	1.28	<i>ns</i>
KDT	1,37	2.81	<i>ns</i>
TTT	1,37	1.00	<i>ns</i>

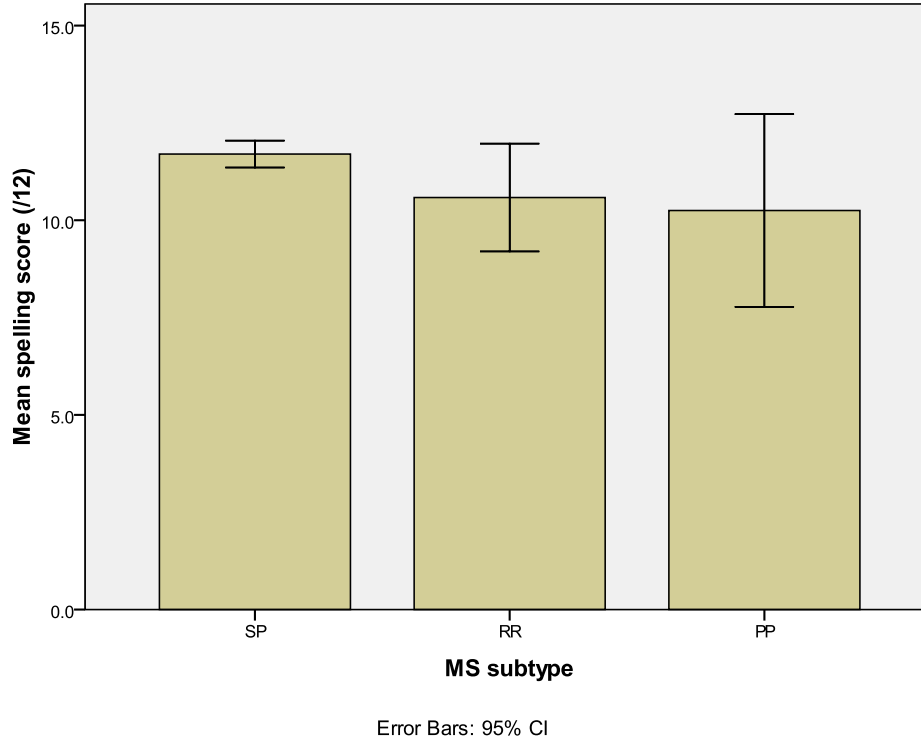
As can be seen by examining the p values, none of these tests showed a significant difference between patients and controls.

6.2 ANALYSIS BY MS SUBTYPE

In this section, the patients were divided into groups based on the MS subtype they exhibited (as reported in Table 1), and compared against each other to see if the course taken by the disease had any effect on the language tests. It is worth noting that the groups thus become very small; the PP subtype, indeed, was represented by only four members. A one-way independent ANOVA revealed that education level was not significantly different between any of the three groups, $F(2,19) = .395$, *ns*. There were, however, age differences. The mean age for the SP group was 55.4 (SD = 5.89), for the PP group, 51.28 (SD = 9.04), and for the RR group, 43.7 (SD = 9.24). A one-way ANOVA shows that there is a significant difference in ages, and *post hoc* Bonferroni tests reveal that it is only the SP and RR groups that differ significantly, ($p < .05$).

The spelling test was again analysed using a factorial mixed ANOVA, with MS subtype as the between subjects factor and test condition (immediate or delay) as the within subjects factor. There was no significant effect of test condition, $F(1,17) = .540$, *ns*, which is to be expected given the results from the group analysis. There was also no significant interaction of test condition and group, $F(1,17) = 1.43$, *ns*. The variance was not the same across the subgroups in the delay condition, $F(2,17) = 6.40$, $p < .05$, calling into doubt the between-subjects measures, so the results from the two tests were averaged for each participant, and a Kruskal-Wallis test was done, with the three subtypes as the between-groups variable. This revealed a significant effect of MS subtype on spelling, $H(2) = 6.23$, $p < .05$. Mann-Whitney U tests with Bonferroni corrections were carried out, and revealed that the difference between the scores of SP and PP patients approached significance, $U = 4.00$, $p = .017$, with the SP patients performing better than the PP patients. The difference in spelling scores between patients with the SP and RR forms of the disease was non-significant, $U = 15.0$, $p > .0167$, as was the difference between patients with RR and PP MS, $U = 11.0$, $p > .0167$. The data is illustrated in Figure 6.

Figure 6. Spelling results according to MS subtype



These results show that the delay condition had no effect on any of the patients, but that the subgroups did have different results on spelling overall, according to the Kruskal-Wallis test, though in the *post hoc* tests it did not quite achieve significance.

The non-word repetition task was also analysed using a factorial mixed ANOVA, with MS subtype as the between-subjects factor and test condition as the within-subjects factor. There was a significant main effect of test condition, $F(1,16) = 5.40$, $p < .05$, which is to be expected given the results in the group analysis. However, there was no main effect of MS subtype, $F(2,16) = .308$, *ns*, and no significant effect of interaction between subtype and test condition, $F(2,16) = .022$, *ns*, indicating that the three subtypes all performed the same on the non-word repetition task, and were not differentially affected by the delay condition.

The minimal pairs task was analysed with a factorial mixed ANOVA with MS subtype as the between-subjects factor, and test condition as the within-subjects factor. There was no significant effect of test condition, $F(1,16) = .209$, *ns*, as in the group

analysis. There was also no significant interaction effect between test condition and MS subtype, $F(2,16) = .188, ns$. Since the variance was not equal in the real word condition, the results for each patient were averaged across the two conditions, and the resulting data was analysed with a Kruskal-Wallis test with the MS subgroup as the between subjects variable. This revealed no significant difference across the MS subtype groups, $H(2) = 2.17, ns$.

The PPT did not meet the requirement of homogenous variance, so a Kruskal-Wallis test was performed on this data. It revealed that there was no significant difference between the performances of the three MS subtypes on this test, $H(2) = 3.30, ns$.

The rest of the tests were analysed with MANOVA tests, using MS subtype as a between-subjects factor. Because one patient did not complete the KDT and TTT, these tests were analysed separately in a second MANOVA. Additionally, one patient did not complete the NART, so this was analysed separately with a one-way independent ANOVA. It was undesirable to have the other test results of these two patients omitted from one large MANOVA, especially as the groups are so small. The results can be seen in Table 3.

Table 3. Analysis of potential difference between MS subtype performance on seven tests

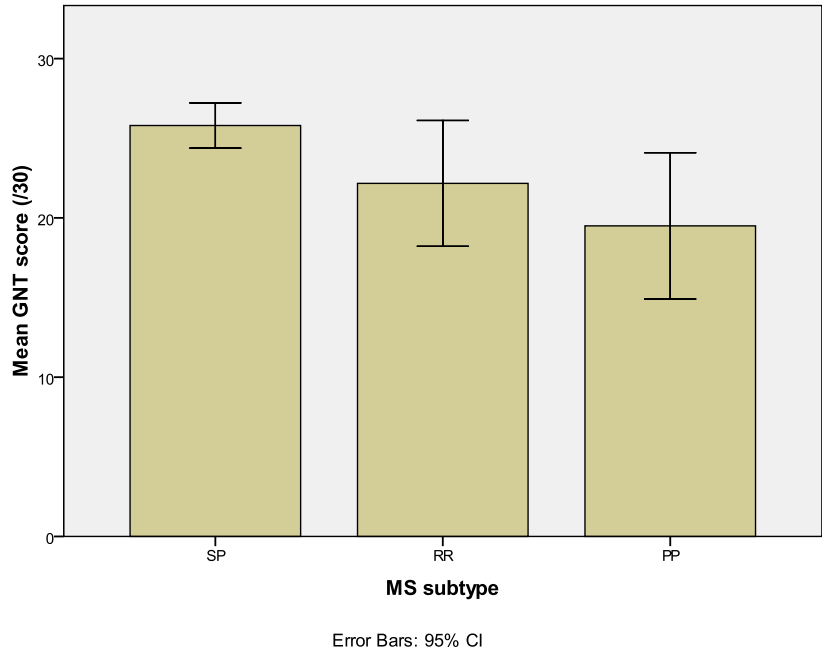
Test	df	F	p
BCT-MLU	2,17	4.05	<i>ns</i>
BCT-CI	2,17	3.40	<i>ns</i>
TROG	2,17	.052	<i>ns</i>
GNT	2,17	8.25	<.05
KDT	2,16	.567	<i>ns</i>
TTT	2,16	2.24	<i>ns</i>
NART	2,18	11.0	<.05

