How do speakers with and without aphasia use syntax and semantics across two discourse genres?

Lucy Dipper[1], Madeleine Pritchard[1], Elizabeth Walkden[2], Madeline Cruice[1]

[1]Division of Language and Communication Sciences, School of Health Sciences, City, University of London

[2]Homerton University Hospital NHS Foundation Trust, Regional Neurological Rehabilitation Unit London, London, UK E9 6SR

Corresponding author

Dr Lucy Dipper

Division of Language and Communication Science

School of Health Sciences

City, University of London

Northampton Square

London EC1V 0HB

UK

l.t.dipper@city.ac.uk

Tel: +44 (0) 207 040 4658

ABSTRACT

Background: Discourse is an increasing focus of assessment in clinical and research settings because it reflects everyday communication. Everyday communication is likely to include a range of different discourse genres, e.g. describing a scene, or reflecting on life experiences. It is likely that speakers use verbs differently in these different discourse genres, but very little is known about this.

Aims*:* To explore whether there were differences in how two groups of speakers (with and without a communication impairment) used verbs in two different discourse genres, in terms of syntax and semantics.

Methods & Procedures*:* Data from people with aphasia (PWA) were taken from an earlier study (Cruice and colleagues, 2010; 2014), and neurologically healthy people (NHP) were recruited for the current study. Participants produced discourses from two genres: a picture description (the *Western Aphasia Battery* 'Picnic Scene') and personal narrative (reflective responses to quality of life questions). Discourses were analysed using measures of argument structure (mean Predicate Argument Structure score), verb weight (% heavy verbs) and verb semantic category (% mental and relational verbs). Comparisons were made for each measure between genre and group using a series of two mixed two-way ANOVAs.

Outcomes & results: Data from 26 PWA and 27 NHP were analysed. For PAS, there was a main effect of genre, significant interaction between group and genre, and main effect of group. For the semantic measures, there was a main effect of genre for % mental verbs but no effects or interactions for % heavy and % relational verbs. Post-hoc correlations explored associations between the variables.

Conclusions: Genre exerts no demonstrable impact on semantic weight, in either speaker group, but does exert influence on the semantic category of verbs used because, for both speaker groups, the picture description genre elicited a smaller percentage of mental verbs than the personal narratives produced in response to QOL questions. For PWA only, genre also exerted an influence on argument structure, with QOL narratives eliciting significantly less complex argument structures. This has implications for clinical assessment. Discourses of different genres should be sampled to fully assess a speaker's syntactic and lexico-semantic skills; and the genre of discourse used for assessment and therapy materials should align with the client's communication goals.

Background

Discourse is a functional communication skill, reflecting communication in everyday situations. Although Halliday (2004) defines discourse as one person producing a monologue used for a specific purpose, in everyday situations discourse appears in more than one guise, such as when people in conversation produce sequences of smaller discourses, or when one person temporarily takes the floor to produce a monologic discourse such as describing a scene, or talking about a recent event. Discourse production is complex and multifaceted process (Levelt, 1989; Sherratt, 2007), with each processing level offering an opportunity for the discourse to be shaped by context.

Discourse is a subject of increasing interest for both assessment and intervention studies in aphasia (Bryant, Ferguson & Spencer, 2016; Pritchard, Hilari, Cocks & Dipper, 2017) and the field of study has reached a point at which discourse is no longer considered a single entity with a single set of linguistic properties. Discourses are heavily influenced by the context in which they are produced (Eggins & Martin, 1997), including social and pragmatic influences; conventions in the way information is organized in different situations; and tendencies in the way language is used for different purposes. This variability poses a challenge to the field of aphasiology because, with the exception of fictional narratives, there is only a sparse evidence base to indicate which linguistic patterns and properties to expect in any given genre and similarly limited information on the range of normal performance. Nevertheless, a variety of discourse genres are used to elicit clinical discourse samples. This poses a problem for evaluating the discourse produced by people with aphasia because there is insufficient evidence to distinguish the effects of discourse genre from discourse impairment. If, for example, the discourse genre used in clinical or research assessment tasks prompts specific verb semantics or specific syntactic structures, there is the potential to skew the profile of a person's linguistic strengths and difficulties.

Discourse and Genre

Discourse is used for different functional purposes. For example, we can use discourse to share a fairy tale, a story from a holiday, describe a beautiful view, argue about politics, and give instructions. To produce an appropriate discourse in each of these situations, a speaker must: identify a suitable opportunity to produce it; select the 'shape' it needs to take; select information to include and information to omit; and then encode this information linguistically (Levelt, 1989; Sherratt, 2007).

Discourse is not produced without a context, and each level of discourse production is likely to be significantly influenced by the situation in which it is produced as well as by its intended purpose (Halliday, 2004). Because of this, context is a key driver for all decisions around the language used in a discourse, including sentence structure and verb selection, although such a relationship is probabilistic rather than deterministic. Whilst there is nothing to stop a speaker producing any verb or syntactic structure in a given context, they are unlikely to produce it if doing so is significantly atypical.

Register and genre theory (Eggins & Martin, 1997) is useful for the study of discourse in aphasia because it explicitly addresses the differences and similarities between discourses produced in different contexts or genres. A key stage in analysis using this approach is the description of the linguistic patterns in each discourse that produce meaning and genre effects. The most commonly used set of analytical tools used in such an approach come from Halliday's framework (e.g. Halliday 2004), in which language is viewed both as a system for making meaning (what does a person want to say?), and as a device to effect a particular purpose (why do they want to say it?). Halliday's analytical framework thereby allows an exploration of what language does and how it does it in any given discourse. One aspect of this analysis is to characterize meaning in terms of 'semantic process', for example to categorise verbs into five process types according to their meaning (*material*, *mental, relational*, *behavioral,* and *verbal*, reflecting meanings related to doing, thinking or feeling, being, physiological processes, and saying). These categories are defined in depth in Halliday (2004) and summarised in Table 1.

Considering Halliday's verb semantic types, and the definitions and examples in Table 1, we would expect all verb types to be used across a range of discourse genres. Halliday (2004) indicates that relational and material verbs are the types most used in English and so we would expect this pattern to hold across genres. For example, in a complex picture description of the type commonly used in aphasia assessment batteries, where a person is describing visually depicted concrete actions, *material* verbs which have meanings related to 'doing' would be expected to occur frequently: '*The man is flying a kite. The woman is pouring coffee into her mug'* and *relational* verbs of 'being' would also be expected to be prevalent: *'It's a park with grass, trees and a lake. There is someone on the pier'*. Similarly in personal narratives in which a reflective account is being produced, we might also expect to see *material* verbs occurring frequently: *'I worked in town, and walked to the office whenever I could.'* As well as frequent use of *relational* verbs: '*I'm keen on big celebrations. Family time is special. It's the time spent together and the memories that are so important'.*

Within this overall pattern, however, we would also expect some genre effects. Given that each discourse will have its own purpose and topic, we would expect some difference in the relative proportions of verb types between genres. For example, although both picture descriptions and personal narratives will contain high proportions of material and relational verbs, we might hypothesise that the picture description will elicit relatively more relational verbs than the personal narrative because of the visual-spatial nature of the pictured scene. By contrast, we might hypothesise that *mental* verbs, which have meanings related to sensing, thinking and feeling, would feature more in personal narrative responses to quality of life questions than the picture: *'We love our garden. I know how much time it takes but I think it is worth it.' My children think so too*.' Although there is no evidence base to provide explicit guidance about the verb types anticipated in each discourse genre, working this through using expectations based on general knowledge provides a useful basis for hypothesizing about the kinds of language a speaker is more likely to use.

