Running head: Picture description in aphasia

Title: ORAL AND WRITTEN PICTURE DESCRIPTION IN INDIVIDUALS WITH APHASIA

Authors: Dorien Vandenborre [1,2], Evy Visch-Brink [3], Kim van Dun [1], Jo Verhoeven[4,5], Peter Mariën [1,6]

Name of institution where study was performed: Clinical and Experimental Neurolinguistics, Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussels, Belgium

Affiliations of authors:

[1] Clinical and Experimental Neurolinguistics, Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussels, Belgium

[2] Thomas More Logopedie & Audiologie, Molenstraat 8, B-2018 Antwerp, Belgium

[3] Department of Neurology and Neurosurgery, Erasmus University Medical Centre, Wytemaweg 80, 3015 CN Rotterdam, the Netherlands

[4] School of Health Sciences, Division of Language and Communication Science, City University London, Northampton Square, London EC1V0HB, UK

[5] CLIPS, University of Antwerp, Lange Winkelstraat 40, B-2000 Antwerp, Belgium

[6] Department of Neurology and Memory Clinic, ZNA Middelheim Hospital, Lindendreef 1, B-2020 Antwerp, Belgium

Abstract (198 words) and Full text (5450 words)

Abstract

- Aim: To explore the differences between the oral and written description of a picture in individuals with chronic aphasia (IWA) and healthy controls. Descriptions were controllled for productivity, efficiency, grammatical organization, substitution behavior and discourse organization.
- Method: 50 IWA and 50 healthy controls matched for age, gender and education, provided an oral and written description of a black and white situational drawing from the Dutch version of the Comprehensive Aphasia Test. Between- and within-group analyses were carried out and the reliability of the test instrument was assessed.
- Results: The language samples of the healthy controls were more elaborate, more efficient, syntactically richer, more coherent and consisted of fewer spoken and written language errors than the samples of the IWA. Within-group comparisons showed that connected writing is more sensitive than connected speech to capture aphasic symptoms.
- Conclusion: The analysis of both modalities (speech and writing) at discourse level allows to simultaneously assess microlinguistic and macrolinguistic skills and their potential interrelations in a given IWA. Connected writing appears to be more sensitive in discriminating IWA from healthy controls than connected speech. This method for analyzing language samples should, however, be used in conjunction with other assessment tools.

Keywords: connected speech, connected writing, aphasia, picture scene, linguistic markers, reliability

This paper aids:

- What is already known on this subject? Aphasia is characterized by difficulties in connected speech/writing

- What this study aids? Within- and between-group comparisons for connected speech and connected writing in IWA and healthy controls
- Clinical implications? Connected writing is more sensitive in discriminating IWA and healthy controls than connected speech

INTRODUCTION

Aphasia is an acquired, multi-modal language disorder resulting from neurological damage (Royal College of Speech and Hearing Therapists, 2016), that does not affect intelligence (American Speech-Language and Hearing Association, 2016). Aphasia is characterized by difficulties in the production of connected speech/writing to efficiently and accurately convey information to listeners/readers. However, little is known about connected writing in aphasia and the potential difference with connected speech. Moreover, there is no consensus about the qualitative method for assessing connected language (del Toro et al., 2008).

Written and spoken language are essential aspects of communication both serving to mediate ideas, but they differ in several aspects (see e.g., Catts & Kamhi, 2005). While speaking is spontaneous and informal, writing enables the interlocutor to edit and to try out several options before completing the final product (Behrns, Wengelin, Broberg & Hartelius, 2009). Differences in the developmental and organizational aspects of writing and speech suggest that both modalities may be differentially affected in individuals with chronic aphasia (IWA). However, it should be kept in mind that there may well be substantial individual differences (e.g., Mortensen, 2005).

Describing a picture is an elicitation method frequently used to generate language samples. The main advantages of this method are that (1) an explicit referent with a ‘supporting’ (Fergatitis, Wright & Capilouto, 2011) content is used and (2) memory and sustained attentional demands are minimized (McNeil et al., 2005). A theoretical motivation for this elicitation method is also grounded in psycholinguistic mechanisms. Lambon Ralph (1999) suggested that object stimuli allow direct access to meaning (the assumption of privileged access) and that, because the relation between word form and word function is arbitrary, knowledge about the structural aspects of an object will be closely linked to the semantic properties that specify function (the assumption of privileged relationships). These two assumptions taken together imply that not only structural but also functional aspects of concept knowledge should be more readily accessed from a picture-scene. There might be compelling evidence for the existence of a layered semantic organization (Vandenborre & Mariën, 2014). Ross and Wertz (1999) have found significant correlations in IWA between changes in picture description information and listener's perception of changes in communicative ability. The Cookie Theft picture of the Boston Diagnostic Aphasia Examination (BDAE: Goodglass & Kaplan, 1972) and the picnic-picture of the Western Aphasia Battery (WAB: Kertesz, 1982) (e.g., Andreetta, Cantagallo & Marini, 2012) are used to examine oral discourse in traditional test batteries. However, the recently developed situational plate of the Comprehensive Aphasia Test (CAT: Swinburn, Porter & Howard, 2005) focuses on both oral and written language.

A number of different analysis techniques for connected oral speech have been proposed in the literature. These include word-by-word analyses which focus on information content, on syntax or on pragmatics (e.g., Nicholas & Brookshire, 1993). In the BDAE and in the WAB subjective ratings of language characteristics are used rather than objective measures. While the use of rating scales in the BDAE and the WAB is straightforward and sufficient for clinical reasons, various key phenomena such as language errors remain unsufficiently explored (Vermeulen, Bastiaanse & Van Wageningen, 1989) and the underlying cause remains unspecified. As a result these rating scales do not provide sufficient information for optimal intervention.