As can be seen, only two tests reached significance: the GNT and the NART.

Bonferroni *post hoc* tests revealed that on the GNT, the SP patients ($M = 25.8, SD = 1.99$) performed significantly better than the PP patients ($M = 19.5, SD = 2.89$), ($p < .05$).

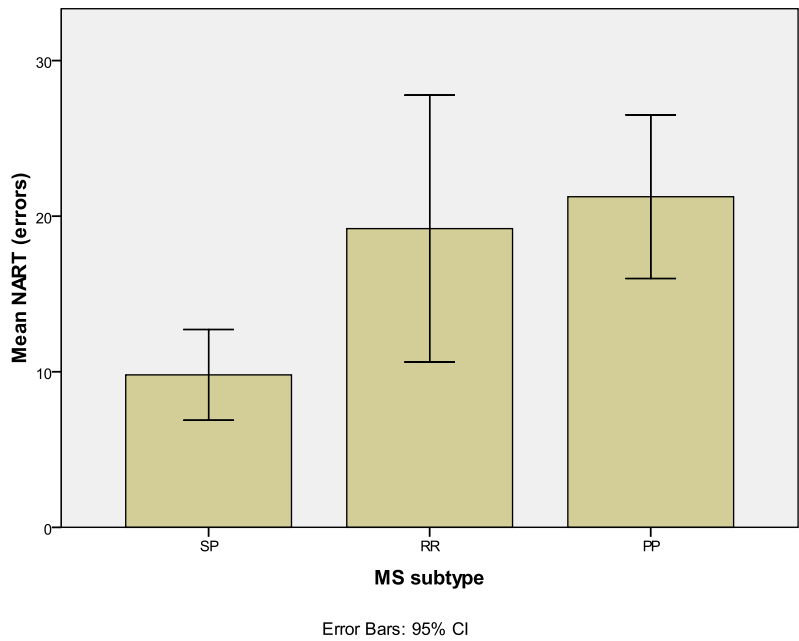
The results are shown in Figure 7.

Figure 7. GNT results according to MS subtype



On the NART, SP patients made many fewer errors ($M = 9.80$, $SD = 4.08$) than either the RR ($M = 19.2$, $SD = 6.91$) or the PP patients ($M = 21.3$, $SD = 3.30$), (both p s < .05). The results are shown in Figure 8.

Figure 8. Mean NART errors according to MS subtype



6.3 INDIVIDUAL ANALYSIS

This section compares each MS patient against the control group on every test, in order to ascertain whether language impairment in MS is due to a subgroup of severely impaired patients or a marginal impairment in the general MS population. To this end, the control mean and standard deviation were obtained for each language measure. The patient scores were compared against these, and any patients whose scores fell outside two standard deviations from the control mean were considered ‘impaired’ on that measure. For most tests, this meant the patients scored two standard deviations below the mean or more, but when the measurement was made in terms of errors, as in the NART and the non-word repetition test, impaired scores were those where the number of errors made by a patient was two standard deviations above the control mean.

There is a competing tradition in research to define ‘impaired’ as performing outside the range of the control data, on the assumption that even 5% of a control population falls outside the two standard deviation range, rendering it an inadequate measure. However, in this study it was deemed more prudent to rely on standard deviations, because just one control failure would vastly change the definition of ‘impaired’. Table 4 identifies impaired patient performances for every language measure used in this study.

The first thing to notice with these results is that there are many instances in which MS patients exhibit a marked deficit, achieving scores beyond a double standard deviation margin away from the control mean. This indicates that the significant results in section 6.1 were likely due to the results of several relatively severely impaired patients, as opposed to a slightly depressed mean for MS patients overall. Indeed, all the tests that revealed a significant difference between patients and controls – TROG, GNT, NART, spelling, and non-word repetition – can be seen here to have one-third to one-half of the patients who took them performing at an impaired level (on at least one condition, for spelling and non-word repetition).

Table 4. Individual patients impaired on language tests

Patient	MS type	SYNTAX			SEMANTICS				PHONOLOGY				WRITTEN LANGUAGE		
		Rec. TROG	Exp. MLU CI		PPT	Rec. KDT	TTT	Exp. GNT	Rec. MPwd	MPnwd	Exp. Rep.1 Rep.2		Rec. NART	Exp. Sp.1 Sp.2	
1.01	SP	n	n	n	n	n	n	n	na	na	na	na	n	n	n
1.02	SP	n	n	n	n	n	n	n	n	n	n	n	n	n	n
1.03	SP	n	n	n	n	n	n	n	n	n	n	n	n	X	n
1.04	SP	X	n	n	n	n	n	n	X	n	X	n	n	X	n
1.05	SP	X	n	n	n	n	n	n	n	n	n	n	n	n	n
1.08	SP	n	n	n	n	n	n	n	n	n	n	n	n	X	n
1.11	SP	X	n	n	n	n	n	X	n	X	X	X	n	n	n
1.15	SP	X	n	n	n	n	n	n	n	n	n	n	n	n	n
1.16	SP	n	n	n	n	X	n	n	n	X	n	X	n	n	n
1.20	SP	X	n	n	n	X	n	n	X	X	X	X	n	n	n
1.06	PP	X	n	n	n	X	n	n	n	n	n	n	X	n	n
1.07	PP	X	n	n	X	n	n	X	n	n	n	X	X	X*	X*
1.10	PP	n	n	n	n	n	n	X	n	n	n	n	X	X	n
1.18	PP	n	n	n	X	n	n	X	n	n	n	n	X	X	n
1.09	RR	X	n	n	n	n	X	X	X	X	n	X	na	X*	X*
1.12	RR	X	n	n	n	n	n	X	n	n	X*	X*	X	X*	n
1.13	RR	n	n	n	na	na	na	n	n	n	n	n	X	X	X
1.14	RR	X	n	n	n	n	n	X	n	n	n	n	X	X	X
1.17	RR	n	n	n	n	n	n	n	n	n	n	n	n	n	n
1.19	RR	n	n	n	n	n	n	X	n	n	n	n	X	n	n

n = no impairment

X = impairment

na = did not complete test

* = reported hearing problem and/or life-long spelling problem that should inform interpretation

It is also of interest that there is no definite pattern followed by all the MS patients, which reflects the variability of this disease. However, certain patterns do begin to emerge, especially intriguing divisions by MS subtype. Most strikingly, NART and GNT impairment seems linked to subtype. All four PP patients and four of five RR patients who took the NART showed impaired reading performance, but no SP patients were impaired on this test. Similarly, three of four PP and four of six RR patients were impaired on the GNT, while there was only one SP patient with impaired naming ability. The spelling tests seem to follow a similar line, with the majority of PP and RR patients impaired, though this time there were a few more impaired performances among the SP population (three of ten). The PP patient group was most affected by difficulty in the receptive semantic tests (the PPT, KDT and TTT), with three of four showing impairment

on one of these tests. In comparison, three of 10 SP and one of five RR patients were affected on one of these tests.