----------Table 1 about here----------

Discourse, language, and aphasia

The current study aimed to explore how speakers use language in discourse of different genres. Language is impaired in the discourse produced by people with aphasia (PWA) (see reviews by Bryant et al., 2016; Linnik, Bastiaanse, & Höhle, 2016). In particular, PWA have marked deficits in their verb usage (e.g., Conroy, Sage & Lambon-Ralph, 2006; Kohn, Lorch & Pearson, 1989). In turn, a difficulty with verbs may affect how effectively speakers can communicate. Armstrong (2001) found that although some PWA use similar patterns of verb use to NHP, some PWA may have a restricted semantic range of verbs, leading to a restricted variety of meanings conveyed in discourse, and for some speakers, restricted communicative functions. In addition, PWA were less able to express their opinions than speakers with aphasia and used more high frequency or general mental verbs (Armstrong, 2005). The evidence base is complex, however. For example, although Cruice, Pritchard, & Dipper (2014) found that PWA used a more limited range of verbs and used the same verbs more frequently than NHP (using type/token analysis), they also found both speakers groups produced similar quantity, weight and semantic type of verbs overall.

Sentence structure is a challenge for PWA. PWA have more difficulty with verbs requiring increasing/complex argument structures compared to control populations (Whitworth, Claessen, Leitão, & Webster, 2015). Verb use and argument structure are likely to be linked, as some verbs are likely to require more complex sentence structure. PWA with semantic verb impairments often produce fewer/short sentences relying largely on noun phrases in comparison to PWA without verb impairments (Berndt, Haendiges, Mitchum, & Sandson, 1997), suggesting that difficulties with producing argument structure arise from verb deficits. As a result, in spontaneous speech, PWA with verb deficits often produce more single phrases than controls (Berndt et al., 1997; Webster, Franklin, & Howard, 2007), tend to rely on one-/two-argument structures according to Webster et al. (2007; which in the current study equates to verbs with no internal arguments), and use fewer complex structures than controls (Cruice, et al., 2014; Malyutina & den Ouden, 2016; Webster et al., 2007). Malyutina and den Ouden (2016) suggested that some or all of this pattern can be accounted for by psycholinguistic factors, such as higher frequency and shorter verbs chosen, leading to particular argument structures. Webster, Franklin and Howard (2004) suggested that noun retrieval issues may also impact on the production of complex argument structures. It is therefore likely that realization of sentence structure is affected by a range of different features, and that sentence structure is likely to be impaired in speakers with a range of different aphasia profiles.

One way of measuring verb use is examining the range of meanings a speaker is able to convey using a verb. For example, we can focus on the kind of information a verb communicates (its 'semantic category') and the amount of information communicated by the verb (its 'semantic weight'). Verbs can be classified as 'light' (e.g. do, go, come, get) or 'heavy' (e.g. fly, cook, surf) (Berndt et al., 1997). There is a common perception evident in the literature that speakers with aphasia tend to rely on light verbs, but no such consensus arises from the evidence base (Barde, Schwartz, & Boronat, 2006; Berndt et al., 1997; Breedin, Saffran, & Schwartz, 1998; Gordon, 2008) and furthermore there is only one published study (Cruice et al., 2014) which compares semantic weight in aphasic versus control speakers. Cruice and colleagues explored the impact of aphasia on speakers' capacity to produce personal narratives by examining 58 speakers' responses to questions about their quality of life (29 aphasic speakers; 29 non-aphasic speakers). Both speaker groups produced similar proportions of heavy and light verbs in this discourse context. The authors suggest that the quality of life narrative genre influenced these findings by providing broader linguistic opportunities than more commonly used clinical elicitation methods (including picture description). One aim of the current study was to

investigate this possibility further by directly comparing the verbs used in responses to quality of life questions with those used for picture descriptions.

Discourse, genre, and aphasia

The most common way to elicit discourse is through picture description tasks (Bryant et al., 2016a), which may lead to an increased focus on 'concrete' action verbs such as 'go' and 'eat' (e.g., Kim & Thompson, 2000). A number of studies suggest that different discourse genres affect the quality and quantity of discourse produced (Coelho, 2002; Olness, 2006; Olness 2007; Olness, Ulatowska, Wertz, Thompson, & Auther-Steffan, 2002; Van Leer & Turkstra, 1999). For example, Olness (2006) found differences in the discourses produced in response to two different kinds of picture description (pictures depicting a specific event sequence narrative vs. complex pictures of the type commonly used in aphasia assessment batteries). Olness found that pictured event sequences were likely to elicit discourses where the verbs were produced in the past tense, and which contained key narrative elements (e.g., background or setting information). By contrast, pictures with no event sequence were more likely to elicit verbs in the present tense, and discourse which communicated fewer narrative functions. These findings indicate that the type of picture stimulus affects the language used in a discourse, and suggest that genre differences might result in similar effects on language.

Genre is also likely to impact the syntactic complexity of language. Ulatowska and colleagues (Ulatowska, Freedman Stern, Weiss Doyel, Macaluso-Haynes, & North, 1983a; Ulatowska, Doyel, Stern, Haynes, & North, 1983b) report on what appear to be the same participants in two separate studies (indicated by identical size participant groups; identical age and education information, and for the group with aphasia, identical time post onset, and aphasia information), in which two different discourse genres are used. Although the studies do not aim to explicitly compare performance across genres, comparisons are possible by considering the results of each study together. This collective perspective from both studies suggests that for both PWA and NHP, the same pattern of more complex syntax (measured via clauses per t-unit, percent of dependent clauses, and percent of non-finite clauses), was used in narrative discourse than in procedural discourse.

In Cruice et al.'s (2014) exploration of a single genre (responses to quality of life questions), both speaker groups produced discourses which were similar in terms of quantity, weight, and type of verbs but which differed in terms of structural complexity. In comparison to NHP, the PWA group had significantly lower mean predicate argument structure scores, and produced discourses with significantly more 0 argument structures and fewer 1 argument structures. However, although this study provides us with evidence that PWA use less complex predicate argument structure than people without language impairment in this discourse genre, what we do not yet know is whether genre impacts on this complexity. The current study extends previous single-genre work, by comparing key aspects of the language used in two different discourse genres (picture description and personal narrative).

Halliday's 'process' types characterizing the verbs in utterances have been used to describe the discourse of PWA (Armstrong, 2001; Armstrong, 2005; Cruice et al., 2014). Findings from these studies have varied, with Armstrong (2001) finding that use of mental and relational verbs by some PWA differed to those used by NHP, whilst Armstrong (2005) and Cruice et al. (2014) found that PWA and NHP used similar verb types. To date, no study has directly explored how genre affects speakers' verb semantics in discourses of different genres.

The current study

The current study aimed explore the differences between two speaker groups (speakers with and without aphasia) across two different discourse genres: a widely-used picture description in a commonly used aphasia battery, the *Western Aphasia Battery-Revised* (WAB-R; Kertesz, 2006); and a personal narrative gained from reflective responses to quality of life (QOL) questions (Cruice, Hill, Worrall, & Hickson, 2010), to answer the research question '*how do two parallel groups of speakers use syntax and semantics across two discourse genres, in terms of argument structure, verb weight, and verb semantic type?'.* Picture description and responses to QOL questions were selected for comparison, as picture descriptions are the most widely used tool to elicit discourse in speakers with aphasia (Bryant 2016b; Linnik et al., 2015), but language beyond basic communication (such as might be used to respond to QOL questions), has been described by PWA as a communication priority by Worrall and colleagues (2011, p314): "*No. Needs, yes, but talk. . . my [points to head], I want to talk is politics and religion"*.