A detailed analysis of written connected language has been performed in etiologically heterogeneous patient groups with neurological damage. There has been research on IWA as a consequence of stroke (e.g., writing: Mortensen, 2005), closed head injury (CHI) (e.g., writing: Wilson & Proctor, 2002), Alzheimer disease (writing and speech: Croisille et al., 1996), nonfluent progressive aphasia (writing and speech: Graham, Patterson & Hodges, 2004) and also on normal young adults (writing and speech: Smith, Heuerman, Wilson & Proctor, 2003). However, these studies have focused on between-group rather than within-group comparisons. For example, Ulatowska et al. (1983) compared spoken and written story-retelling of IWA and healthy controls. They concluded that IWA made the same kind of errors than healthy controls but on a higher rate. Mortensen (2005) analyzed personal letters written by IWA and non-brain-damaged controls. Results indicated that all the IWA had mild to moderate impairment of writing following a single left hemisphere stroke. Their connected writing is equally coherent as that of the controls, but less productive. Only Behrns et al. (2009) compared writing and speech with a focus on within-group and between-group comparisons. The study included 8 IWA (mean age 42.5 years) and 10 healthy controls (mean age 23.5 years). The written output of the IWA was characterized by a less adequate vocabulary and a less complex syntax than in speech, but it was more coherent. Compared with the controls, IWA had a low production rate in writing, made proportionally more word-level errors and their narratives were less structured. When analyzing the time needed for the writing process, the texts of the IWA were the result of hard and time-consuming work. Despite these efforts, however, the written and spoken narratives of the IWA included the most essential parts of the story.

The aim of this study is to analyze and compare connected speech and writing in individuals with chronic aphasia collected by means of a picture description task. The analysis parameters are productivity, efficiency, grammatical organization, substitution behavior and discourse organization. It is hypothesized that (1) both modalities are affected in IWA and (2) that there are clear differences in speech and writing within IWA and within the control group in terms of the parameters used.

METHOD

Participants

In total, 100 participants took part in the study: 50 individuals with chronic aphasia (31 men, 19 women; mean age: 63.4; mean level of education: 1.9; mean duration of aphasia: 26.0 months post-onset neurological symptoms) and 50 healthy controls matched for age (mean age: 63.8), gender (31 men and 19 women) and education (mean level of education: 1.9) (Table 1). These healthy controls were carefully matched out of a data collection of 404 healthy individuals. This group was added to investigate the sensitivity of the chosen variables and the causuality of the possible differences. Are the differences between connected speech and connected writing in individuals with aphasia typical for this population or typical for the different output modalities. The IWA were consecutively recruited from a hospital (Ziekenhuis Netwerk Antwerp) and a rehabilitation centre (CEPOS Duffel), thereby collecting a gamut of severity ranges. In order to be included in the study, participants had to (1) be native speakers of Belgian Dutch; (2) be adults (> 18 years of age); (3) have no previous history of neurological, psychiatric illness or dementia as indicated by the medical files and self-report; (4) be right-handed as determined by a score of ≥+80 on the Edinburgh Handedness Inventory (EHI: Oldfield, 1971); and (5) have adequate hearing and vision as indicated by the medical files. Three additional criteria were used to select the study population; (6) aphasia had to be due to a vascular etiology (hemorrhage or ischemic stroke);

(7) aphasia had to be present for at least 12 months; (8) participants had to be able to produce at least 10 intelligible words in oral as well as in written response to the picture (Nicolas & Brookshire, 1993); and (9) concomitant motor speech impairment or apraxia of speech were excluded. All controls obtained a normal score on the Mini Mental State Examination (MMSE: Folstein, Folstein & McHugh, 1975). The IWA were examined by the same speech and language therapist (SLT), while the controls were assessed by final year students (master level SLT). The demographic characteristics of the participants are summarized in table 1:

Insert* Table 1 *near here

Data collection

Participants completed the Dutch version of the CAT (Swinburn et al., 2004; CAT-NL: Visch-Brink, Vandenborre, De Smet & Mariën, 2014) in a quiet room in the hospital in a single session lasting 60-90 minutes. This study only focuses on two subtests of the CAT-NL, i.e. spoken (subtest 19) and written picture description (subtest 27). Maximum 15 minutes elapsed between the two description tasks. The ideal description of the used picture, a black and white drawing of a situational scene, is: a man is sleeping in an armchair, while a cat is trying to catch a fish in a fish bowl. The cat pushes some books from the shelf which are going to fall on the sleeping man. A child, sitting on the floor, playing with its toys, is trying to warn the man about the falling books.

Participants were instructed to orally describe (subtest 19) and write (subtest 27) in as much detail as possible what they observed in the picture. All participants were given a maximum of five minutes to describe the picture as indicated in the manual of the CAT-NL. Speech samples were recorded by means of a high-quality digital voice recorder (smartphone: Apple iPhone 5s) that was placed next to the target picture. The samples were transcribed in standard orthography, with neologisms and phonemic language errors transcribed phonetically using the narrow IPA transcription. Written samples were provided on paper. The research was conducted in accordance with the Helsinki Declaration and the ethics committee of Ziekenhuis Netwerk Antwerp which approved the study.

Data-analysis

The language assessment protocol consisted of the CAT-NL and the Token Test as part of the Aachen Aphasia Test (AAT-TT: Graetz, de Bleser & Willmes, 1992). The CAT is a standardized comprehensive language test which consists of three components, i.e. a cognitive screen, a language test and a disability questionnaire. The TT is administered to determine aphasia severity (De Renzi & Vignolo, 1962). The AAT-version consists of 50 items and results in an error score, so lower scores indicate better performance on the test. As indicated in the manual of the AAT four levels are used: no aphasia (Token Test 0-6), mild aphasia (Token Test 7-23), moderate aphasia (Token Test 24-40), severe aphasia (Token Test 41-50).