However, not all the tests demonstrated these patterns, and impaired performance on the TROG seems to dramatically ignore subtype distinctions in its distribution. Exactly half of each MS subtype was impaired on the TROG. Phonological impairment, as represented by impaired performance on at least one of the phonological tests, has a similar distribution, patterning equally across MS subtypes. It is of interest that six of seven of those showing impaired performance on at least one phonological test were also impaired on the TROG. These tests may not pattern according to MS subtype, but it is conceivable that they pattern together according to other factors.

6.4 CASE STUDY

As mentioned above, there was one volunteer participant who had to be excluded from the control group. This was due to a suspicion of MS-related symptoms in her medical history, combined with a family history of MS, and noticeable ‘outlier’ status on some of the language tests administered. The data was not discarded, however, because it is of interest to determine whether these tests, if sensitive to MS, might also be sensitive to early stages of MS onset. There is, of course, no way of knowing whether this participant will develop diagnosable MS. This is merely an exploratory analysis.

According to Crawford and Howell (1998), case studies should be handled with care in order to draw valid conclusions. They should be compared to control populations large enough to give a good estimate of typical performance, but when $N < 50$, the usual practice of converting the patient’s scores to z-scores is not advised (Crawford & Howell, 1998). To avoid an unacceptably large incidence of Type I error, it is best to treat the data of a single person as a set of statistics, rather than parameters. This method is robust in the face of control data that does not conform to normal parameters, and it provides an estimate of the percentage of the control population that would be expected to obtain a lower score than the case study individual (Crawford & Garthwaite, 2002). Crawford’s software, PROFLIMS.EXE (from his website, <http://www.abdn.ac.uk/~psy086/dept/>) was used to analyse the results of the case study participant in the current study. The tests

had to be divided into two batches of seven, as the software does not allow for more than 12 tests to be analysed at one time. The results are displayed in Table 5.

Table 5. Case study performance compared against controls

Test	z-score	Discrepancy from mean z	t for discrepancy	Percentage of population below	95% lower CL on percentage	95% upper CL on percentage
TROG	-3.08	-2.03	-1.98	3.19	0.25	11.10
MLU	1.10	2.49	2.43	98.70	94.05	99.97
CI	0.14	1.45	1.42	91.30	78.57	98.19
GNT	-4.19	-3.22	-3.14	0.28	0.00	1.88
PPT	-1.12	0.09	0.09	53.45	35.88	70.52
KDT	-1.92	-0.77	-0.75	23.15	10.10	40.24
TTT	0.63	1.99	1.94	96.56	88.33	99.71
Spelling 1	-6.00	-6.32	-6.16	0.00	0.00	0.00
Spelling 2	-1.32	-1.26	-1.23	11.68	3.12	25.90
*NART	3.83	4.29	4.18	99.97	99.76	100.00
MP word	0.75	0.97	0.94	82.12	65.95	93.42
MP non-word	0.73	0.95	0.92	81.54	65.24	93.06
*Repetition 1	-0.11	0.04	0.04	51.69	34.22	68.91
*Repetition 2	1.09	1.34	1.30	89.54	75.87	97.44

Percentage of population below: This figure represents the estimated percentage of the population that would obtain a lower discrepancy score

95% CLs: The last two columns provide the upper and lower confidence limits on the above percentage

Figures in the 'Percentage of population below' column marked in **bold** represent significantly impaired results

* the NART and the non-word repetition test scores are counted in terms of errors, meaning that having a large percentage of the population score below the case study represents poor performance

As we can see by inspecting the ‘Percentage of population below’ column, this participant, who was meant originally to be a control, performed very poorly on most of the tests that caused trouble for the MS population, and none that did not: she was impaired (performed worse than over 95% of the control population) on the TROG, the NART, the GNT, and one spelling test. The confidence intervals indicate that on the GNT, spelling test and NART, it is highly unlikely that she could perform within the ‘unimpaired’ range encompassed by a two standard deviation distance from the mean.

7. DISCUSSION

7.1 GROUP ANALYSIS

In this study it became apparent, first and foremost, that MS patients show impairment on a wide variety of language tests, in comparison with controls. Within all the domains tested – syntax, semantics, phonology and written language – at least one test brought back significant results in the group comparison. There were, however, some potential sources of error that need to be addressed in order to permit a more informed analysis of these results.

It is important to note is that this study, unlike many of those cited in the literature review, did not specifically contrast ‘cognitively impaired’ versus ‘cognitively preserved’ patients. Data on this point is available from the ACE-R scores, although two patients did not complete the test. Of the remainder, four scored at or below 88, the conservative cut-off point on the ACE-R, and none scored below 82, the definitive cut-off point for dementia. There was one of six RR and one of four PP patients in this intermediate zone, as well as two of 10 from the SP group. Due to the relatively even spread between the subtypes and the moderate scores, cognitive impairment is not considered a major concern in this study. However, other potential confounding factors that were not considered include duration of disease and whether or not the patient was currently experiencing an exacerbation of their symptoms, both of which could have factored into the results in some unknown capacity. There were also many issues more specific to the tests and participants of this study that need to be considered.

Firstly, as documented in Table 1, some of the patients reported having lifelong problems with spelling, which could bias the results of the spelling test against the

patients without due cause. If we look at the nature of these complaints we see that patient 1.09 reported having dyslexia. Patient 1.12 and patient 1.07 simply stated that they have “always been bad at spelling,” with patient 1.07 having been held back a grade in school, which she attributes to her spelling. It was decided to keep their results in the analysis mainly because self-report is not a very reliable means of gathering data; two of the controls also alluded to spelling difficulties before the test, but they both performed near ceiling. In the individual analysis, all these patients performed outside the double standard deviation (along with seven other patients who had reported no spelling difficulties), and their results surely contributed to the significant difference noted between patient and control spelling data. This result is therefore questionable: how much did prior spelling concerns contribute to the patients’ performance? It can be argued that the patient with dyslexia should have been excluded, at least, but his results were roughly equivalent to 1.12 and higher than 1.07, so prior to running the statistical tests it was decided to retain his data. Another key factor in the significant results obtained by this test is the excellent performance of controls: the ceiling effect was so great in the first spelling test that even one mistake was enough to put the patients outside the double standard deviation boundary. Overall, the spelling test results are perhaps not to be ascribed too much weight. Ideally, future studies would use a harder spelling test to offset ceiling effects, and ask for documentation of spelling-related diagnoses.

There are some potential sources of error in the phonological tests also. Notably, four of the patients reported hearing problems, though none of the controls did. Patient 1.01 had hearing problems so severe that she was unable to take these tests, so the inclusion of her data is not a concern. However, the data from the other three patients (1.02, 1.03 and 1.12), unlike in the spelling test, did not seem to line up with the self-reported impairment. Only one of the three performed outside the double standard deviation cut-off on any phonological tests, and this only on the non-word repetition. Of the two phonological tests, the minimal pairs test might be considered the more ‘pure’ test of hearing, since it requires the participant to recognise the smallest possible differences between words, while the repetition task includes additional productive demands. Meanwhile, there was a control who performed outside the double standard deviation level on the first minimal pair test and another who performed outside this level

on the first non-word repetition test. It is quite possible that unacknowledged hearing impairment is responsible for these results, rather than strictly phonological impairment, and without running hearing tests it is impossible to distinguish between the two. It is therefore concluded that hearing impairment, if it is to be considered in future research (which it should be), should be measured quantitatively by the examiner. Self-report is simply not reliable in this situation.