Very limited published work in the field of aphasia has described the differences between genres, making it problematic to hypothesise about the linguistic differences that might be expected in the different genres under investigation here. For this reason, the hypotheses outlined below are motivated by linguistic and genre theory (Eggins & Martin, 1997; Halliday, 2004), combined with the following assumptions about the language prompted by each discourse context: 1) that to succeed at a picture description task, a speaker must use language to communicate about a scene which has explicit/ concrete targets, for example, describing what people are doing; and 2) that to succeed at responding to questions about QOL, a speaker must communicate key aspects of QOL as they conceptualise it, and therefore use language to indicate their own evaluations. This content is likely to be highly individual, abstract, and reflect a speaker's sense of self, feelings, values, and attitudes.

Based on these assumptions, we hypothesised the following:

1. The differences in complexity of information in the two discourse genres would be reflected in the same pattern of more structurally complex language in the responses to QOL questions than in the picture description, so we would predict a between-genre difference. Also, structural complexity would be impacted by aphasia, based on previous work suggesting PWA have more difficulty than NHP (Cruice et al 2014) so we would predict a between-group difference.
2. Both discourses would elicit similar proportions of semantically 'heavy' verbs in speakers with and without aphasia, because these verbs are likely to be needed for both picture description and in answer to questions about QOL. We anticipated that in the picture description discourse,

these verbs might convey information about what people in the scene were doing, using verbs such as *running, fishing, standing, building*; and in the QOL questions, we anticipated that these might convey opinions and feelings, using verbs such as *think*, *feel*, and *know*. For heavy verbs we predict no between-genre not between-group difference.

We also hypothesised that the semantic process types reflected in the verbs that speakers used would differ between the discourses:

3. For both speaker groups, the increased focus on thinking and reflecting in the QOL questions would elicit a greater proportion of *mental* verbs than the picture description discourse. We therefore predict a between-genre (but not between-group) difference.
4. For both speaker groups, the visual-spatial nature of the scene prompting the picture description discourse would elicit more *relational* verbs locating people and objects than the personal narrative responses to the QOL questions. We therefore predict a between-genre (but not between-group) difference.

Method

Recruitment

Data used the current study were from two sources: Cruice (2002) in Australia, and Walkden in the UK. Cruice (2002) recruited from university aphasic clinics, three metropolitan hospital speech pathology departments (discharged patients), community stroke groups, and the state stroke association. This formed a part of a larger study on QOL in aphasia (Cruice et al, 2010). Inclusion criteria were that participants had aphasia at the time of stroke and ongoing aphasic difficulties; had reliable yes/no response and moderate comprehension at time of interviewing (determined by clinical assessment scores); were more than 12 months post-onset; and had no concomitant neurological disease (confirmed by hospital file checks, clinical observation, and self-report at interview). Walkden recruited neurologically healthy participants (NHP) for the current study from two privately-run nursing homes based in the north west of England. Inclusion criteria were that participants did not have depression; had no self-reported neurological history; and matched as closely as possible to the PWA group demographics in terms of age and educational background.

Ethics approval was granted to Cruice (2002) by the ethics committees of the three hospitals involved in recruitment of participants with aphasia and the university. Ethics approval for collecting data from NHP and completing the analyses in the current study was granted to Walkden and co-authors (LD and MC) by the university department proportionate review ethics committee in October 2011.

Materials

Screening measures

All participants completed the *Geriatric Depression Scale* (GDS: Brink & Yesavage, 1982) to screen for normal emotional health. Participants were excluded from the study if they obtained scores between 5-15, indicating depression. PWA also completed the 'comprehension' subsection of the *Western Aphasia Battery* to screen for moderate comprehension levels (cut-offs: ≥16/20 on the Yes/ No section; and ≥5/10 overall for the comprehension subtest). Diagnosis of aphasia at stroke onset and self-reported ongoing difficulties were used to identify aphasia.

Discourse samples

Participants completed two discourse tasks from two different genres: the picture description task from the *Western Aphasia Battery*, elicited using the prompt question '*Can you tell me what's going on in this picture?*' and six open-ended QOL questions (Cruice et al., 2010), which were considered as one sample.

(1) How would you describe the quality of your life? And why do you say that?

(2) What things give your life quality?

(3) What things take quality away from your life?

(4) What would make the quality of your life better?

(5) What would make the quality of your life worse?

(6) Does communication have an impact on the quality of your life? If yes, then how?

Analysis

What was analysed when

Discourse samples were audio recorded, and verbatim transcription took place from each recording. Transcription of the PWA was completed by one author, and transcription of the NHP was completed by another (see Table 2). All data was coded by one author, according to the procedures outlined below. A proportion of the data was also re-coded by a second author (see *Reliability* section below).

----------Table 2 about here----------

Analysis mirrored that completed for and reported in Cruice et al. (2014), using two semantic and one syntactic analyses:

1) Argument structure
2) Semantic weight
3) Semantic category

Argument structure. Argument structures were identified by identifying the main verbs or verb groups in transcripts, and then the internal arguments of the verbs. Therefore, subject noun phrases were excluded from analysis, and not included in any numerical data reported, for example, the 'I', in [I] go [to the park]. Verbs were then categorised as having 0, 1, and 2 internal arguments (matching the 1, 2 and 3 arguments identified in Byng & Black, 1989, and Webster et al., 2007, both of whom counted the verb phrase-external subject noun phrase as well as verb phrase-internal phrases). Complex arguments were coded as a single argument for example, [I] know [I need the rest]. Following this, the total number of predicates and arguments produced by each participant were tallied, and an average predicate argument structure (PAS) score was calculated for each participant, using a similar process to Webster et al. (2007), using the formula (total number of arguments produced/ total number of predicates produced). Scores from this calculation describe the average complexity of utterances produced.

Semantic weight. The main verb, which carried the weight of meaning in each utterance was identified, for example, 'am' in 'I am happy with my life', and 'getting' in 'we are getting much better at things', which served to remove auxiliary verbs from the analysis. The main verbs were then classified as 'light' or 'heavy', based on the analysis used by Berndt et al. (1997). The 'light' verbs include *go, be, do, have*, and are characterised by high frequency and communicating limited semantic information. By contrast, 'heavy' verbs communicate more semantic information, for example, *think, slump, inhale*. The number of verbs of each type in each discourse sample was tallied for each participant, and then converted into a percentage of their overall total verb use.

Semantic category. Main verbs were identified as described above, and then classified using the five process types, representing semantic categories, outlined by Halliday (2004). The number of verbs of each type in each discourse sample was tallied for each participant, and then converted into a percentage of their overall total verb use.

Statistics & planned comparisons

All numerical data were entered into an MS Office Excel spreadsheet, and analysed using IBM SPSS Statistics, version 24 for Mac. Descriptive statistics were generated, and outliers were removed (these were scores more than 2SDs above or below the mean) (see Appendices 1, 2, 3, and 4). Descriptive statistics from the new dataset (outliers removed) were then generated for inclusion in the *Results* section below.

A series of two-way mixed ANOVAs were completed, with a separate ANOVA for each measure, using the between-groups factor of groups, and discourse genre as the within-groups factor. Parametric tests were applied because more than 50% of the measures for the participant group were normally distributed, with distributions outside the range -1 to +1 considered skewed. *p* was therefore adjusted for multiple comparisons, using a Bonferroni correction to allow for this 0.05/ 4= 0.0125.