Variables included in the scoring system

In this study, a more detailed scoring on different language processing levels was used instead of the original scoring of the CAT(-NL) (Table 2). The original scoring of the CAT-NL is a qualitative measure taking into account the following parameters : accuracy and adequacy of information (oral and written description), grammatical correctness (oral and written description), grammatical diversity (oral description), fluency (oral description) and an overall score on a ten-points scale (oral and written description). The variables included in the current scoring system are chosen so there is no bias between fluent speech (productive, less efficient, grammatical variety) and non-fluent speech (less productive, more efficient, less grammatical variety) (Visch-Brink et al., 2014).

Insert* Table 2 *near here

Productivity was expressed as the total number of words (TNWs) and the mean length of utterances (MLU). The TNW-count included all verbalizations, i.e. not only phonologically correct words, but also nonwords, phonological language errors, repetitions and minimal responses (such as ‘yes’ and ‘well’). Comments and stereotypes were also included in the word count and this is consistent with the Analysis of Spontaneous Speech in Aphasia (ASTA: Boxum, van der Scheer & Zwaga, 2010). ASTA is a frequently used scoring tool in clinical practice in the Netherlands.

The second indicator of productivity was MLU (Wagenaar, Snow & Prins, 1975). For each language sample the total number of utterances was counted (ASTA: Boxum et al., 2010). An utterance is defined as a complete thought which is separated from other utterances on the basis of grammatical completeness, content and intonation contour/visual segmentation. MLU was calculated by dividing the TNWs by the number of utterances (ASTA: Boxum et al., 2010).

Discourse efficiency was established by counting the number and percentage of correct information units (CIU and % CIU)**. CIUs** are defined as ‘content words and function words that are intelligible in context, accurate in relation to the picture(s) or the topic, and relevant to and informative about the content of the picture(s)’ (Nicholas & Brookshire, 1993, p. 348). CIUs capture the degree to which words are consistent with the overall semantic content of the language sample. Since the number of CIUs is a quantitative measure, a more balanced measure was also added: %CIUs expresses the number of CIUs relative to the TNW.

The number of **correct sentences and** the number of **compound sentences** were taken as an indication of grammatical organization. A sentence was considered correct if all the arguments required by the verb were correctly inserted (ASTA). A compound sentence was defined as a sentence which contains two or more independent clauses joined by a coordinating or/and subordinating conjunction and sentences which contain one or more relative clauses. A grammatical index was calculated by dividing the number of correct compound sentences by the number of utterances and then multiplying this value by 100 (Thompson, Shapiro, Tait, Jacobs, & Schneider, 1996).

The percentage of language errors provides an indication of substitution behavior. The percentage was derived by dividing the total number of language errors by the TNW and then multiplying this value by 100. The following language errors were included: (1) phonemic errors; (2) verbal phonological errors; (3) verbal semantic errors; (4) unrelated errors; (5) morphological errors; and (6) neologisms (Table 2) (Mariën, Mampaey, Vervaet, Saerens & De Deyn, 1998). All these errors were transcribed in the oral and written descriptions, but those representing phonemic errors in the oral descriptions are actually ‘spelling errors’ in the written samples.

Cohesion markers were counted and cohesive adequacy was evaluated. Cohesion ratings were based on Van Leer and Turkstra’s (1999) coding method. A cohesion marker is a word which connects the meaning expressed in the ongoing utterance with that already expressed, or which will be expressed in subsequent utterances (Marini et al., 2011). Three types of cohesion markers were studied: (1) lexical; (2) reference; and (3) conjunction markers (Table 2). Both the absolute and the relative (to the TNW) number of cohesive markers were quantified. In order to study how efficient a speaker/writer maintains meaning across the discourse, **cohesion adequacy** was assessed. Individual cohesive ties were labeled as being correct or incorrect (Liles, Coelho, Duffy & Zalagens, 1989). Cohesive adequacy was calculated as a percentage by dividing the number of correct ties by the total number of ties and multiplying this value by 100.

Reliability of analysis method

Inter-rater and inter-subject reliability were measured by means of intra-class correlation coefficients (ICC: Shrout & Fleiss, 1979). Inter-rater reliability was measured by comparing the scores of two independent raters (one SLT and one SLT-student). Intra-subject or test-retest reliability was assessed using the same population of twenty participants. They were tested twice in identical locations with a delay of approximately one month between testing sessions.

Statistical analysis

The mean, standard deviation and the range of different measures were obtained for the two groups and for the two modalities.

Performance between the two groups was compared by means of paired t-tests. In order to assess the strength of the relationship between aphasia severity and the different linguistic markers, Pearson's correlation coefficients were calculated between the scores of the IWA on the Token Test and the different linguistic markers. The impact of demographic variables on the results was also assessed.

The results of the analyses are summarized in table 3:

Insert* Table 3 *near here

RESULTS

The severity level of aphasia differed in the study group. Based on the scores of the Token Test: 30 IWA had mild aphasia (Token Test score between 7-23), 9 IWA had moderate aphasia (Token Test score between 24-40) and 11 IWA had severe aphasia (Token Test score between 41-50).