There is one other testing limitation that should be discussed, regarding the control group rather than the patients. Most of the controls were drawn from a list of willing volunteers held by the psychology department at the University of Edinburgh. The people on this list are often contacted by researchers, and several had taken part in many studies. For the experimental tests and the receptive semantic tests this was not an issue, nor for the TROG, which is predominantly used as a test for children. However, the NART and the GNT are commonly used adult tests, and a minority of the controls had come across them before, predominantly the NART. Exposure to the experimental items may have prompted them to seek out the answers on their own, or the researchers may have given the correct answers at the end of the test.

It could also be that the significant results obtained for the GNT and the NART in the present study were due to an average difference in IQ. Though education level was well-matched, it is often the case that those who volunteer for subject pools are particularly bright and motivated learners. The GNT and the NART also happen to be tests that are particularly sensitive to differences in IQ. Indeed, the GNT was designed to correlate with other measures of IQ (McKenna & Warrington 1980), and the NART is often used as a test of pre-morbid IQ in populations where there is no suspicion of specific language deterioration (Crawford et al., 2001). If the control scores were much higher than standardised norms, we might assume that this control population represented either a well-practiced or unusually intelligent subsection of society.

As it turns out, GNT scores were above the standardised average from McKenna and Warrington (1980), which were $M=22.45$, $SD=4.3$. The mean and standard deviation in the current study were $M=26.16$, $SD=1.7$. The mean is high, but still within one standard deviation from the normalized mean. The NART error scores show a much larger discrepancy from standardised norms obtained from Crawford et al. (1989), which

were $M = 20.5$, $SD = 5.8$. These figures in the current study were $M = 7.95$, $SD = 4.45$. The control results from the GNT and the NART could therefore be somewhat suspect, though the GNT is merely at the high end of normal. However, before discounting them, we should consider that although the concerns regarding the randomness of the control sample do not apply to the patient sample, nonetheless the patient results for the GNT ($M = 23.5$, $SD = 3.69$) and the NART ($M = 14.7$, $SD = 7.01$) are also better than the normalized data which is, admittedly, several decades old in both cases. It is questionable whether, as with IQ standards, the norms for these tests need to be raised periodically. In order to combat these issues it would be preferable in future research to perform IQ tests, and where possible, to choose control participants from a less compromised population.

Having established that some of the results need to be viewed cautiously, it is now appropriate to address the meaningful information that can be gleaned from the data. Several tests did not achieve significance; none of the receptive semantic (the PPT, KDT and TTT) or phonological tests (minimal pair discrimination), or expressive syntactic tests (the MLU and CI measures obtained from the Boston Cookie Theft language samples) showed significant differences between patients and controls. It is interesting that a mixture of expressive and receptive tests was necessary to locate the multiple impairments demonstrated by the MS population. This emphasizes the fact that the processing demands in the different language domains are different, and that no one overarching test can be pinpointed that tests ‘language.’

Let us examine the results from the syntactic tests. The receptive TROG forces people to contend with the most difficult aspects of syntax: relative clauses and reversible passives, sometimes in combination. In this difficult situation, controls performed much better than patients. However, in the expressive language sample, participants were allowed to choose their own words, and thus could avoid the most difficult structures. In this scenario, MS patients used subordinate clauses as often as did controls, and in both groups there were individuals who chose not to use any dependent clauses at all. This tells us that the syntactic ability is certainly not extensively impaired – no one would suggest that MS patients are not able to use language to communicate functionally on a day to day basis. However, it also tells us that given unusually complex syntactic input, it is not always possible for them to convert it correctly into real-world meaning.

In the semantic tests on the other hand, it was the expressive naming test that caused difficulty, as opposed to the receptive tests in which patients had to gauge association strengths between pictures. In this situation, they seemed to have very little difficulty putting concepts together, but when it came to accessing the lexical store and producing a word, they had less success. This indicates that the semantic association network is intact, but that the associations between meanings and their corresponding linguistic items are somewhat impaired. This is the result we would expect with a language-specific impairment, rather than a degeneration of the semantic concepts themselves.

The phonological tests were significant only in their expressive form: non-word repetition. This would seem to indicate that hearing and processing phonemes is not a problem (the scattered poor results on the minimal pairs test may well reflect a smattering of mild hearing problems, also present in the controls), but it seems that the processing involved in turning those sounds into a spoken word causes difficulty. Given that MS is known to affect motor skills, this could reflect a deficit in motor planning and execution. However, memory, also known to be affected in MS (Thornton & Raz 1997; Rao et al., 1989a), may play a role; the task requires detailed recollection of a sequence of sounds. In the condition with the 15 second delay, more errors were made by control participants as well as patients, implicating short term memory, and particularly the articulatory loop, in the skills required for the task (see Jacquemot et al., 2011). Since the patients performed worse than controls on both conditions, this may reflect a deficit in short-term memory and/or the articulatory loop, motor planning, or sequencing. However, the hypothesis outlined in section 4.1.3b concerning subvocal rehearsal is not confirmed: the patients were not affected more than the controls by the delay, and, moreover, the suppression condition did not impact the controls more than did the simple delay. It seems this test can be performed without any reliance on subvocal rehearsal.

The interpretation of the written language tests should perhaps be the most cautious. The presence of patients with written language issues across the lifespan, and of controls who performed abnormally well on the NART, together foster a high probability that the significant results obtained in both tests are not due purely to MS. However, if MS is responsible, the presence of the irregular word reading impairment represents a

difficulty either in word-to-concept mapping, resulting in a regularised pronunciation due to failure to access specific semantic associations, or else to a grapheme-to-phoneme mapping issue: no distinction was made in the marking, but some patients did produce attempts that were neither correct nor regularized, rather seeming to have been guessed from a few of the letters in the words. The spelling results would seem to support the former explanation, since the answers given were virtually all phonologically accurate.

On the whole, the results of this study reinforce and extend what has already been found in the literature reported in section 2.3. The non-word production difficulties found by Kujala (1969) are reproduced, as are the single word reading difficulties found by Jambor (1969), though the limitations of this study's NART assessment must be kept in mind. The receptive phonology task achieved no significant result, like that of Ivnik (1978). As with King (2009), language samples showed no syntactic impairment, but Grossman's (1995) subset of patients who performed poorly on receptive syntactic tests was extended to a fully significant patient-control difference on the TROG. The naming test garnered significant results, like those of Jambor (1969), Beatty et al. (1988) and Pozzilli et al., (1991), though due to the possible control-participant issues with the GNT, this is perhaps not a decisive addition to a testing domain that has thus far achieved mixed results. In a similar vein, the spelling results support Kujala's finding of an impairment, rather than Jambor's findings of intact spelling ability, though the pre-morbid spelling skills of the patients in this study may be in doubt.