Reliability

Three of the transcripts for each participant group (approx. 10% of the data) were reanalysed by a second coder. Transcripts were randomly selected for re-coding using a random two-digit number generator. Agreement for the PWA picture description task, and the NHP picture description discourse and responses to the QOL questions is reported in Table 3, alongside reliability data from the analysis of the PWA responses to QOL questions are reported from Cruice et al. (2014). Agreement was excellent (agreement >80%) for each analysis.

----------Table 3 about here----------

Results

Participants

The PWA participant pool from Cruice (2002) (n=31) was examined against the inclusion criteria for the current study. One participant was excluded due to writing his responses, and four participants (#12, 13, 23, and 31) were excluded for only producing one of the two discourse samples. Therefore 26 PWA were included in the current study (Female= 14; Male= 12), with a mean age of 71.62 years (SD= 8.57, Range= 57-88), and a mean education of 10.65 years (SD = 3.75, Range = 6-20). Twenty-four PWA were below the WAB cut-off of 93.8, and the two PWA who exceeded this demonstrated a profile of ongoing anomic aphasia and reported negative impact of communication difficulties.

Thirty-three (33) NHP expressed an interest in the research and were assessed. Of these, six were excluded due to neurological/ cerebrovascular disease or because they did not produce both discourse samples (#9, 15, 16, 27, 32, and 33). Twenty-seven (27) NHP were therefore included in the current study (Female= 20; Male= 7), with a mean age of 85.41 years (SD = 5.96, Range = 67-92), and a mean education of 11.93 years (SD = 2.99, Range = 9-16).

Descriptive statistics

Full descriptive statistics for each measure are reported for PWA in Appendices 1 and 2, and for NHP in Appendices 3 and 4. Notably, for each measure, there was a wide range in scores and large standard deviations. Descriptive data indicate that for the PWA, the picture description genre (in comparison to the responses to the QOL questions) elicited more structurally complex language in terms of mean PAS score (Picture description PAS= 1.28; QoL PAS= 0.83), similar percentages of heavy verbs (Picture description= 55.75%; QoL= 50.94%), a smaller percentage of mental verbs (Picture description= 10.38%; QoL= 22.39%), and a similar percentage of relational verbs (Picture description= 39.32%; QoL= 35.82%) (Figure 1A).

For the NHP, descriptive data indicated that the picture description task (in comparison to the responses to QOL questions) elicited similar structurally complex language in terms of mean PAS score (Picture description PAS= 1.2; QoL PAS= 1.29); a similar percentage of heavy verbs (Picture description= 49.8%; QoL= 51.55%); a smaller percentage of mental verbs (Picture description=

12.66%; QoL= 22.45%); and a greater percentage of relational verbs (Picture description= 48.36%; QoL= 34.28%) (Figure 1B).

Inferential statistics

Comparisons

ANOVA were used to compare use of syntax and semantics in the two genres. The use of this parametric test was appropriate given that none of the NHP variables was skewed and only 3/8 PWA variables were skewed (Picture Description PAS; Picture Description % mental verbs; and QOL % behavioural verbs).

For PAS, there was an effect of genre F (1,51)= 14.81, p<0.0125; an interaction between genre and group F (1, 51)= 34.8, p<0.0125; and an effect of group F(1, 51)= 20.49, p<0.0125. For the % heavy verbs, there was no effect of genre F (1, 51)= 0.076, p>0.0125, no interaction between genre and group F (1, 51)= 1.78. p>0.0125, and no effect of group F (1, 51)= 0.03, p>0.0125. For the % mental verbs, there was an effect of genre F (1, 51)= 13.021, p<0.0125; no interaction between genre and group F (1,51)= 0.205, p>0.0125; and no effect of group F (1, 51)=0.47, p>0.0125. For % Relational verbs, there was no effect of genre F (1, 51)= 3.22, p>0.0125; no interaction between genre and group F (1, 51)= 2.29, p>0.0125; and no effect of group (1, 51)= 0.095, p>0.0125.

Planned post- hoc analysis using a paired t-test indicated that there was a difference between PAS scores for PWA t(25)= 5.26, p<0.0125, but not NHP t(26)= -2.27, p>0.00125.

----------Tables 4 and 5 about here----------

----------Figure 1 about here----------

Correlations

Given the above findings, a set of post-hoc Pearson product-moment correlation coefficients were computed to assess the relationships between the experimental variables of interest in our hypotheses (Table 6). Analyses revealed a positive correlation between mental verbs and heavy verbs used in QOL questions, for both groups (strong for PWA and borderline-moderate for NHP). There was also a strong negative correlation between the relational verbs and heavy verbs used by PWA in both genres. This relationship did not hold for the NHP group; instead the correlational analysis revealed a negative correlation between relational verbs and PAS score (borderline-moderate for picture description, and moderate for QOL responses).

----------Table 6 about here----------

Summary of Results

In the context of large ranges of scores and standard deviations, three findings were noted. Overall, there was no impact of genre or speaker group on the semantic weight of verbs (% heavy) that speakers used. For PAS, there was a main effect of genre, interaction of group and genre, and of group, suggesting that aphasia and discourse genre influence the semantic complexity of language speakers use. For the mental verbs, there was a main effect of discourse genre only, reflecting that there were more mental verbs used in response to quality of life questions than in describing the picture. There were also patterns of association between some of the variables: between mental and heavy verbs for both speaker groups for the QOL questions only; between relational and heavy verbs in both discourse genres produced by PWA; and between relational verbs and PAS in both discourse genres produced by NHP.

Discussion

The current study compared how people with aphasia and neurologically healthy speakers used syntax and semantics when producing two different discourse genres. The first of our hypotheses was rejected because the responses to QoL questions did not elicit structurally more complex language than the picture descriptions. In fact, the reverse was true: the more structurally complex argument structure (PAS) was produced for the picture descriptions. Post-hoc analysis indicated that aphasia caused this genre effect, as the difference was significant for the PWA but not the NHP group. The second hypothesis was fully supported in that both discourses elicited similar proportions of semantically 'heavy' verbs. The two hypotheses relating to semantic category were partially supported, with pictures eliciting a significantly smaller proportion of mental verbs than the responses to QoL questions, for both groups; and with pictures showing a non-significant trend towards eliciting a greater proportion of relational verbs than the Qol responses, but only for the NHP.

We had predicted a difference in the structural complexity of the language that would be prompted by each discourse genre, anticipating more structurally complex language in the quality of life responses than in the picture description for both groups. This was not the case. The descriptive statistics and post-hoc analysis indicated that the NHP produced structurally similar language in both discourses, and the PWA produced more structurally complex language in the picture description discourses than in the quality of life discourses. Starting with the lack of difference in structural complexity for the NHP, two possible explanations are proposed. Firstly, our hypothesis was motivated by a synthesis of the results of two separate studies by Ulatowska and colleagues (Ulatowska et al., 1983a; 1983b), in which more complex syntax was used in narrative discourse than in *procedural* discourse, so the unexpected findings may be due to the different discourses genres under investigation. It is possible that procedural discourses prompt structural difference (i.e. less complex language) but that other discourse genres do not do so significantly. Other studies have also found that procedural narratives have specific attributes, including structurally simpler language and particular grammatical constructions not found so often in other genres (Pritchard, Morgan, Dipper & Cocks, 2015; Shadden, Burnette, Eikenberry, & Dibrezzo, 1991; Ulatowska, North, & Macaluso-Haynes, 1981). Secondly, Ulatowska and colleagues (1983a & 1983b) used a number of different syntactic complexity analyses, including clauses per t-unit, whilst the current study only used mean PAS complexity. However, for either explanation, a similar effect on the data from the both NHP and PWA groups would have been expected, which is not demonstrated here.