Productivity

Between-group comparison of TNW indicates that the sample length was significantly shorter in the group of IWA than in the control group in writing only (written: t(49)=-4.39, *p*=0.00; oral: t(49)=-0.82, *p*=0.42). However, with regard to **MLU,** IWA used shorter utterances than the controls in both modalities, (oral MLU: t(49)=-4.82, p=0.00; written MLU: t(49)=-4.25, *p*=0.00) (Table 4). Within both groups the oral samples counted more words than the written samples (aphasic group: t(49)=7.58, *p*=0.00; controls: t(49)=5.49, *p*=0.00) (Table 3). The output modalities did not significantly influence the MLU (aphasic group: t(49)=1.00, *p*=0.33; controls: t(49)=--0.38, *p*=0.70) (Table 4).

Insert* Table 4 *near here

Severity of aphasia was negatively related with the TNW in writing only, and with the MLU in both modalities (Table 5). The TNW in writing and the MLU in speech and writing were lower in more severe aphasia.

Insert* Table 5 *near here

Efficiency

The counts for efficiency are summarized in table 6:

Insert* Table 6 *near here

Between-group comparisons show that both in speech and in writing, IWA produced fewer CIUs and a significantly lower proportion of **CIUs** (%CIU) than the controls (e.g. oral CIUs: t(49)=-3.09, *p*=0.00; oral %CIUs: t(49)=-6.46, *p*=0.00) (Table 6). Within-group analysis indicate that both groups produced significantly more CIUs in speech than in writing (aphasic group CIUs: t(49)=6.26, *p*=0.00; controls CIUs: t(49)=4.83, *p*=0.00), but the percentage of CIUs was lower in speech than in written output (aphasic group %CIUs: t(49)=-1.99, *p*=0.05; controls %CIUs: t(49)=-4.22, *p*=0.00) (Table 6). Furthermore, there was a significantly negative correlation between aphasia severity and production of CIUs (%CIUs) in both modalities (e.g. oral CIUs-TT: R=-0.32, p=0.02) (Table 5); a greater range of CIUs was produced with less severe aphasia.

Grammatical organization

Between-group analysis showed, except for the number of oral correct sentences (t(49)=-1.62, p=0.11), significant differences to the detriment of the IWA between written correct sentences (t(49)=-3.51, p=0.00), oral and written compound sentences (e.g. t(49)=-4.18, p=0.00) and oral and written grammatical index (e.g., t(49)=-5.43, p=0.00) (Table 7). Within-group comparisons showed that both groups used significantly more **correct sentences** in speech than in writing (aphasic group: t(49)=7.16, p=0.00; controls: t(49)=5.08, p=0.00), and more **compound sentences** in speech than in writing (aphasic group: t(49)=4.68, p=0.00; controls: t(49)=4.15, p=0.00) (Table 7). There was no difference concerning **grammatical index** between both modalities (aphasic group: t(49)=1.03, p=0.31; controls: t(49)=-0.28, p=0.78) (Table 7). Furthermore, there was a strong negative correlation between grammatical organization and aphasia severity in both modalities for all parameters (oral grammatical index-TT: R=-0.37, p=0.01; written grammatical index-TT: R=-0.40, p=0.00; oral correct sentences-TT: R=-0.37, p=0.01; written correct sentences-TT: R=-0.40, p=0.00; oral compound sentences-TT: R=-0.34, p=0.03; written compound sentences-TT: R=-0.41, p=0.00) (Table 7).

Insert* Table 7 *near here

Substitution behavior

Between-group comparisons indicate that in the study group of IWA, the absolute number of language errors did not differ significantly between modalities (t(49)=1.50, *p*=0.14), but only in relation to the TNW (%language errors) (t(49)=-2.84, *p*=0.01) (Table 8). The controls made significantly more language errors in speech than in writing (language errors: t(49)=3.13, *p*=0.00; %language errors: t(49)=2.59, *p*=0.01) (Table 8). It should be noted that controls often corrected their errors themselves, while self-corrections were rarely observed in the individuals with aphasia. The within-group analysis showed that IWA made significantly more language errors than the controls, and this applies to both modalities (oral: t(49)=5.77, *p*=0.00; written: t(49)=5.69, *p*=0.00). This was also the case for the percentage of language errors (oral: t(49)=6.80, *p*=0.00; written: t(49)=4.00, *p*=0.00). A strong negative correlation was found between substitution behavior and aphasia severity in both modalities (e.g., spoken language errors-TT: *R*=0.38, *p*=0.01; spoken% language errors-TT: *R*=0.48, *p*=0.00; written language errors-TT: *R*=0.38, *p*=0.01; written %language errors-TT: *R*=0.57, *p*=0.00) (Table 5).

Insert* Table 8 *near here

Discourse organization

Cohesion adequacy significantly differed between both groups in both modalities (oral: t(49)=-3.17, p=0.00; written: t(49)=-5.62, *p*=0.00) (Table 9) in favor of the controls. However, only in writing the IWA used less **cohesive markers** than the controls (t(49)=-3.32, *p*=0.00). In speech no significant difference between both groups in terms of the absolute or relative number of cohesive markers (CM) was found (oral CM: t(49)=-1.21, *p*=0.23; oral %CM: t(49)=-0.64, *p*=0.52). As far as within-group comparisons are concerned, both groups used more cohesion markers in speech than in writing (e.g., aphasic group oral mean: 5.8, sd: 3.8; aphasic group written mean: 2.3, sd: 2.4) (Table 3): the difference was only significant in absolute numbers (CM: t(49)=5.74, *p*=0.00) and not in percentage (%CM: t(49)=-0.09, *p*=0.93) (Table 9). The oral and written samples of controls were cohesive to the same extent (cohesion adequacy: 100%), whereas the speech of the IWA was significantly more cohesive than the writing (t(49)=2.25, *p*=0.03) (Table 9). A strong negative correlation was found between the number of cohesion markers in writing and the severity of aphasia (written CM-TT: R=-0.34, p=0.02; written %CM-TT: *R*=-0.58, *p*=0.00)). No correlations were found between oral discourse organization and the severity of aphasia (oral CM-TT: *R*=0.06, *p*=0.66) or between written cohesive adequacy and the severity of aphasia (written cohesive adequacy-TT: *R*=-0.13, *p*=0.26) (Table 5).