There is one area in which this study disagrees with the literature, and this is on receptive semantics. Laatu et al. (1999) found that MS patients had more trouble than controls in providing descriptions of various nouns, whereas in the PPT, KDT and TTT patients showed no significant impairment associating concepts. This discrepancy is probably due to the very different nature of the tasks: the PPT, KDT and TTT involve no language whatsoever, and are simply designed to test the associative networks linking concepts, rather than words. The task used by Laatu and her colleagues required the patients to produce words, which would have taxed their word-finding abilities (which naming tests, including those in the current study, tend to indicate impairment). The difference is between a conceptual impairment and a purely linguistic one.

The main point to be aware of is that there does seem to be evidence of a wide variety of language-related impairments in people with MS. Conventionally, the cognitive abilities of people with MS are not questioned in clinical assessments, but the results from the group analysis of this study indicate that multiple aspects of language are significantly impaired in the MS population. This is important for clinical treatment because without language testing, individuals may be experiencing a reduced quality of communication when there is potential that the attentions of a speech-language therapist could help them. It is also of interest neurologically, since clear evidence is given that even with the subcortical brain damage typical of MS, multiple aspects of language can be affected, which goes against the view that only damage to the 'language centres' can impact linguistic ability.

7.2 ANALYSIS BY SUBTYPE

At this point we consider the results of the analysis comparing patients with different MS subtypes. The previous studies in the literature often did not consider this point as it relates to language, but the results of the present study would seem to indicate that there are differences between the language abilities of the three groups. It is important to note, however, that in dividing the MS population, two of the groups became very small: there were only six RR patients, and only four PP, which casts some doubt on the validity of using them as representatives of their subtype. In addition, there was a small but significant age difference between the groups: the SP patients were older than the RR patients (the PP group was somewhere in the middle, and did not differ significantly from either group). It would be an improvement to the study to use larger groups for each MS subtype, and to concentrate on a more restricted age range, but this was not feasible due to the time constraints of this project.

In the analysis by subgroup, three results showed a significant difference between the MS types. Firstly, there was a significant effect of group on spelling scores as identified in a Kruskal-Wallis test, but *post hoc* tests revealed only a difference that approached significance between the SP and PP populations, with the SP group being the better spellers. It is relevant to point out that one of the patients who self-reported having always been a bad speller (1.07) was one of four in the PP group, whereas no one in the

SP group had reported spelling difficulties, which would cast doubt on this result even if it had reached significance.

More interestingly, there were strong results with both the GNT and the NART. It was discussed above that controls might have achieved overly high results on both these tests, but in the current analysis the controls are excluded, and there is no reason to believe that any of the patients would have come across these tests before, or that they might, as a group, have an abnormal IQ distribution. Therefore, it is particularly intriguing that there were notable differences between the groups on these tests. On the GNT, the SP patients scored significantly higher than the PP patients. On the NART, the SP patients significantly outperformed both the RR and PP patients.

Even in studies that consider MS subtype, dividing out SP patients is rare: the usual comparison is between RR and 'chronic progressive' (CP) patients, and the definition of CP varies. For instance, Heaton et al. (1985) clearly defined both SP and PP patients as CP, but Lethlean & Murdoch (1997) seemed to test only PP and RR patients, describing no one that could fit the SP description. The difference between SP and PP is that SP MS develops out of an RR disease course, in which the patient suffers from acute attacks and remissions of their MS symptoms; if the patient eventually stops experiencing any remissions, they have entered into the SP category. PP patients, on the other hand, experience a disease-course which is steady, without remissions, right from the beginning. In most of the studies which compared CP with RR patients, the CP patients were found to have more cognitive deficit. In some cases no difference was found, but there were no studies that found an advantage for CP patients. In light of this, the current study's results are rather telling. There seems to be a significant difference between the two subtypes commonly grouped into CP MS, which indicates that this grouping method is inappropriate, and may have been a confounding factor in previous studies. More surprisingly, the results of the current study indicate that SP patients tended to perform better than RR patients, with a significant advantage noted in one test. It is relevant at this point to consider the effect each type of MS has on the brain, in order to better understand the subtype analysis.

On the neurological level there are several differences between the MS subtypes. RR patients are characterised by active lesions (Thorpe et al., 1996), while SP patients

tend to have large confluent brain lesions (Thompson et al., 1991). PP patients have fewer, smaller lesions, even though their condition is marked by severe pervasive disability (Thompson et al., 1991). In a study aimed at unraveling this last somewhat counter-intuitive finding, Lycklama à Nijeholt et al., (1998) performed a series of different MRI scans on 28 RR, 32 SP and 31 PP patients. They looked beyond lesions and found distinctive patterns of degeneration demarcating the three subtypes. The PP patients were more likely to have diffuse abnormalities in the brain and spinal cord than RR and SP patients, most of these areas being located in the parietal periventricular white matter. The SP group, on the other hand, was distinguished by a higher number of lesions in the brain, more ventricular enlargement and more spinal cord atrophy than either of the other groups.

As pertains to the current study, the significant differences between language performance in PP and SP patients might reflect the major underlying neurological differences between PP and SP disease courses, even though the presence of sustained physical symptoms has previously caused researchers to group these subtypes together. However, the difference observed between RR and SP patients is a little more difficult to explain. According to extent of neurological damage we would expect the RR patients to be outperforming the SP group. Other factors not measured in this study, such as disease duration and severity of disability should also be expected to favour the RR patients, as their disease is essentially the earlier stage, experienced previously by all of the SP patients. It is possible that this slight advantage of SP over RR patients, which reached significant levels only in the NART, is related to the possibility (not measured) that RR patients could be experiencing acute relapses.

More likely, however, is the possibility that the discrepancy is not related to MS at all. The only variable measured in this study that might account for the difference was age. As mentioned above, the SP group was on average about 12 years older than the RR group. Usually, age effects favour younger participants, but in this case there may have been a reverse age effect due to the nature of the test items. Being older than the RR patients, the SP group might have enjoyed a slight advantage on the NART, which admittedly contains a fair number of somewhat arcane items, more suited to an older

generation. Alternatively, it is worth remembering that the sample groups are very small, and that random variation might account for a good deal of the observed effect.

There also remains the issue of the patients who self-identified with dyslexia and spelling problems. Two of these were in the RR group (N = 6) – though one did not complete the NART – and one in the PP group (N = 4), and this may have impacted the NART as well as the spelling test, since reading and spelling are not dissimilar exercises. However, looking at the individual analysis reveals that these few patients cannot be the only factor involved.

7.3 INDIVIDUAL ANALYSIS

Table 4 displays an interesting spread of language impairment across the MS subtypes, and patterns for individual patients as well. Very clearly, reading impairment as measured by the NART is almost universal among PP and RR patients, despite the fact that only two of those who took the NART (one PP and one RR) have any claim to pre-morbid written language impairments. No SP patients were impaired on the NART. Moreover, the GNT patterns the same way, with three out of four PP and four out of six RR patients showing impaired performance on this test. Meanwhile, only one SP patient was impaired.

These results clearly demarcate the SP population as the least impaired on both naming and reading. Recalling that the NART is specifically a test of irregular words, and that access to semantics is assumed to be required in order to read them correctly, the pattern here is quite striking. RR and PP patients appear to have difficulty relating words with their meanings, and vice versa. Why this effect should spare SP patients, who by definition emerge out of RR patients is, again, unclear, but it may have to do with reverse age effects, as mentioned above. The GNT, like the NART, may favour older participants because several of the test items, such as the retort, a tool no longer used in modern chemistry labs, and the bellows, in a time where wood-burning fireplaces are less common, could favour the realm of experience of older participants. Though experiencing similar disease-related decline to the RR patients, this factor could have kept them slightly above the threshold of ‘impairment’ (which is admittedly arbitrary), while the RR patients fell just below it. This hypothesis is supported by the fact that in

the analysis by subtype there was no significant difference between the SP and RR patients on the GNT, meaning that the seemingly dramatic impairment differences in the individual analysis are actually based on only very slightly different raw scores. On the other hand, SP patients were significantly better at both the NART *and* the GNT than the PP patients, and there is no age difference between these groups. It seems that differentiating between SP and PP patients based on disease course alone is likely warranted, and the results of the individual analysis merely reinforce the clarity of this divide.

