ASSESSING HIGH-LEVEL LANGUAGE IN INDIVIDUALS WITH MULTIPLE SCLEROSIS: A PILOT STUDY

Karin Brunnegård, Katja Laakso, Lena Hartelius, Elisabeth Ahlsén***

*Dept. of Logopedics and Phoniatrics, **Dept. of Linguistics, Göteborg University, Göteborg, Sweden

ABSTRACT

This study describes the development of a test battery to assess high-level language function in Swedish and a description of the test performances of a group of 9 individuals with multiple sclerosis (MS). The test battery included tasks such as repetition of long sentences, understanding of complicated logico-grammatical sentences, naming famous people, resolving ambiguities, recreating sentences, understanding metaphors, making inferences, defining words. The MS group included individuals with self-reported language problems as well as individuals without any such problems. Their performances were compared to a group of 7 control subjects with a Kruskal-Wallis one-way ANOVA which indicated significantly different total mean scores. Post hoc analysis with Mann-Whitney U-tests revealed that the group with self-reported language problems had significantly lower mean scores when compared to control subjects and to MS subjects without self-reported language problems. None of the language difficulties were detected by a standard aphasia test.

1. INTRODUCTION

Individuals recovering from the effects of stroke or traumatic brain injury may still experience language problems. There is also increasing evidence that individuals with certain progressive neurological disorders such as Parkinson's disease (1) and multiple sclerosis (2, 3-5) may not only have speech production deficits, but also experience language deficits. These problems may be very discrete but nevertheless interfere greatly with the professional and/or social lives of these individuals, that is their functional capacity. Individuals suffering from subtle deficits in verbal expression may exhibit difficulties in conveying their thoughts and needs to others (6). Since these language problems usually neither are detected by the speech-language pathologist nor the neuropshychologist, the clinical management of these individuals is often insufficient. The main reason for this is the lack of instruments to assess these functions and because the more obvious clinical syndromes are not present (6). The so called high-level language (HLL) deficits may also be a sensitive indicator of brain pathology or a marker for certain degenerative dementias (7, 8). If subtle language deficits are detected, language intervention can begin at an early stage. There is thus a need for a more complex and comprehensive assessment tool in Swedish sensitive enough to detect the subtle language deficits that are often neglected, for example in the MS population. The literature on the topic indicates the lack of such a tool in other languages as well.

Multiple sclerosis (MS) is a chronic progressive demyelinating disease which predominantly affects the white matter of the central nervous system. Dysarthria is a commonly reported finding in this population (9, 10) but aphasia has been reported as rare (11). However, recent research has demonstrated the existence of high level language dysfunction in MS (3, 5). These studies show that individuals with MS produced significantly more naming errors than control subjects and that they had difficulties understanding ambiguous sentences and metaphoric expressions, making inferences, and recreating sentences. They also exhibited relatively poor performance on vocabulary and semantic tasks compared to control subjects. Impaired sentence comprehension has also been reported (2).

The aims of the present study were to develop a test battery to assess high-level language functions in Swedish, and use this test battery with a group of individuals with multiple sclerosis and matched controls.

2. METHODS

2.1 THE TEST BATTERY

The assessment material includes ten types of tasks that will assess subtle language deficits according to current research. Prior to the development of the test, a number of considerations were made which are summarised below:

- Not more than one hour to administer
- Variability in subtest character
- Subtests not heavily dependent on visual acuity
- Time restrictions on some subtests
- Qualitative information accounted for in scoring
- Subtests carefully selected with respect to linguistic/cognitive factors
- High item difficulty
- Approximately ten items on each subtest
- Clear and precise instructions for administration and scoring

• Practice item before each subtest

The ten subtests were:

1. Repetition of long sentences (15-28 syllables).

Repetition of relatively short sentences is a common task in aphasia testing and considered the most elementary form of expressive language. Example: *“Continue straight ahead and then turn to the right in the intersection by the store.”*

2. Comprehension of logico-grammatical sentences

This task assesses the comprehension of complex grammatical constructions, for example passive voice, inverted sentences, instructions in several steps and double negations. Example: *“Tell me if the mother’s sister and the sister’s mother are two individuals or one and the same.”*

3. Naming famous people

As proper names have been found to be more vulnerable to anomia, we have chosen to include naming famous people instead of naming common nouns.