Insert* Table 9 *near here

Demographic variables

In this study level of participants' education, age and gender only had a significant impact on the performance of IWA in the following combination: (1) oral %cohesive markers and education (*R*=-0.32, *p*=0.02), (2) written MLU and age (*R*=-0.31, *p*=0.03) and (3) written cohesive adequacy and gender (*R*=0.34, *p*=0.02) (Table 10).

However, more significant correlations were found between linguistic markers and demographic variables for the controls (Table 10). Education and language output correlated significantly with productivity, efficiency, grammatical organization and discourse organization (e.g., oral TNW: *R*=0.39, *p*=0.01) (see shaded areas in Table 10). There was no significant correlation between education and substitution behavior (e.g., spoken language errors: *R*=-0.00, *p*=0.98). A different pattern emerged when evaluating the impact of age for the controls. Age correlated significantly with substitution behavior: older subjects made more language errors in speech than in writing (e.g., spoken %language errors: *R*=0.35, *p*=0.01) (Table 10). There was no significant correlation between age and productivity, efficiency, grammatical organization nor discourse organization, except for written %CIUs and written cohesion markers (written %CIUs-age: R=-0.28, p=0.05; written CM-age: R=-0.37, p=0.01) (Table 10). There was no gender-related impact on the performances of the controls (Table 10).

Insert* Table 10 *near here

Reliability of the analytical method

Inter-rater and inter-subject reliability was measured by means of intra-class correlation coefficients (ICC: Shrout & Fleiss, 1979). A good to excellent consistency was obtained for inter-rater reliability (ICC ranged from 0.85 to 1.00) (Table 11). Intra-subject reliability was lower (ICC ranged from 0.58 to 0.91). More details are found in table 11.

Insert* Table 11 *near here

DISCUSSION

Oral and written picture description of 50 individuals with chronic aphasia and 50 matched healthy controls were analyzed in terms of productivity, efficiency, grammatical organization, substitution behavior and discourse organization.

Between-group differences

In line with the literature (e.g. Gordon, 2006) the output of IWA was impaired in both modalities and there was a significant difference between the IWA and the controls for all investigated dimensions. Compared to the controls, the speech and writing of the IWA were impaired at all levels of investigation. In line with Mortensen (2005) and Behrns et al. (2009), both modalities of the IWA were less productive (lower MLU), less efficient (lower number of CIUs and a lower %CIUs), less well-organized grammatically (fewer compound sentences, lower grammatical index) and consisted of more language errors than the speech and writing of the controls. In contrast to Mortensen (2005) and Behrns et al. (2009), but in line with Ulatowska et al. (1983), writing was less cohesive (less cohesive adequacy) than speech. In speech, there was no significant difference between both groups regarding the TNW, the number of correct sentences, the number of cohesion markers and the percentage of cohesion markers. In writing, on the other hand, the percentage of cohesion markers was the only parameter which did not discriminate the IWA from the controls. These results conform the findings of other studies (e.g., Ulatowska et al., 1983; Marini et al., 2011) in which IWA were found to have greater difficulty in expressing relations and causal links between units of information. It remains controversial as to whether these difficulties are due to pragmatic deficits or whether they are an epiphenomenon of a lexical deficit (e.g., Andreetta et al., 2012).

The study results indicate that aphasia has an impact on the microstructure and the macrostructure. The problems experienced by IWA involve both the specific impaired linguistic level (e.g. fluency, content, syntax, phonology, coherence) and the relationship between these different levels. This can be accounted for in various ways: (1) IWA may have ‘lost’ or cannot access the desired algorithm to formulate connected language; (2) IWA cannot perform all the operations needed to formulate connected language since they have limitations in the ability to do multiple operations (‘resource reduction’). As computational load increases, errors do as well (Caplan et al., 2007). Even in healthy subjects, connected writing is a more time-consuming activity than connected speech. Despite this, it can be concluded that in terms of the parameters used in the present study, writing appears to be more sensitive in discriminating IWA from healthy people than the oral picture description. This is an important finding, since written picture description is rarely used in the diagnostic work-up of aphasia. Frequently used comprehensive language tests, such as the BDAE (Goodglass & Kaplan, 1972), do not include a written picture description task. Moreover, there is no golden standard to assess a written narrative, nor is there a principled method to analyze the samples. Since in our study connected writing appeared to be more sensitive than connected speech in discriminating IWA from healthy people, it might be a valuable tool in the assessment of individuals with mild aphasia in view of other etiologies. For example in individuals with brain tumors a decrease of everyday language skills is a frequent complaint, but when tested the usual language tasks such as naming do not always reveal abnormalities, nor do they capture all the communication problems encountered by these individuals (Satoer, Vincent, Smits, Dirven & Visch-Brink, 2013).

Within-group differences

The greater sensitivity of writing is illustrated very clearly from the more pronounced relation with aphasia severity in writing than in speech. A negative correlation was found in connected writing between aphasia severity and productivity, efficiency, grammatical organization and discourse organization. A positive correlation was found in the written version between aphasia severity and substitution behavior. These findings indicate that output in severe aphasia is less productive, efficient and grammatically complex with fewer cohesion markers and with more language errors.