Turning to the TROG which, along with the first spelling test, identified the greatest number of impaired patients, we see a completely different pattern. Here, exactly half of each MS subtype is impaired. This emphasizes the fact that all of ‘language’ is not processed together in the brain. Double dissociations can be identified between individuals with impaired syntax but no impaired semantics (eg. 1.04, 1.05, and 1.15) and individuals having impaired semantics, but not syntax (eg. 1.10, 1.18, and 1.19). Moreover, there were patterns here on the individual level: of the seven who exhibited phonological impairment (on at least one of the four phonological tests), all but one showed impairment on the TROG. Conversely this means that six of 10 who were impaired on TROG had a phonological issue as well. This is not a perfect overlap, but it does seem to be a trend, and may implicate a problem in processing information received via auditory channels.

These findings have important clinical implications. It was mentioned above that language testing should be considered in MS due to the obvious presence of language problems. However, the individual analysis makes it plain that one simple ‘language’ test is inappropriate for MS patients. The variability of the disease makes it such that different subsets of patients are impaired on different components of language. Even if they were all to be administered a cognitive test battery with a language component (ie. naming or fluency), this would miss the syntactic issues measured by the TROG, as well as potential problems with phonology and written language. Notably, on naming tests most SP patients would be considered linguistically preserved, but the present study has demonstrated that they are as likely as anyone else to exhibit a syntactic impairment.

7.4 CASE STUDY

The case study was an unplanned addition to this study, but allows for the framing of some very interesting questions. The participant in question has a sibling with MS and has a suspicion of MS in her medical history. She was at the young end of the scale for this study (age 28.8) and had relatively fewer years of education (10). Her performance on the language tests, as illustrated in section 6.4, was markedly poor, and followed similar patterns to the MS patients. She was impaired on the TROG, the GNT, the NART, and one spelling test. This represents almost the full range of tests that achieved significant results in the group analysis between the MS patients and the controls (the only one missing was the non-word repetition), indicating that any impairments she has might well be drawn from the same source as those displayed by the patients.

It is worth emphasizing again that there is no certainty that this participant will develop MS. She does not currently have MS, and a family history may increase her risk, but not drastically. It is possible that her poor performance is due to random personal factors, perhaps including a different neurological condition. Alternately, it could reflect a combination of age effects (as discussed previously, there is a possibility that the GNT and the NART favour older participants) and her education level. However, if this participant is indeed in the early stages of MS, that means that these language tests could be sensitive early indicators of the disease. Clinically Isolated Syndrome is the name for symptoms that do not amount to a diagnosis of MS, but that are suggestive of the disease. The results of this case study suggest that perhaps language should be considered as one of the possible indicators of emergent MS.

It is not unknown for some form of language degeneration to be an early warning sign of neurological degradation. As mentioned in section 2.2, foreign language syndrome can emerge as the first symptom of organic brain disease (Luzzi et al., 2008). Isolated difficulty with word finding is also a common precursor to neurodegenerative conditions, especially the progressive aphasia (Rohrer et al., 2007). Additionally, there are many conditions that can present with similar symptoms to MS, and occasionally render diagnosis according to clinical and laboratory guidelines unclear (McDonald et al., 2001). Therefore, the possibility that the initial onset of MS may include language effects is interesting, and should be pursued with an eye to enhancing diagnostic procedures.

8. CONCLUSION

The findings of this study clearly reinforce previous reports from the literature, giving evidence for language impairment in MS. The most sensitive tests, identifying the greatest number of MS patients as being outside the normal range of control scores, are the GNT, TROG, non-word repetition, NART and spelling, though these last two were particularly problematic in the current study and would benefit from further investigation.

The analysis by subtype should also be treated cautiously due to the small numbers in the RR and PP groups, but there was certainly a suggestion of distinctive language profiles associated with MS subtype. The naming test and the reading test were sensitive in this regard, revealing an advantage of SP patients over PP patients. This finding indicates that the practice in previous studies of combining the two progressive-type MS groups is inappropriate, and is probably reflective of the different levels and types of brain damage observed in these patients. A significant advantage for the SP patients over the RR patients was also found in the NART, which is somewhat mysterious given the brain-damage profiles of each subtype; the result might be ascribable to reverse age effects, but it is doubtless a question for further research.

In the individual analysis, some patients were identified who had no impairment with language, while a sizable subset was shown to have one or more areas of significant difficulty. There was also an intriguing observable double dissociation between the language domains: naming and reading were almost exclusively impaired in the RR and PP subtypes, but syntactic difficulties affected exactly half of each subtype. Individuals can be identified having impairment in one of either semantics or syntax, but not the other. This variability is not surprising in the face of the notoriously variable disease course of MS, and merely reinforces the need to bear in mind the linguistic perspective that language is not an indivisible module. There are different domains that can be differentially affected by brain degeneration, and it is important to use multiple tests to truly evaluate linguistic ability.

The overall results indicate that MS patients should be investigated for language difficulties clinically, as they may well benefit from speech and language therapy. Furthermore, evidence of language impairment following from subcortical damage is of interest to the field of linguistics and neurology, and might prove fruitful for investigators

interested in the neural substrates of language. It is also suggested by the case study that for some people, language impairment could be an early warning sign of MS, and this possibility is worthy of further investigation.

9. REFERENCES

- Achiron, A., Ziv, I., Djaldetti, R., Goldberg, H., Kuritzky, A. & Melamed, E. (1992). Aphasia in multiple sclerosis: Clinical and radiologic correlations. *Neurology*, *42*, 2195-2197.
- Alexander, M.P., Naeser, M.A. & Palumbo, C.L. (1987). Correlations of subcortical CT lesion sites and aphasia profiles. *Brain*, *110*, 961-991.
- Arnett, P.A., Smith, M.M., Barwick, F.H., Benedict, R.H.B. & Ahlstrom, B.P. (2008). Oralmotor slowing in multiple sclerosis: Relationship to neuropsychological tasks requiring an oral response. *Journal of the International Neuropsychological Society*, *14*, 454-462.
- Bak T., Hodges J.R. (2003). 'Kissing and Dancing'—a test to distinguish the lexical and conceptual contributions to noun/verb and action/object dissociation. Preliminary results in patients with frontotemporal dementia. *Journal of Neurolinguistics*, *16*, 169-81.
- Beatty, W.W. Goodkin, D.E., Monson, N., Beatty, P.A. & Hersgaard, D. (1988). Anterograde and retrograde amnesia in patients with chronic progressive multiple sclerosis. *Archives of Neurology*, *45*, 611-619.
- Bishop, D. V. M. (1983). *Test for Reception of Grammar*. Published by the author at Department of Psychology, University of Manchester.
- Blackwood, H.D., La Pointe, L.L., Holtzapple, P., Pohlman, K., & Graham, L.F. (1991). Lexical-semantic abilities of individuals with multiple sclerosis and aphasia. In R. E. Prescott (Ed.), *Clinical Aphasiology* (pp. 121-127). Austin, TX: Pro-Ed
- Callanan, M.M., Logsdail, S.J., Ron, M.A. & Warrington, E.K. (1989). Cognitive impairments in patients with clinically isolated lesions of the type seen in multiple sclerosis. *Brain*, *112*, 361-374.
- Crawford, J.R., Deary, I.J., Starr, J. & Whalley, L.J. (2001). The NART as an index of prior intellectual functioning : a retrospective validity study covering a 66-year interval. *Psychological Medicine*, *31*, 451-458.
- Crawford, J.R. & Garthwaite, P.H. (2002). Investigation of the single case in neuropsychology: Confidence limits on the abnormality of test scores and test score differences. *Neuropsychologia*, *40*, 1196-1208.
- Crawford, J.R. & Howell, D.C. (1998). Comparing an Individual's Test Score Against Norms Derived from Small Samples. *The Clinical Neuropsychologist*, *12*, 482-486.