Turning next to the reverse finding for PWA, we found that this group produced more structurally complex language in the picture description discourses than in the quality of life discourses. A possible reason for this is the nature of the visual stimulus in the former, which may scaffold complex language by providing a prompt about the content that can be included. For PWA, in the picture description discourses, the mean PAS score is increased by a relatively larger number of V + 2 argument structures which, in our data, are generally verbs followed by a noun phrases and a prepositional phrase. These two internal argument phrases describe the spatial and physical properties of the picnic scene, for example 'a man on a pier', 'trees all around there'. A second explanation is that the syntax of PWA is less robust and therefore may be compromised when a speaker is also faced with processing the higher conceptual demands of the quality of life questions.

As hypothesised, speakers used similar proportions of heavy (and light) verbs regardless of genre. This outcome was expected because semantically light verbs are common in everyday discourse for both people with unimpaired language and those with aphasia (Cruice et al., 2014) and because semantically heavy verbs are likely to be needed for both picture description and in answer to questions about QOL. As we had anticipated, in the picture description discourse, the heavy verbs conveyed information about what people in the scene were doing, using verbs such as *running, fishing, standing, building*; and in the QOL questions, the heavy verbs (such as *think*, *feel*, and *know*) helped convey opinions and feelings. Although there is a perception evident in the literature that speakers with aphasia use more light verbs, there is no consensus arising from the evidence base, and only one published study compares light verb use in aphasic versus control speakers in which no such difference was found. Cruice et al. (2014) found no difference between speaker groups in terms of semantic weight for personal narratives, and the present study adds to this by also providing evidence of no difference in verb weight between speaker groups for picture descriptions. So, neither genre nor aphasia appear to impact on semantic weight.

The hypotheses about the relative proportion of verb semantic types differing across the discourse genres were partially upheld. Firstly, we argued that the increased focus on thinking and reflecting in the QOL questions would elicit a greater proportion of mental verbs than the picture description discourse, and this was the case for both speaker groups. In the quality of life discourses, both groups used mental verbs such as '*like*', and '*know*'; the PWA used verbs such as '*think*', and '*suppose*'; and the NHP used verbs such as '*expect*' and '*hope*'. In line with this finding, post-hoc analysis revealed a positive link between proportions of mental verbs and heavy verbs, further reinforcing the suggestion that personal narratives produced as reflective responses to quality of life questions require speakers to use more complex semantics.

We also argued that the visual-spatial nature of the scene prompting the picture description discourse would elicit more relational verbs locating people and objects than the personal narrative responses to the QOL questions, but this was not the case for either group. Inspection of the data reveals that both groups used relational verbs in the predicted way to describe the picture ('*he's on a pier*', '*they're in a park*', '*it's by the sea*'), but they used them in equal proportion in their responses to quality of life questions. Neither genre not group had an impact here.

The post-hoc analysis revealed a negative association between the proportion of relational verbs and heavy verbs used by PWA across genres, indicating that either more relational verbs co-occurred with less semantic weight or that the same verb type (light relational verbs) is being identified in both analyses. The latter option reflects that the PWA were relying on the semantically light verb '*be*' to convey relational meanings, and although they did produce some semantically heavy relational verbs (e.g. '*seem*'), they did so in smaller proportions than semantically light ones. The lack of a similar correlation for the NHP, suggests that this latter group had a wider range of relational verbs to draw from, some of which were semantically heavy (such as '*feel*') and they produced the heavy ones in proportions equal to semantically light relational verbs.

Theoretical implications

Our hypotheses were driven by the language framework described by Halliday (2004) and cited within the theoretical discourse production model proposed by Sherratt (2007). These frameworks and models suggest that language is likely to be driven by the context in which it occurs, which in this study led us to hypothesise that the syntax and semantics of language would differ between the two discourse genres: the picture description, and the personal narratives produced in response to quality of life questions. Overall the findings indicated that genre influenced language as we had expected, however the finding was stronger for semantics than for syntax. Only for PWA did genre influence argument structure (with picture descriptions prompting more complexity than responses to questions about quality of life). The effect of genre on language relating to verb semantics, on the other hand, was seen in both groups (with questions about quality of life prompting more mental verbs than picture description). This finding is in line with evidence from the single-genre work in Cruice et al. 2014, where syntax differentiated speaker group more strongly than semantics. Taken together, these findings raise the possibility that semantics is the domain most likely to indicate genre effects whereas syntax is the domain most likely to identify impairment.

Limitations and next steps

The current study is limited, in that it uses participants from two distinct geographical areas (Australia and the UK), and of different ages (the NHP in the current study are older than the PWA group), which may have impacted on findings. These geographic and age differences could be an additional source of the group effects observed, in addition to the presence of aphasia. Furthermore, the study compared only two discourse genres, across a relatively small number of micro-linguistic discourse variables. Further research should aim to profile how more variables differ across discourse genres for neurologically healthy speakers and speakers with aphasia.

Conclusions and Clinical implications

The findings from the current study indicate that there are key differences between genres in terms of verb semantics, which has implications for assessment and therapy for PWA. The findings suggest that discourse is genre- driven, and that a speaker's discourse skills in one genre do not necessarily reflect those which might be present in another genre. Therefore, if a clinician or researcher is looking to fully assess a speaker's discourse skills, and in particular the lexico-semantics of their

discourse, then discourses of different genres should be sampled. If a client's discourse goals focus on a specific genre, then the discourse used in assessment should align with this, or there is likely to be a mismatch between a speaker's skills and performance, due to the genre difference.

References

Armstrong, E. (2000). Aphasic discourse analysis: The story so far. *Aphasiology, 14*(9), 875-892.

Armstrong, E. (2001). Connecting lexical patterns of verb usage with discourse meanings in aphasia. *Aphasiology, 15*(10-11), 1029-1045.

Armstrong, E. (2005). Expressing opinions and feelings in aphasia: Linguistic options. *Aphasiology, 19*(3-5), 285-295.

Barde, L., Schwartz, M., & Boronat, C. (2006). Semantic weight and verb retrieval in aphasia. *Brain and Language, 97*, 266-278.

Berndt, R., Haendiges, A., Mitchum, C., & Sandson, J. (1997b). Verb retrieval in aphasia 2: Relationship to sentence processing. *Brain & Language, 56*(1), 107-137.

Breedin, S. D., Saffran, E. M., & Schwartz, M. F. (1998). Semantic factors in verb retrieval: An effect of complexity. *Brain and language*, *63*(1), 1-31.

Brink, T. L., Yesavage, J. A., Lum, O., Heersema, P. H., Adey, M., & Rose, T. L. (1982). Screening tests for geriatric depression. *Clinical Gerontologist*, *1*(1), 37-43.

Bryant, L., Ferguson, A., & Spencer, E. (2016). Linguistic analysis of discourse in aphasia: A review of the literature. *Clinical Linguistics & Phonetics*, *30*(7), 489-518.

Bryant, L., Spencer, E., & Ferguson, A. (2016). Clinical use of linguistic discourse analysis for the assessment of language in aphasia. *Aphasiology*, 1-22.

Byng, S., & Black, M. (1989). Some aspects of sentence production in aphasia. *Aphasiology*, *3*(3), 241-263.