All the participants produced stories in which the written version consisted of fewer words than the spoken version and this is consistent with other studies (e.g., Behrns et al. (2009). It might be that written description is a more demanding task for individuals with aphasia, since it does not only involve cognitive and linguistic skills, but also writing-specific abilities such as motor action and control (Graves-Wright et al., 2004). The greater number of CIUs in speech might be related to the reported higher TNW in speech. However, del Toro et al. (2008) pointed out that it is more informative to look at the %CIUs because of the differences in TNW produced. The fact that the oral speech task elicits more words than the written version together with the observation that more CIUs and lower %CIUs characterize speech than writing in both groups illustrates the basic difference between speech and writing: a more unplanned and repetitive versus a more planned and better controlled language modality.

Comparing both modalities within groups, the IWA and the controls appear to have a rather similar profile with an overall better quality of speech than writing. Both groups realized longer descriptions, more CIUs, more correct sentences, more compound sentences and more cohesive markers in speech than in writing. No significant difference was found between speech and writing in MLU, grammatical index and the percentage of cohesion markers. Cohesion adequacy (better in speech than in writing for IWA) and the number of language errors (in the control group more in speech than in writing) were the only parameters with a selective difference in one of the groups compared.

The results of this study indicate that IWA used more correct sentences and more compound sentences in speech than in writing. The higher number of correct sentences and compound sentences in oral speech might be due to the fact that speech was more productive than writing. These results contrast with the findings of Behrns et al. (2009) and Ulatowska et al. (1983) who reported relatively higher syntactic complexity in writing. This difference might be due to the analytical methods used: whereas Behrns et al. (2009) and Ulatowska et al. (1983) focused on words and clauses per T-unit (a main clause plus any clauses subordinate to it) this study only counted the number of correct and compound sentences.

The fact that more cohesive markers were used in speech than in writing by both groups could also be related to the greater productivity in speech than in writing. However, the fact that the speech of IWA resulted in a higher percentage of cohesive adequacy than writing is not in agreement with the observations of Mortensen (2005) and Behrns et al. (2009), who observed no difference in coherence between IWA and controls and more coherence in writing than in speech.

Impact of demographic variables

A striking difference was found between both groups. Although for the IWA hardly any linguistic marker correlated significantly with demographic variables, the opposite was true for the controls. Three positive significant correlations were found for the IWA: (1) education and oral %cohesive markers, (2) age and written MLU and (3) gender and written cohesive adequacy. In the control group a significant correlation between education and TNW, MLU, number of CIUs, number of compound sentences, written cohesive markers and oral cohesive adequacy was found, indicating that higher education leads to more productive, efficient, grammatically complex and cohesive oral and written output. These findings are in line with Mackenzie, Brady, Norrie and Poedjianto (2007) and Behrns et al. (2009), confirming a significant positive correlation between education and productivity and discourse organization.

Contradictory findings are reported in the literature regarding age-related factors. In the control group, and in line with Shewan and Henderson (1988), age correlated with substitution behavior. In line with e.g., Mackenzie (2000), age correlated with written %CIUs and written cohesive markers. In the literature significant correlations were found between productivity and age (e.g., Harris, Kiran, Marquadt & Fleming, 2008), efficiency and age (e.g. Capilouto, Wright & Wagovich, 2005), grammatical organization and age (e.g., Ulatowska, Allard & Chapman, 1990), and discourse organization and age (e.g., Harris et al., 2008). These findings were not replicated in this study, what could be due to the fact that our study group was relatively young (mean age: 63.4 years).

Reliability of the test instruments

As in Forbes-MacKay and Venneri (2005), the results of this study suggest that a picture description task is a reliable measure for connected language. High intra-class correlations were found for inter-rater reliability, although lower scores were found for intra-subject reliabilitiy. This is in line with the findings of Goodglass and Kaplan (1983) since reliability in the sense of repeatability of results on retesting any individual with a stroke varies amongst IWA to a degree rarely found in other types of patients. However, as Nicolas and Brookshire (1993) pointed out, even though performance on individual items may be inconsistent, overall scores across a set of items can show a high degree of consistency.

Study Limitations

A counterbalanced procedure might have given different results. In this study, participants first described the picture scene orally and afterwards provided a written description, thinking that they can omit information which they had already given orally. However, it may be that the knowledge of the picture and the experience in connected speech the quality of connected writing exceeds the quality of connected speech. It might be that given in another order connected writing was even more sensitive than connected speech. Since conflicting results are reported in the literature about the impact of demographic variables on normative data, a larger study group is required.

Other methodological limitations of the study are: (1) the elicitation method; (2) the length of the samples; (3) the absence of speech/writing rate; (4) the absence of longitudinal data for the IWA, (5) the absence of a measure for estimating experience in writing, (6) the absence of measuring motor speech disorders (dysarthria and apraxia of speech) and limb apraxia and (7) the impact of different variables such as writing skills and intelligence on connected writing. Firstly, in the literature a variety of techniques has been used to elicit connected speech, such as the use of pictures (e.g., Capilouto et al., 2005), interactive conversation (e.g., Coelho, Youse & Le, 2002) and monologic recounts (e.g., McNeil et al., 2007). These methods differ in terms of the cognitive and linguistic demands required from the speaker/writer (Bliss & McCabe, 2006). Secondly, it is still unclear what constitutes a representative writing sample (Armstrong, 2000). In speech, a sample of 300 words appears to be a reliable volume for a quantitative linguistic analysis (Prins & Bastiaanse, 2004). Thirdly, participants were given a maximum of 5 minutes to describe the picture, but the actual time used to describe the scene was not registered. Aphasia might impair the speed with which a message is conveyed (e.g. Graham et al., 2004). Fourthly, it should be considered whether the method for analysis is sensitive enough to capture changes in aphasic discourse over time. Fifthly, writing as an artificial learned activity is more susceptible to individual variation than speech, especially since the introduction of digital communication (Catts & Kamhi, 2005). Sixthly, mild motor disorders may influence writing ability (Behrns et al. 2009) and motor adequacy was not assessed. Seventhly, it is expected that for experienced writers writing is a more automatic skill so that both the cognitive demand for writing and the motoric writing skills diminish implicating that these people can focus more on their linguistic skills (such as efficiency and grammatical organization) (Beeson, Higginson & Rising, 2013).