- Crawford, J.R., Parker, D.M., Stewart, L.E., Besson, J.A.O. & De Lacey, G. (1989). Prediction of WAIS IQ with the National Adult Reading Test: Cross-validation and extension. *British Journal of Clinical Psychology*, 28, 267-273.
- Curtiss, S., Fromkin, V., Krashen, S., Rigler, D. & Rigler, M. (1974). The Linguistic Development of Genie. *Language*, 50, 528-554.
- Damasio, A.R., Damasio, H., Rizzo, M., Varney, N. & Gersh, F. (1982). Aphasia with nonhemorrhagic lesions in the basal ganglia and internal capsule. *Archives of Neurology*, 39, 15-20.
- Dapretto, M., & Bookheimer, S. Y. (1999). Form and Content: Dissociating Syntax and Semantics in Sentence Comprehension. *Neuron*, 24, 427-432.
- Day, T.J., Fisher, A.G. & Mastaglia, F.L. (1987). Alexia with agraphia in multiple sclerosis. *Journal of the Neurological Sciences*, 78, 343-348.
- Eubank, L., & Gregg, K. (1999). Critical periods and (second) language acquisition: Divide et impera. In D. Birdsong (Ed.), *Second language acquisition and the Critical Period Hypothesis* (pp. 65–99). Mahwah, NJ: Erlbaum.
- Flege, J. E. (1999). Age of learning and second-language speech. In D. Birdsong (Ed.), *Second language acquisition and the Critical Period Hypothesis* (pp. 101–131). Mahwah, NJ: Erlbaum.
- Franklin, G.M., Heaton, R.K., Nelson, L.M., Filley, C.M. & Seibert, C. (1988). Correlation of neuropsychological and MRI findings in chronic/progressive multiple sclerosis. *Neurology*, 38, 1826-1829.
- Friedman, J.H., Brem, H. & Mayeux, R. (1983). Global aphasia in multiple sclerosis. *Annals of Neurology*, 13, 222-223.
- Friend, K.B., Rabin, B.M., Groninger, L., Deluty, R.H., Bever, C. & Grattan, L. (1999). Language functions in people with multiple sclerosis. *The Clinical Neuropsychologist*, 13, 78-94.
- Goodglass, H., & Kaplan, E. (1983). The assessment of aphasia and related disorders. 2nd ed. Lea & Febiger, Philadelphia
- Grossman, M., Robinson, K.M., Onishi, K., Thompson, H., Cohen, J. & D'Esposito, M. (1995). Sentence comprehension in multiple sclerosis. *Acta Neurologica Scandinavica*, 92, 324-331.

- Heaton, R.K., Nelson, L.M., Thompson, D.S., Burks, J.S. & Franklin, G.M. (1985). Neuropsychological findings in relapsing-remitting and chronic-progressive multiple sclerosis. *Journal of Consulting and Clinical Psychology, 53*, 103-110.
- Henry, J.B. & Beatty, W.W. (2006). Verbal fluency deficits in multiple sclerosis. *Neuropsychologia, 44*, 1166-1174.
- Hopp, H. (2009). The syntax-discourse interface in near-native L2 acquisition: Off-line and online performance. *Bilingualism: Language and Cognition, 12*, 463-483.
- Howard, D. & Patterson, K. (1992). *The Pyramid and Palm Trees Test: A Test of Semantic Access From Words and Pictures*. Bury St. Edmunds: Thames Valley Test Company.
- Ivnik, R.J. (1978). Neuropsychological stability in multiple sclerosis. *Journal of Consulting and Clinical Psychology, 46*, 913-923.
- Jacquemot, C., Dupoux, E. & Bachoud-Lévi, A-C. (2011). Is the word-length effect linked to subvocal rehearsal? *Cortex, 47*, 484-493.
- Jambor, K.L. (1969). Cognitive functioning in multiple sclerosis. *The British Journal of Psychiatry, 115*, 765-775.
- Jennekens-Schinkel, A., Lanser, J. B., Van der Velde, E. A., & Saunders, E. A. (1990). Performances of multiple sclerosis patients in task requiring language and visuoconstruction: Assessment of outpatients in quiescent disease stages. *Journal of Neurological Sciences, 95*, 89-103.
- Kan, I.P., Biran, I., Thompson-Schill, S.L. & Chatterjee, A. (2006). Letter selection and letter assembly in acquired dysgraphia. *Cognitive Behavioural Neurology, 19*, 225-236.
- Kennedy, M. & Murdoch, B.E. (1989). Speech and language disorders subsequent to subcortical vascular lesions. *Aphasiology, 3*, 221-247.
- King, K.A. (2009). Syntactic complexity in persons with multiple sclerosis. Unpublished master's dissertation. Brigham Young University.
- Kujala, P., Portin, R. & Ruutinen, J. (1996). Language functions in incipient cognitive decline in multiple sclerosis. *Journal of the Neurological Sciences, 141*, 79-86.
- Laakso, K., Bruunegård, K., Hartelius, L. & Ahlsén, E. (2000). Assessing high-level language in individuals with multiple sclerosis: a pilot study. *Clinical Linguistics and Phonetics, 14*, 329-349.

- Laatu, S., Hamalainen, P., Revonsuo, A., Portin, R. & Ruutiainen, J. (1999). Semantic memory deficit in multiple sclerosis; impaired understanding of conceptual meanings. *Journal of the Neurological Sciences*, 162, 152-161.
- Lethlean, J. B. & Murdoch, B. E. (1997). Performance of subjects with multiple sclerosis on test of high-level language. *Aphasiology*, 11, 39-57.
- Long, M. (1990). Maturation constraints on language development. *Studies in Second Language Acquisition*, 12, 251-285.
- Lucchelli, F. & Papagno, C. (2005). Is slowly progressive anarthria a “pure motor-speech disorder? Evidence from writing performance. *Neurocase*, 11, 234-241.
- Lucchinetti, C., Brück, W., Parisi, J., Scheithauer, B., Rodriguez, M. & Lassman, H. (2000). Heterogeneity of multiple sclerosis lesions: Implications for the pathogenesis of demyelination. *Annals of Neurology*, 47, 707-717.
- Luzzi, S., Viticchi, G., Piccirilli, M., Fabi, K., Pesallacchia, M., Bartolini, M., Provinciali, L. & Snowden, J.S. (2008). Foreign accent syndrome as the initial sign of primary progressive aphasia. *Journal of Neurology, Neurosurgery and Psychiatry*, 79, 79-81.
- Lycklama à Nijeholt, G.J., van Walderveen, M.A.A., Castelijns, J.A., van Waesberghe, J.H.T.M., Poman, C., Scheltens, P., Rosier, P.F.W.M., Jongen, P.J.H. & Barkhof, F. (1998). Brain and spinal cord abnormalities in multiple sclerosis: Correlation between MRI parameters, clinical subtypes and symptoms. *Brain*, 121, 687-697.
- McAllister, W.S., Belman, A.L., Milazzo, M., Weisbrot, D.M., Christodoulou, C., Scherl, W.F., Preston, T.E., Cianciulli, C. & Krupp, L.B. (2005). Cognitive functioning in children and adolescents with multiple sclerosis. *Neurology*, 64, 1422-1425.
- McDonald WI, Compston A, Edan G, et al. (2001). Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. *Annals of Neurology*, 50, 121-127.
- McKenna, P. & Warrington, E.K. (1980). Testing for nominal dysphasia. *Journal of Neurology, Neurosurgery, and Psychiatry*, 43, 781-788.
- McKenna P. & Warrington, E.K. (1983). *The Graded Naming Test*. Windsor: NFER-Nelson.
- Mioshi E, Dawson K, Mitchell J, Arnold R, Hodges JR. (2006). The Addenbrooke's Cognitive Examination Revised (ACE-R): a brief cognitive test battery for dementia screening. *International Journal of Geriatric Psychiatry*, 21, 1078-85.