4. Comprehension of ambiguous sentences

To comprehend ambiguous sentences (syntactical and lexical) a person needs to make multiple interpretations of a sentence. Example: *“He likes Malin more than Robert.”*

5. Word fluency

Difficulties with word fluency tasks can be a sensitive indicator of word finding difficulties in some patients whose tested naming is otherwise unimpaired. The phonemic categories used were words beginning with the letter S and T and the semantic category used was animals. Time limit was one minute for each task.

6. Recreating sentences

This task assesses the ability to plan and formulate speech acts that are semantically and syntactically correct sentences. Example: *Context given: at the restaurant. Words given: or, pie, have.*

7. Comprehension of metaphors

Comprehension of figurative language includes the ability to understand abstract language, to understand the features of a concept and apply it to something else, that is categorisation and generalisation. Test items included both common and poetic expression.

8. Making inferences

The subtest evaluates the ability to make inferences and understand implied relationships, depending on recognition and recovery of missing links in the underlying causal chain of a script of a text. The subtest consists of four types of inference: pragmatic, dialogue, narrative and emotional.

9. Similarities/Dissimilarities

This task evaluates verbal concept formation. To solve this task the subjects need access to semantic information about categories and subcategories. Example: *“What is the difference between a train and a tram?”*

10. Word definitions

The task requires the ability to recognise critical semantic attributes of words. The test score reflects both the extent of recall vocabulary and the effectiveness of speaking vocabulary.

Interjudge reliability was measured between the consensus judgement of two examiners and the assessment of two external judges who made independent evaluations. The external judges’ agreement with the consensus judgement was 84% and 87% respectively.

To exclude the possibility that the subjects had a language disorder of the type captured by a conventional aphasia examination, we first tested all subjects with a widely used neurolinguistic aphasia test called A-ning.

A specialised neuropsychologist made a neuropsychological screening on all multiple sclerosis subjects. This was made in order to determine of the performance on these linguistic measures may be explained by fatigue, low attention span or other neuropsychological abilities detected in a neuropsychological screening test. Assessed functions were: verbal memory span, visual memory, attention span/working memory, tempo/attention, visual perception/tempo.

2.2 SUBJECTS

Subject inclusion criteria were (a) a definite diagnosis of MS, (b) a chronic progressive stage of the disease, (c) Swedish as native language, (d) no known brain pathology except for MS, (e) maximum age of 75 years, (f) none or minimal dysarthria, (g) good vision and hearing abilities. The subjects with MS included in our study were 9 women with a mean age of 64.6 (range=51-74, SD=6.4). The control subjects were pair matched for age, gender and education level.

3. RESULTS

All subjects included were able to complete language assessments. No subjects had language difficulties of aphasia type as tested with the neurolinguistic aphasia test A-ning.

Table 1. Mean percent correct on the test battery for MS group and control group.

MS	1	2	3	4	5
Subjects	**79.4**	**87.9**	**86.4**	**97.7**	**90.3**
	6	7	8	9	*Group mean*
	81.7	**45.6**	**79.2**	**83.0**	**81.2**
- -					
Matched	1	2	3	4	5
controls	**82.9**	**83.4**	**82.2**	**95.6**	**90.7**
	6	7	8	9	*Group mean*
	92.6	**62.8**	**62.1**	**87.9**	**82.2**

The scores on the test were relatively low for both groups, although a few subjects had high scores (Table 1). Means for the groups were very similar. No correlation between mean results on the test and age or education was found, when compared with a Pearson correlation measure.

The MS group was divided into two groups: Those with self-reported language problems, and those who considered themselves free of such problems. MS subjects 1, 7, 8 and 9 fell in the former group and 2, 3, 4, 5 and 6 in the latter. What they reported was mainly word finding difficulties of different kinds. The subjects with self-reported language problems were also identified by the neuropsychological screening test having defective performances of variable degrees. The main finding was below average functioning of working memory.