Future directions

Further research is needed to elicit the potential interface between microlinguistic and macro-linguistic aspects of connected language. This study focused on the analysis of the product, i.e. the written output. Further research might focus on the production process of writing itself (e.g. keystroke logging) so that planning, generation and revision could also be assessed (e.g., Behrns et al., 2009).

ACKNOWLEDGEMENT

Conflict of interest: The manuscript has not been previously published and is not under

consideration to be published anywhere.

REFERENCES

- ANDREETTA, S., CANTAGALLO, A. and MARINI, A., 2012, Narrative discourse in anomic aphasia. Neuropsychologia, 50, 1787-1793.
- ARMSTRONG, E., 2000, Aphasic discourse analysis: the story so far. Aphasiology, 14(9), 875-892.
- BEESON, P.M., HIGGINSON, K., and RISING, K. , 2013, Writing treatment for Aphasia: a texting approach. Journal of Speech, Language and Hearing Research, 56, 945-955.
- BEHRNS, I., WENGELIN, A., BROBERG, M. and HARTELIUS, L., 2009, A comparison between written and spoken narratives in aphasia. Clinical Linguistics and Phonetics, 23(7), 507-528.
- BLISS, L.S., and MCCABE, A., 2006, Comparison of Discourse Genres: Clinical Implications. Contemporary Issues in Communication Science and Disorders, 33, 126-137.
- BOXUM, E., VAN DER SCHEER, F. and ZWAGA, M., 2010, Analyse voor Spontane Taal bij Afasie; standaard in samenspraak met de VKL. Vereniging voor Klinische Linguïstiek.
- CAPILUOTO, G.J., WRIGHT, H.H. and WAGOVICH, S.A., 2005, CIU and main event analyses of the structured discourse of older and younger adults. Journal of Communication Disorders, 38(6), 431–444.

- CAPLAN, D., WATERS, G., DEDE, G., MICHAUD, J. and REDDY, A., 2007, A study of syntactic processing in aphasia I: Behavioral (psycholinguistic) aspects. Brain and Language, 101, 103-150.
- CATTS, H.W. and KAMHI, A.G., 2005, Language and reading: convergences and divergences. In H.W. Catts and A.G.,Kamhi (Eds.), *Language and reading disabilities* (Boston: Pearson), pp. 1–25.
- COELHO, C.A., YOUSE, K.M. and LE, K.N., 2002, Conversational discourse in closed-head -injured and non-brain-injured adults. Aphasiology, 16 (4-5-6), 659-672.
- CROISILE, B., SKA, B., BRABANT, M., DUCHENE, A., LEPAGE, Y., AIMARD, G., and TRILLET, M., 1996,, Comparative Study of Oral and Written picture description in patients with Alzheimer's Disease. Brain and Language, 53, 1-19.
- DEL TORO, C.M., ALTMANN, L.J.P., RAYMER, A.M., LEON, S., BLONDER, L.X. and ROTHI, L.J.G., 2008, Changes in aphasic discourse after contrasting treatments for anomia. Aphasiology, 22(7-8), 881-892.
- DE RENZI, E. and VIGNOLO, L.A., 1962, The Token Test: a sensitive test to detect receptive disturbances in aphasia. Brain, 85, 665-678.
- FERGADITIS, G., WRIGHT, H.H. and CAPILOUTO, G.J., 2011, Productive vocabulary across discourse types. Aphasiology, 25(10), 1261-1278.
- FOLSTEIN, M.F., FOLSTEIN, S.E. and MCHUGH, P.R., 1975, Mini-mental state: a practical method for grading the cognitive state of patients for the clinician. Journal of Psychiatric Research, 12, 189-198.
- FORBES-MCKAY, K.E. and VENNERI, A.,2005, Detecting subtle spontaneous language decline in early Alzheimer's disease with a picture description task. Neurological Sciences, 26, 243-254.

- GOODGLASS, H. and KAPLAN, E.F., 1972, Boston Diagnostic Aphasia Examination (BDAE). *(*Philadelphia: Lea and Febiger).
- GOODGLASS, H.,and KAPLAN, E., 1983, The assessment of aphasia and related disorders. (Philadelphia: Lea and Febiger).
- GRAETZ, P., DE BLESER, R. and WILLMES, K., 1992, *Akense Afasietest, Nederlandse versie*. (Lisse: Swets and Zeitlinger).
- GRAHAM, N.L., PATTERSON, K. and HODGES, J.R., 2004, When more yields less: Speaking and Writing deficits in nonfluent progressive aphasia. Neurocase, 10(2), 141-155.
- GRAVES-WRIGHT, K., NEILS-STRUNJAS, J., BURNETT, R. and O'NEILL, M.J., 2004, A comparison of verbal and written language in Alzheimer's disease. Journal of Communication Disorders, 37, 109-130.
- HARRIS, J.L., KIRAN, S., MARQUADT, T.P. and FLEMING, V.B., 2008,. Communication wellness check-up: age-related changes in communicative abilities. Aphasiology, 22(7-8), 813-825.
- KERTESZ, A., 1982, The Western Aphasia Battery. (New York: Grune and Stratton).
- LEER, VAN, E. and TURKSTRA, L., 1999, The effect of elicitation task on discourse coherence and cohesion in adolescents with brain injury – ages nine through nineteen. Journal of Communication Disorders, 32*,* 327–349.
- LILES, B.Z., COELHO, C.A., DUFFY, R.J. and ZALAGENS, M.R., 1989, Effects of elicitationprocedures on the narratives of normal and closed head-injured adults. Journal of Speech and Hearing Disorders, 54 (3), 356-366.
- MACKENZIE, C., 2000, Adult spoken discourse: the influences of age and education. International Journal of Language and Communication Disorders, 35(2), 269-285.