Coelho, C. A., Youse, K. M., & Le, K. N. (2002). Conversational discourse in closed-head-injured and non-brain-injured adults. *Aphasiology*, *16*(4-6), 659-672.

Conroy, P., Sage, K., & Lambon Ralph, M. A. (2006). Towards theory-driven therapies for aphasic verb impairments: A review of current theory and practice. *Aphasiology*, *20*(12), 1159-1185.

Cruice, M. (2002). *Communication and quality of life in older people with aphasia and healthy older people*. Unpublished doctoral dissertation. Division of Speech Pathology, University of Queensland, Australia.

Cruice, M., Hill, R., Worrall, L. & Hickson, L. (2010). Conceptualising quality of life for older people with aphasia. *Aphasiology*, *24*(3), 327-347.

Cruice, M., Pritchard, M. & Dipper, L. (2014). Verb use in aphasic and non-aphasic personal discourse: What is normal? *Journal of Neurolinguistics*, *28*, 31-47.

Eggins, S., and Martin, J.R. (1997). Genres and registers of discourse. In van Dijk, T. (Ed.), *Discourse as structure and process: Discourse studies: A multidisciplinary introduction*. London: Edward Arnold.

Gordon, J. (2008). Measuring the lexical semantics of picture description in aphasia. *Aphasiology, 22*(7-8), 839-852.

Halliday, M. (2004). *An Introduction to Functional Grammar*. London: Edward Arnold.

Kertesz, A. (2006). *The Western Aphasia Battery WAB-R*. London: Grune and Statton.

Kohn, S. E., Lorch, M. P., & Pearson, D. M. (1989). Verb finding in aphasia. *Cortex*, *25*(1), 57-69.

Kim, M., & Thompson, C. K. (2000). Patterns of comprehension and production of nouns and verbs in agrammatism: Implications for lexical organization. *Brain and Language*, *74*(1), 1-25.

Levelt, W. (1989). *Speaking: From Intention to Articulation*. Cambridge: MIT Press.

Linnik, A., Bastiaanse, R., & Höhle, B. (2016). Discourse production in aphasia: a current review of theoretical and methodological challenges. *Aphasiology*, *30*(7), 765-800.

Malyutina, S., & den Ouden, D. B. (2016). What is it that lingers? Garden-path (mis) interpretations in younger and older adults. *The Quarterly Journal of Experimental Psychology*, *69*(5), 880-906.

Olness, G. S. (2006). Genre, verb, and coherence in picture-elicited discourse of adults with aphasia. *Aphasiology*, *20*(2-4), 175-187.

Olness, G. (2007) *Narrative Structure in Aphasia: Impact of Task and Aphasia Severity,* Clinical Aphasiology Conference, Scottsville, Arizona, May 22-26, 2007.

Olness, G. S., Ulatowska, H. K., Wertz, R. T., Thompson, J. L., & Auther, L. L. (2002). Discourse elicitation with pictorial stimuli in African Americans and Caucasians with and without aphasia. *Aphasiology*, *16*(4-6), 623-633.

Pritchard, M., Hilari, K., Cocks, N., & Dipper, L. (2017). Reviewing the quality of discourse information measures in aphasia. *International Journal of Language and Communication Disorders*, early online.

Pritchard, M., Morgan, G., Dipper, L., & Cocks, N. (2015). Language and iconic gesture use in procedural discourse by speakers with aphasia. *Aphasiology, 29*(7), 826-844.

Shadden, B., Burnette, R., Eikenberry, B., & Dibrezzo, R. (1991). All discourse tasks are not created equal. In T. Prescott (Ed.), *Clinical Aphasiology, 20*, pp. 327-341. Austin TX: Pro-Ed.

Sherratt, S. (2007). Multi-level discourse analysis: A feasible approach. *Aphasiology*, *21*(3-4), 375-393.

Ulatowska, H., Freedmam Stern, R., Weiss Doyel, A., Macaluso-Haynes, S., & North, A. (1983a). Production of narrative discourse in aphasia. *Brain and Language, 19*, 317-334.

Ulatowska, H. K., Doyel, A. W., Stern, R. F., Haynes, S. M., & North, A. J. (1983b). Production of procedural discourse in aphasia. *Brain and Language*, *18*(2), 315-341.

Ulatowska, H., North, A., & Macaluso-Haynes, S. (1981). Production of narrative and procedural discourse in aphasia. *Brain & Language, 13*, 345-371.

Van Leer, E., & Turkstra, L. (1999). The effect of elicitation task on discourse coherence and cohesion in adolescents with brain injury. *Journal of Communication Disorders*, *32*(5), 327-349.

Webster, J., Franklin, S., & Howard, D. (2004). Investigating the sub-processes involved in the production of thematic structure: An analysis of four people with aphasia. *Aphasiology*, *18*(1), 47-68.

Webster, J., Franklin, S., & Howard, D. (2007). An analysis of thematic and phrasal structure in people with aphasia: What more can we learn from the story of Cinderella?. *Journal of neurolinguistics*, *20*(5), 363-394.

Whitworth, A., Claessen, M., Leitão, S., & Webster, J. (2015). Beyond narrative: Is there an implicit structure to the way in which adults organise their discourse? *Clinical Linguistics & Phonetics*, *29*(6), 455-481.

Worrall, L., Sherratt, S., Rogers, P., Howe, T., Hersh, D., Ferguson, A., & Davidson, B. (2011). What people with aphasia want: Their goals according to the ICF. *Aphasiology*, *25*(3), 309-322.

Figure 1: A) Percentage of verb types used in the two discourses by A) PWA and B) NHP. Error bars represent 1 SD