The two MS groups' and the control group's total mean scores were compared. Uneven group sizes necessitated the use of non-parametrical statistical procedures in the form of Kruskal-Wallis one-way ANOVAs. When using the Kruskal-Wallis one-way ANOVA initially no significant differences were recorded across the three groups. However, the control group was screened for outliers, and two subjects (number 7 and 8) and their scores were removed from the dataset. Their scores differed more than 1.5 SD from the total mean score. Of these two control subjects, one subject expressed extreme nervousness during testing, the other one did not complete all subtests due to fatigue. Kruskal-Wallis one-way ANOVA was used once more to compare the three groups. Significant differences in mean scores were found ($H=6.4958$, $p=0.0389$). Post hoc analysis using Mann-Whitney U-Tests indicated significant differences among the three subject groups. MS subjects with self-reported language problems had significantly lower total mean scores than MS subject with no reported language problems ($z=-2.2045$, 2-tailed $p=0.0275$) and control subjects ($z=-2.2678$, 2-tailed $p=0.0233$). No significant differences were found between the results of the group with no reported language problems and the control subjects.

Comparing the mean raw scores of each subtest for the MS group and the control group with a t-test for paired samples revealed no statistically significant differences in their performances. The t-test was chosen because this would increase the power compared to using a non-parametric test. However, the same comparison was made with a Wilcoxon non-parametric test. The results were very similar and revealed no significant differences.

4. DISCUSSION

It is our hope that the findings of the present study will be used in further research on developing a test of high-level language in Swedish, for a variety of patient groups suffering from subtle language deficits. We do not see the test battery used in this study as an assessment material for a specific population. Lesion sites could be focal/diffuse, cortical/subcortical or in the right or the left hemisphere. Lesions could be of different aetiology. Cross-linguistic research in the area of high-level language could be rewarding, concerning different populations and from different language groups.

Overall the administration of the test battery worked well. The dimension of the test battery was such that it was administered in one hour or less to most subjects, which was within the time limits set up. Instructions seemed to be clear and seldom needed to be repeated. Instructions for scoring were mostly sufficiently detailed. Some alterations to the test battery are suggested. Some time restrictions used should be revised, most likely time limits should be shortened in order to discover deficiencies more easily. Furthermore, test items should all be arranged according to item difficulty, in order to be able to discontinue any single subtest if the subject fails to respond correctly on a number of items after several attempts. In addition to this, subtests should be ordered to give a variation in difficulty and character.

Subjects with multiple sclerosis and control subjects performed in a nearly equivalent fashion on the test battery, when comparing the matched pairs and groups (MS vs. control group) on mean results on subtests, as well as when comparing the groups on subtest items. These results contrast with those in some previous studies, which have found differences, in performance on language tasks (2-5). There are several factors that may have induced these results, of which the low number of subjects in the present study is one main factor. Previous studies have mainly been group studies with larger samples (3, 5). Hence, the presence of individuals with preserved language functioning would not have influenced the results as much as in a study with a lower number of subjects. The prevalence of linguistic difficulties amongst subjects may also depend on a variety of variables such as duration of disease (12). Other factors that may have elicited these results are that the control subjects may not have been representative of a normal population and the possibility that the test battery does not differentiate between individual with subtle language deficits and those without such difficulties. Characteristics in both MS subjects and control subjects as well as the test battery and inclusion criteria are discussed in the following, as variables possibly affecting the outcome. Our belief is that the low scores of the control group were uncharacteristic for the normal population. This may be the main reason why no group differences were found. The findings in this study should therefore not be interpreted as supporting the thesis that the multiple sclerosis populations do not exhibit language deficits.

The control subjects tended to be more nervous during testing and very eager to perform well. The multiple sclerosis subjects had all participated in several studies before and appeared more relaxed during testing.