- MACKENZIE, C., BRADY, M., NORRIE, J. and POEJIANTO, N., 2007, Picture description in neurologically normal adults: Concepts and topic coherence. Aphasiology, 21(3/4), 340–354.
- MARIEN, P., MAMPAEY, E., VERVAET, A., SAERENS, J. and DE DEYN, P.P., 1998, Normative Data for the Boston Naming Test in native Dutch-speaking Belgian Elderly. Brain and Language, 65, 447-467.
- MARINI, A., GALETTO, V., ZAMPIERI, E., VORANO, L., ZETTIN, M. and CARLOMAGNO, S., 2011, Narrative language in traumatic brain injury. Neuropsychologia, 49, 2904-2910.
- MCNEIL, M.R., MATTHEWS, C.T., HULA, W.D., DOYLE, P.J., RUBINSKY, H.J. and FOSSETT, T.R.D., 2005, A dual-task tool for quantifying normal comprehension of aphasic connected speech production: a constructive replication. Aphasiology, 19 (3/4/5), 473-484.
- MCNEIL, M.R., SUNG, J.E., YANG, D., PRATT, S.R., FOSSETT, T.R.D., DOYLE, P.J. and PAVELKO, S., 2007, Comparing connected language elicitation procedures in persons with aphasia: concurrent validation of the Story Retell Procedure. Aphasiology, 21 (6/7/8), 775-790.
- MORTENSEN, L., 2005, Written discourse and acquired brain impairment: Evaluation of structural and semantic features of personal letters form a Systematic Functional Linguistic perspective. Clinical Linguistics and Phonetics, 19(3), 227-2417.
- NICHOLAS, L.E. and BROOKSHIRE, R.H., 1993, A system for quantifying the informativeness and efficiency of the connected speech of adults with aphasia. Journal of Speech and Hearing Research, 36(2), 338-350.
- OLDFIELD, R.C., 1971,The assessment and analysis of handedness: the Edinburgh Inventory. Neuropsychologia, 9, 97-113.

- PRINS, R. and BASTIAANSE, R., 2004, Review: Analyzing the spontaneous speech of aphasic speakers. Aphasiology, 18(12), 1075-1091.
- ROSS, K.R. and WERTZ, R.T.,1999, Comparison of impairment and disability measures for assessing severity of, and improvement in aphasia. Aphasiology, 13, 113-124.
- SATOER, D., VINCENT, A., SMITS, M., DIRVEN, C. and VISCH-BRINK, E., 2013, Spontaneous speech of patients with gliomas in eloquent areas before and early after surgery. Acta Neurochirugica, 155 (4), 685-692.
- SHEWAN, C.M. and HENDERSON, V.L., 1988, Analysis of spontaneous language in the older normal population. Journal of Communication Disorders, 21, 139–154.
- Shrout, P.E., & Fleiss, J.L. (1979). Intraclass Correlations: uses in assessing rater reliability.*Psychological Bulletin, 86(2),* 420-428.
- SMITH, R., HEUERMAN, M., WILSON, B.M. and PROCTOR, A., 2003, Analysis of normal discourse patterns. Brain and Cognition, 53, 368-371.
- SWINBURN, K., HOWARD, D. and PORTER, G.,2005, *CAT: Comprehensive Aphasia Test.* (Hove, UK: Psychology Press).
- THOMPSON, C.K., SHAPIRO, L.P., TAIT, M., JACOBS, B. and SCHNEIDER, S.L., 1996, Training wh-question production in agrammatic aphasia: analysis of argument and adjunct movement. Brain and Language, 52, 175–228.
- ULATOWSKA, H.K., FREEDMAN-STERN, R., WEISS-DOYLE, A., MACALUSO-HAYNES, S. and NORTH, A.J., 1983, Production of procedural discourse in aphasia. Brain and Language, 18, 315-341.
- ULATOWSKA, H.K., ALLARD, L. and CHAPMAN, S.B., 1990, Narrative and procedural discourse in aphasia. In Y. Joanette and H.H. Brownell (Eds.), Discourse ability and brain damage. Theoretical and empirical perspectives. (New York: Springer-Verlag), pp. 180-198.

- VANDENBORRE, D. and MARIEN, P., 2014, Broca meets Wernicke in a single case. Journal of Neurolinguistics, 29, 17-30.
- VERMEULEN, J., BASTIAANSE, R. and WAGENINGEN, VAN, B.V., 1989, Spontaneous speech in aphasia: a correlation study. Brain and Language, 36, 252-274.
- VISCH-BRINK, E.G., VANDENBORRE, D., DE SMET, H.Y. and MARIEN, P., 2014, De Comprehensive-Aphasia-Test Nederlands (CAT-NL). (Amsterdam: Pearson).
- WAGENAAR, E., SNOW, C. and PRINS, R., 1975, Spontaneous speech of aphasic patients: A psycholinguistic analysis. Brain and Language, 2, 281-303.
- WILSON, B.M. and PROCTOR, A. ,2002, Written discourse of adolescents with closed head injury.Brain Injury, 16(11), 1011-1024.