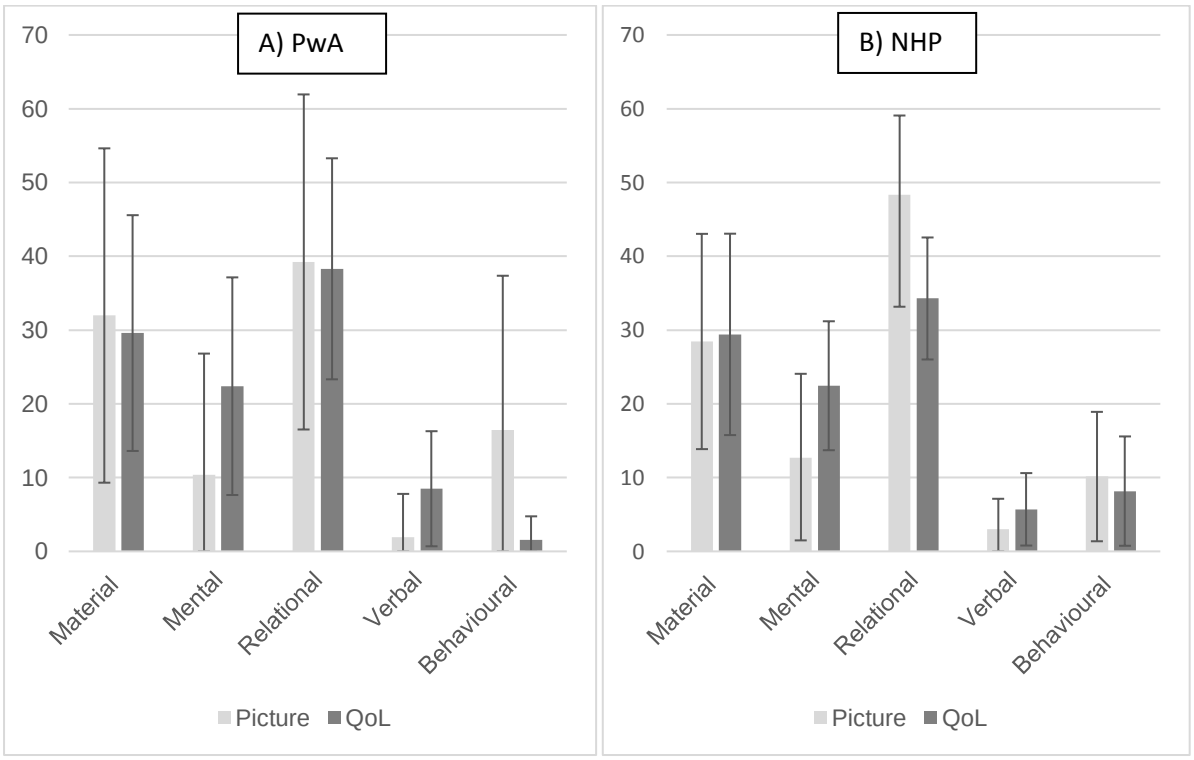

Table 1. Verb semantic process types (Halliday, 2004)

	Definition	Example
Material	Processes of **doing and happening**: change in the flow of events taking place through some input of energy; concrete action	He works too much I used to walk everywhere
Mental	Processes of **sensing**: our experience of the world of our own consciousness.	We love our garden I know it's different now
Relational	Processes of **characterising and identifying**	The man has a bottle Family is important
Behavioural	Processes of **physiological** or **psychological behaviour**	We dreamt of retiring early I never worry about the future
Verbal	Processes of **saying** and exchanging meaning	My children tell me to slow down I had to explain it all again

Table 2. Data and analysis in the current study

	PWA		NHP	
	Participants	Analysis	Participants	Analysis
Picture description	Cruice (2002)	Novel for the current study	Novel for the current study	Novel for the current study
Quality of life questions	Cruice (2002)	Results previously reported in Cruice et al. (2014)	Novel for the current study	Novel for the current study

Table 3. Inter-rater reliability levels for each analysis, participant group, and discourse sample

	PWA	NHP
Picture description (from the WAB)	Argument structure: 97.7% Semantic weight: 100% Semantic category: 97.1%	Argument structure: 96.3% Semantic weight: 97.5% Semantic category: 100%
Personal Narrative (reflective responses to quality of life questions)	Argument structure: 89% Semantic weight: 100% Semantic category: 98.6%	Argument structure: 93.1% Semantic weight: 100% Semantic category: 100%

Table 4: Information about participants in the current study.

	NHP	PWA
N	27	26
Male	7	12
Female	20	14
Mean age (SD; range)	85.57 (5.92; 67-92)	71.62 years (8.57; 57- 88)
Mean education (SD; range)	11.89 (2.94; 9-16)	10. 65(3.75; 6-20)

Table 5. Statistical comparisons for PWA

Measure	Genre	Mean	SD	Range	Statistical comparison between genres[1]
PAS	Picture	1.28	0.24	1-2	$t(22)=9.5$, $p<0.000001$[2]
	QOL	0.83	0.14	0.59-1.07	
% Heavy verbs	Picture	55.75	24.35	17.64-100	$t(24)=2.4$, $p>0.0125$[3]
	QOL	50.94	10.81	29.63-66.67	
% Relational verbs	Picture	39.32	22.71	0-70	$t(24)=0.58$, $p>0.0125$[4]
	QOL	35.82	14.98	4.88- 62.53	
% Mental verbs	Picture	10.38	16.43	0-62.5	$t(25)=-2.78$, $p<0.00001$[5]
	QOL	22.39	14.75	0-57.14	

[1] Differing degrees of freedom are the result of outlier removal
[2] Checked with non-parametric Wilcoxon test: z= 4.198, p<0.00001
[3] Checked with non-parametric Wilcoxon test: z= -0.74, p>0.0125
[4] Checked with non-parametric Wilcoxon test: z= -0.87, p>0.0125
[5] Checked with non-parametric Wilcoxon test: z= -2.89, p<0.0004

Table 6. Statistical comparisons for NHP

Measure	Genre	Mean	SD	Range	Statistical comparison between genres
Mean PAS complexity	Picture	1.2	0.14	1-1.5	*t*(26)= -2.27, p>0.0125
	QOL	1.29	0.16	1-1.65	
% Heavy verbs	Picture	49.8	11.77	28.57-75	*t*(24)= -0.51, p>0.0125
	QOL	51.55	7.65	35.9- 62.63	
% Relational verbs	Picture	48.36	10.72	33.33-75	*t*(24)= 0.58, p<0.0001
	QOL	34.28	8.27	19.51-48.49	
% Mental verbs	Picture	12.66	11.41	0-34.78	*t*(25)= -2.78, p<0.01
	QOL	22.45	8.74	7.35- 40.48	

Table 7. Correlations between discourse variables

		PWA			NHP		
		Picture description			*Picture description*		
		% Heavy	% Mental	% Relational	% Heavy	% Mental	% Relational
Picture description	Mean PAS	.14	-.00	-.19	.30	.27	-.48*
	% Heavy		-.09	-.78**		.23	-.31
	% Mental			-.24			.02
		PWA			NHP		
		Quality of life			*Quality of life*		
		% Heavy	% Mental	% Relational	% Heavy	% Mental	% Relational
Quality of life	Mean PAS	-.29	.09	.26	.06	.26	-.71*
	% Heavy		.68**	-.76**		.49*	-.01
	% Mental			.18			-.11

* p<.05; **p<.01

r: < 0.25= no correlation; 0.25-0.5= mild; 0.51- 0.74= moderate; >0.75= strong

Appendix 1. PWA: Responses to Picture Description task

		Total data set (n=26)				Outliers		Experimental variables with outliers removed				
		Total	Mean	SD	Range	N	Values	N	Mean	SD	Range	Skewness
Verbs produced		335	24.81	7.72	2-28	-	-	-	-	-	-	-
Mean PAS complexity		-	1.27	0.39	0-2.2	2	0, 2.2	24	1.28	0.24	1-2	1.43
#	0 argument structures	72	2.77	2.37	0-9	-	-	-	-	-	-	-
	1 argument structures	230	8.85	5.1	0-18	-	-	-	-	-	-	-
	2 argument structures	36	1.38	1.86	0-7	-	-	-	-	-	-	-
%	0 argument structures	-	24.16	21.14	0-100	-	-	-	-	-	-	-
	1 argument structures	-	67.58	21.03	0-100	-	-	-	-	-	-	-
	2 argument structures	-	8.26	9.9	0-29.17	-	-	-	-	-	-	-
#	Heavy verbs	166	6.39	4	1-18	-	-	-	-	-	-	-
	Light verbs	169	6.5	5.29	0-17	-	-	-	-	-	-	-
%	Heavy verbs	-	55.75	24.35	17.64-100	0	-	26	55.75	24.35	17.64-100	0.24
	Light verbs	35	44.25	24.35	0-82.35	-	-	-	-	-	-	-
#	Material	107	4.12	3.56	0-16	-	-	-	-	-	-	-
	Mental	26	1	1.39	0-5	-	-	-	-	-	-	-
	Relational	150	5.77	4.85	0-16	-	-	-	-	-	-	-
	Verbal	4	0.17	0.48	0-2	-	-	-	-	-	-	-
	Behavioural	48	1.92	1.82	0-8	-	-	-	-	-	-	-
%	Material	-	31.97	22.66	0-10	-	-	-	-	-	-	-
	Mental	-	10.38	16.43	0-62.5	-	-	26	10.38	16.43	0- 62.5	2