Self-reported language problems seem to be related to lower total score on the present test. This is indicated by the four multiple sclerosis subjects who reported word finding difficulties. Kruskal-Wallis one-way ANOVAs showed significant differences between the groups. Post-hoc analysis with Mann-Whitney U-Test revealed that the MS group with self-reported language problems produced significantly lower total mean scores than the other groups. Even though mainly word finding difficulties were reported by the subjects, results on the test indicated the presence of language problems in other areas as well. The raw scores on all ten subtests were lower for the group with self-reported language problems. The difference between raw scores was highest in the following subtests: Inference, Word definitions, and Comprehension of Ambiguities. Surprisingly, the difference was only marginal in the subtests Naming and Word fluency.

As the damage to the brain in multiple sclerosis is very disperse it its probably very difficult to find an MS language syndrome with a distinct pattern of deficits. High interpatient variability is also found to be characteristic of patterns and severity of cognitive disorders. Therefore, one may conclude that the variability is applicable to language deficits as well. Since the clinical manifestations of high-level language are variable and there is no conclusive linguistic theoretic framework available for high-level language, a test battery for these functions will aim at specific linguistic abilities, in order to detect subclinical symptoms in for example incipient language decline.

Since difficulties on the neuropsychological screening in this study co-occurs with lower scores on language testing one might conclude that an assessment tool for high-level language is unnecessary. In our opinion this is not the case. Even though speech/language pathologists and neuropsychologists have found problems in the same patients, the problems detected are not necessarily the same, and they often need to be treated by one or both professionals form their different points of view.

When the present test battery has been tried out in larger populations, norms for individuals functioning at different levels can be set. It is also desirable to set correction factors for age and education on raw scores, as we assume that these variables affect performance on language tasks in the test battery.

5. REFERENCES

1. Lewis FM, LaPointe LL, Murdoch BE, Chenery HJ. Language impairment in Parkinson´s disease. Aphasiology 1998; 12(3):193-206.
2. Grossman M, Robinson KM, Onishi K, Thompson H, Cohen J, DEsposito M. Sentence comprehension in multiple sclerosis. Acta Neurologica Scandinavica 1995;92:324-331.
3. Lethlean JB, Murdoch BE. Language problems in multiple sclerosis. Journal of medical speech-language pathology 1993;1(1):47-50.
4. Lethlean JB, Murdoch BE. Naming errors in multiple sclerosis; support for a combined semantic/perceptual deficit. Journal of Neurolinguistics 1994;8(3):207-223.
5. Lethlean JB, Murdoch BE. Performance of subjects with multiple sclerosis on test of high-level language. Aphasiology 1997;11:39-57.
6. Crosson B. Assessment of subtle language deficits in neuropsychological batteries: Strategies and implications. In Sburdone RJ and Long CJ editors. Ecological validity of neuropsyhological testing. Delray, FL:GR Press/St Lucie press, Inc;1996.
7. Faber-Langendoen K, Morris JC, Knesevich JW, LaBarge E. Miller JP, Berg L. Aphasia in Senile Dementia of the Alzheimer Type. Annals of Neurology. 1988;23(4):365-370.
8. Keenan JS. The detection of minimal dysphaisa. Archives of physical medicine and rehabilitation 1971;52:227-232.
9. Darley FL, Brown JR, Goldstein NP. Dysarthria in multiple sclerosis. Journal of Speech and Hearing research 1972;15:229-245.
10. Hartelius L, Runmarker B, Andersen O. Prevalence and charateristics of dysarthria in a multiple sclerosis incidence cohort: in rlation to neurological data. 1997. manuscript accepted for publication in Folia Phoniatrica et Logopaedica.
11. Olmos-Lau N, Ginsberg M, Geller J. Aphasia in Multiple Sclerosis. Neurology 1977;27:623-626.
12. Beatty WW, Goodkin DE, Hertsgaard D, Monson N, Clinical and Demographic predictors of Cognitive performance in Multiple Sclerosis: Do Diagnostic Type, Disease Duration, and Disability Matter? Archives of Neurology 1990,47:305-308.