- Montrul, S., & Slabakova, R. (2000). Competence similarities between native and near native speakers: An investigation of the preterit/imperfect contrast in Spanish. *Studies in Second Language Acquisition*, 25, 351-398.
- Moro, A., Tettamanti, M., Perani, D., Donati, C., Cappa, S.F. & Fazio, F. (2001). *NeuroImage*, 13, 110-118.
- Mummery, C.J., Patterson, K., Wise, R.J.S., Vandenberg, R., Price, C.J. & Hodges, J.R. (1999). Disrupted temporal lobe connections in semantic dementia. *Brain*, 122, 61-73.
- Murdoch, B.E. & Theodoras, D.G (Eds). (2000). *Speech and Language Disorders in Multiple Sclerosis*. London: Whurr.
- Olerup, O., Hillert, J., Fredrikson, S., Olsson, T., Kam-Hansen, S., Moller, E., Carlsson, B. & Wallin, J. (1989). Primarily chronic progressive and relapsing/remitting multiple sclerosis: two immunogenetically distinct disease entities. *Proceedings of the National Academy of Sciences of the USA*, 86, 7113-7117.
- Olmos-Lau, N., Ginsberg, M.D. & Gellar, J.B. (1977). Aphasia in multiple sclerosis. *Neurology*, 27, 623-626.
- Peyser, J.M., Edwards, K.R., Poser, C.M. & Filskov, S.B. (1980). Cognitive function in patients with multiple sclerosis. *Archives of Neurology*, 37, 577-579.
- Pinker, S. (1994). *The Language Instinct*. New York: Harper Collins.
- Polman, C. H., Reingold, S. C., Edan, G., Filippi, M., Hartung, H.-P., Kappos, L., Lublin, F. D., et al. (2005). Diagnostic criteria for multiple sclerosis: 2005 revisions to the “McDonald Criteria”. *Annals of Neurology*, 58, 840-846.
- Pozzilli, C., Passafiume, D., Bernardi, S., Pantano, P., Incoccia, C., Bastianello, S., Bozzao, L., Lezi, G.L. & Fieschi, C. (1991). SPECT, MRI and cognitive functions in multiple sclerosis. *Journal of Neurology, Neurosurgery and Psychiatry*, 54, 110-115.
- Rao, S.M., Leo, G.J. & St. Aubin-Faubert, P. (1989). On the nature of memory disturbance in multiple sclerosis. *Journal of Clinical and Experimental Neuropsychology*, 11, 699-712.
- Rao, S.M., St. Aubin-Faubert, P. & Leo, G.J. (1989). Information processing speed in patients with multiple sclerosis. *Journal of Clinical and Experimental Neuropsychology*, 11, 471-477.

- Robin, D.A. & Schienberg, S. (1990). Subcortical lesions and aphasia. *Journal of Speech and Hearing Disorders*, 55, 90-100.
- Rohrer, J.D., Knight, W.D., Warren, J.E., Fox, N.C., Rossor, M.N. & Warren, J.D. (2007). Word-finding difficulty: a clinical analysis of the progressive aphasias. *Brain*, 131, 8-38.
- Ron, M.A., Callanan, M.M. & Warrington, E.K. (1991). Cognitive abnormalities in multiple sclerosis: a psychometric and MRI study. *Psychological Medicine*, 21, 59-68.
- Sasanuma, S. & Monoi, H. (1975). The syndrome of Gogi (word-meaning) aphasia: Selective impairment of Kanji processing. *Neurology*, 25, 627-632.
- Schneider, P., Dubé, R. V., & Hayward, D. (2005). The Edmonton Narrative Norms Instrument. Retrieved 2 Aug 2011 from University of Alberta Faculty of Rehabilitation Medicine website: <http://www.rehabmed.ualberta.ca/spa/enni>.
- Seliger, H. (1978). Implications of a multiple critical period hypothesis for second language learning. In W. Ritchie (Ed.), *Second language acquisition research: Issues and implications* (pp. 11–25). San Diego, CA: Academic Press.
- Snow, C.E. & Hoefnagel-Hohle, M. (1978). The Critical Period for Language Acquisition: Evidence from Second Language Learning. *Child Development*, 49, 1114-1128.
- Thompson, A.J., Kermode, A.G., Wicks, D., MacManus, D.G., Kendall, B.E., Kingsley, D.P., et al. (1991). Major differences in the dynamics of primary and secondary progressive multiple sclerosis. *Annals of Neurology*, 29, 53–62.
- Thompson, A.J., Polman, C.H., Miller, D.H., McDonald, W.I., Brochet, B., Filippi, M., Montalban, X. & De Sa, J. (1997). Primary progressive multiple sclerosis. *Brain*, 120, 1085-1096.
- Thornton, A.E. & Raz, N. (1997). Memory impairment in multiple sclerosis: A quantitative review. *Neuropsychology*, 11, 357-366.
- Thorpe, J.W., Kidd, D., Moseley, I.F., Kendall, B.E., Thompson, A.J., MacManus, D.G., et al. (1996). Serial gadolinium-enhanced MRI of the brain and spinal cord in early relapsing–remitting multiple sclerosis. *Neurology*, 46, 373–378.

- van der Hulst, E-J., Abrahams, S. & Bak, T.H. (2010, May). *Semantic deficits in amyotrophic lateral sclerosis: An impairment of action knowledge, or action sequencing?* Poster session presented at the biennial meeting of the World Federation of Neurology Research Group on Aphasia and Cognitive Disorders, Istanbul.
- Wallace, G.L., & Holmes, S. (1993). Cognitive-linguistic assessment of individuals with multiple sclerosis. *Archives of Physical Medicine & Rehabilitation*, 74, 637-643.
- White, L., & Genesee, F. (1996). How native is near-native? The issue of ultimate Attainment in adult second language acquisition. *Second Language Research*, 12, 238-26.
- Wishart, H. & Sharpe, B. (1997). Neuropsychological aspects of multiple sclerosis: A quantitative review. *Journal of Clinical and Experimental Neuropsychology*, 19, 810-824.
- Zurif, E., Swinney, D., Prather, P., Solomon, J. & Bushell, C. (1993). An On-Line Analysis of Syntactic Processing in Broca's and Wernicke's Aphasia. *Brain and Language*, 45, 448-464.