	Relational	-	39.23	22.71	0-70	-	-	26	39.32	22.71	0-70	-0.43
	Verbal	-	1.94	5.85	0-25	-	-	-	-	-	-	-
	Behavioural	-	16.46	20.89	0-100	1	100	25	13.13	12.33	0-40	0.75

Appendix 2. PWA: responses to QOL questions

		Total data set (n=26)				Outliers		Experimental variables with outliers removed				
		Total	Mean	SD	Range	N	Values	N	Mean	SD	Range	Skewness
Verbs produced		849	32.65	26.87	2-101	-	-	-	-	-	-	-
Mean PAS complexity		-	0.79	0.18	0.37-1.3	1	0.36	25	0.83	0.14	0.59-1.07	-0.15
#	0 argument structures	209	8.03	6.39	0-23	-	-	-	-	-	-	-
	1 argument structures	599	23.33	20.78	2-79	-	-	-	-	-	-	-
	2 argument structures	41	1.58	3.75	0-18	-	-	-	-	-	-	-
%	0 argument structures	-	24.37	14	0-63.64	-	-	-	-	-	-	-
	1 argument structures	-	72.47	14.71	36.36-100	-	-	-	-	-	-	-
	2 argument structures	-	3.18	5.35	0-20.93	-	-	-	-	-	-	-
#	Heavy verbs	417	16.03	13.12	0-4.9	-	-	-	-	-	-	-
	Light verbs	432	16.62	14.52	1-52	-	-	-	-	-	-	-
%	Heavy verbs	-	49.02	13.12	0-66.67	1	0	25	50.94	10.81	29.63-66.67	-0.37
	Light verbs	-	50.97	14.54	33.33-100	-	-	-	-	-	-	-
#	Material	270	10.38	9.27	0-29	-	-	-	-	-	-	-
	Mental	172	6.62	5.87	0-20	-	-	-	-	-	-	-
	Relational	309	11.88	11.15	1-43	-	-	-	-	-	-	-
	Verbal	78	3	3.7	0-14	-	-	-	-	-	-	-
	Behavioural	20	0.77	1.7	0-8	-	-	-	-	-	-	-
%	Material	-	29.01	16.06	0-72.5			-	-	-	-	-

	Mental	-	22.39	14.75	0-57.14	0	-	26	22.39	14.75	0-57.14	0.42
	Relational	-	39.21	18.74	5-100	1	100	25	35.82	14.98	4.88-62.53	-0.17
	Verbal	-	8.17	7.69	0-27.28			-	-	-	-	-
	Behavioural	-	1.53	3.09	0-11.76	2	12.5, 7.77	24	0.85	1.19	0-7.69	2.2

Appendix 3. NHP: Responses to Picture Description task

		Total data set (n=27)				Outliers		Experimental variables with outliers removed				
		Total	Mean	SD	Range	N	Values	N	Mean	SD	Range	Skewness
Verbs produced		554	17.19	9.66	4-51	-	-	-	-	-	-	-
Mean PAS complexity		-	1.2	0.14	1-1.5	0	-	27	1.2	0.14	1-1.5	0.59
#	0 argument structures	90	3.33	3.57	0-18	-	-	-	-	-	-	-
	1 argument structures	376	13.93	7.63	4-42	-	-	-	-	-	-	-
	2 argument structures	78	2.89	2.79	0-9	-	-	-	-	-	-	-
%	0 argument structures	-	15.59	9.14	0-33.33	-	-	-	-	-	-	-
	1 argument structures	-	70.17	12.48	30.43-92.31	-	-	-	-	-	-	-
	2 argument structures	-	12	9.28	0- 34.62	-	-	-	-	-	-	-
#	Heavy verbs	281	10.41	7.57	3-41	-	-	-	-	-	-	-
	Light verbs	272	10.07	5.69	1-28	-	-	-	-	-	-	-
%	Heavy verbs	-	51.13	13.46	28.57-85.71	1	85.71	26	49.8	11.77	28.57-75	0.06
	Light verbs	-	48.87	13.46	14.29-71.43	-	-	-	-	-	-	-
#	Material	142	5.26	2.31	0-11	-	-	-	-	-	-	-
	Mental	82	3.04	3.97	0-19	-	-	-	-	-	-	-
	Relational	257	9.52	5.67	0-26	-	-	-	-	-	-	-
	Verbal	17	0.61	0.84	0-3	-	-	-	-	-	-	-
	Behavioral	56	2.07	2.46	0-11	-	-	-	-	-	-	-
%	Material	-	28.45	14.59	0-66.67	-	-	-	-	-	-	-
	Mental	-	12.66	11.41	0-34.78	0	-	27	12.66	11.41	0-34.78	0.4

Relational	-	45.36	15.2	0-75	2	0, 14.29	25	48.36	10.72	33.33-75	0.86
Verbal	-	2.98	4.15	0-14.29	-	-	-	-	-	-	-
Behavioural	-	10.15	8.77	0-28.57	1	28.57	26	7.12	5.06	0-17.5	0.6

Appendix 4. NHP: Responses to QOL questions

		Total data set (participant n=27)				Outliers		Experimental variables with outliers removed				
		Total	Mean	SD	Range	N	Values	N	Mean	SD	Range	Skewness
Verbs produced		1450	53.7	24.88	14-112	-	-	-	-	-	-	-
Mean PAS complexity		-	1.29	0.16	1-1.65	0	-	27	1.29	0.16	1-1.65	0.6
#	0 argument structures	321	11.8	7.27	0-28	-	-	-	-	-	-	-
	1 argument structures	1101	37.44	18.32	7-82	-	-	-	-	-	-	-
	2 argument structures	99	3.8	2.9	0-10	-	-	-	-	-	-	-
%	0 argument structures	-	21.65	9.33	0-39.44	-	-	-	-	-	-	-
	1 argument structures	-	71.33	8.09	57.75-100	-	-	-	-	-	-	-
	2 argument structures	-	7.02	4.65	0-17.39	-	-	-	-	-	-	-
#	Heavy verbs	761	28.19	14.5	8-62	-	-	-	-	-	-	-
	Light verbs	689	25.52	11.7	3-52	-	-	-	-	-	-	-
%	Heavy verbs	-	52.55	9.13	35.9-78.57	1	30	26	51.55	7.65	35.9- 62.63	-0.73
	Light verbs	-	47.45	9.13	21.43-64.1	-	-	-	-	-	-	-
#	Material	438	16.22	10.04	0-39	-	-	-	-	-	-	-
	Mental	334	12.37	8.12	2-36	-	-	-	-	-	-	-
	Relational	492	18.22	9.17	4-48	-	-	-	-	-	-	-
	Verbal	75	2.78	2.42	0-9	-	-	-	-	-	-	-
	Behavioural	111	4.11	3.7	0-18	-	-	-	-	-	-	-
	Material	-	29.41	13.65	0-57.14	-	-	-	-	-	-	-

%	Mental	-	22.45	8.74	7.35-40.48	-	-	27	22.45	8.74	7.35-40.48	0.39
	Relational	-	34.28	8.27	19.51-48.49	0	-	27	34.28	8.27	19.51- 48.49	0.17
	Verbal	-	5.7	4.91	0-17.14	-	-	-	-	-	-	-
	Behavioural	-	8.17	7.42	0-35.71	1	30	26	7.1	5.07	0-17.5	0.45