Published in final edited form as:
J Cogn Neurosci. 2021 November 05; 33(12): 2494–2511. doi:10.1162/jocn_a_01772.

Content word production during discourse in aphasia: Deficits in word quantity, not lexical-semantic complexity

Reem S. W. Alyahya[1,2,3], **Ajay D. Halai**[1], **Paul Conroy**[4], **Matthew A. Lambon Ralph**[1]

[1]MRC Cognition and Brain Sciences Unit, University of Cambridge, Cambridge, United Kingdom

[2]King Fahad Medical City, Riyadh, Saudi Arabia

[3]Faculty of Medicine, Alfaisal University, Riyadh, Saudi Arabia

[4]Division of Neuroscience and Experimental Psychology, University of Manchester, Manchester, United Kingdom

Abstract

Although limited and reduced connected speech production is one, if not the most, prominent feature of aphasia, few studies have examined the properties of content words produced during discourse in aphasia, in comparison to the many investigations of single-word production. In this study, we used a distributional analysis approach to investigate the properties of content words production during discourse by 46 participants spanning a wide range of chronic post-stroke aphasia and 20 neuro-typical adults, using different stimuli that elicited three discourse genres (descriptive, narrative and procedural). Initially, we inspected the discourse data with respect to the quantity of production, lexical-semantic diversity, and psycholinguistic features (frequency and imageability) of content words. Subsequently, we created a ‘lexical-semantic landscape’, which is sensitive to subtle changes and allowed us to evaluate the pattern of changes in discourse production across groups. Relative to neuro-typical adults, all persons with aphasia (both fluent and non-fluent) showed significant reduction in the quantity and diversity of production, but the lexical-semantic complexity of word production directly mirrored neuro-typical performance. Specifically, persons with aphasia produced the same rate of nouns/verbs and their discourse samples covered the full range of word frequency and imageability, albeit with reduced word quantity. These findings provide novel evidence that, unlike in other disorders (e.g., semantic dementia), discourse production in post-stroke aphasia has relatively preserved lexical-semantic complexity but demonstrates significantly compromised quantity of content word production. Voxel-wise lesion-symptom mapping using both univariate and multivariate approaches revealed left frontal regions particularly the pars opercularis, insular cortex, and central and frontal opercular cortices supporting word retrieval during connected speech, irrespective of word class or their lexical-semantic complexity.

Correspondence to: Reem S. W. Alyahya.

Correspondence to: Reem S. W. Alyahya, rsalyahya@gmail.com, Communication and Swallowing Disorders Department, King Fahad Medical City, Riyadh, Kingdom of Saudi Arabia, P.O. Box. 59046, Riyadh 11525, Phone: (+966) 11 288 9999 Ext: 27620.

Keywords

aphasia; connected speech; lexical-semantics; discourse; nouns and verbs; lesion-symptom mapping

Introduction

Word retrieval deficits, which are common following brain damage, undermine single-word, sentence and discourse production; which impact the engagement in conversations. In comparison to the many investigations of single-word production, there have been fewer explorations of the properties of content words produced during discourse, how these might vary across sub-groups of aphasia (i.e., an acquired language impairment following brain damage), and whether the results depend on the elicitation stimuli. These enquiries were addressed in the current study.

Discourse production provides a rich data source to assess linguistic content/structure and language use in a more naturalistic context (Dipper & Pritchard, 2017). Different elicitation stimuli tends to elicit different discourse genres, and studies vary in the employed elicitation stimuli. Composite picture description is commonly used in clinical practice and research (e.g., Goodglass & Kaplan, 1983; Stark, 2019), due to advantages in time and consistency of responses across neuro-typical adults, resulting in a reliable baseline to compare the language profiles of patient groups. This task, however, might elicit more concrete words compared to other more natural modes of connected speech, such as storytelling narratives (Bird, Howard, & Franklin, 2000). Storytelling based on personal past experiences, familiar stories, or using sequential pictures has been employed in research more than clinical settings (Bird & Franklin, 1996; Stark, 2019). Personal narratives, however, can result in inconsistent responses at the lexical and semantic levels across and within individuals. Another natural mode of connected speech is to use procedural discourse (description of a familiar task, e.g., changing a car tyre) (e.g., Alyahya, Halai, Conroy, & Lambon Ralph, 2020b; Basilakos et al., 2014; Stark, 2019), although this has been utilised less frequently in the aphasiology literature and clinical practice compared to picture description and personal story-telling, as indicated by a systematic review on studies that utilised discourse samples from people with aphasia (Bryant, Ferguson, & Spencer, 2016). It has been shown that the elicitation stimuli might results in different language patterns in persons with aphasia. Specifically, there is evidence to indicate that storytelling probes greater word quantity, lexical diversity and propositional density than picture description in people with aphasia and neuro-typical adults (Alyahya et al., 2020b; Stark, 2019). However, no studies explored whether the elicitation stimuli is important to the varying psycholinguistic properties of the words produced during connected speech. This was addressed in the current study.

Discourse production relies heavily on the production of content words, especially of the two main word classes, nouns and verbs. Whilst persons with aphasia generally present with greater verb deficits compared to nouns, there are strong indications that these noun-verb differences at single-word level might be attributable to the semantic and psycholinguistic distinctions between them, with verbs tending to be semantically more complex and more

Europe PMC Funders Author Manuscripts

abstract than nouns (Alyahya, Halai, Conroy, & Lambon Ralph, 2018a; Vigliocco, Vinson, Druks, & Cappa, 2011). Word class effects have also been mapped onto the neural substrates in order to expand our knowledge of the language system. There are evidence to indicate that nouns and verbs are underpinned by shared neural correlates (e.g., Alyahya et al., 2018a; Crepaldi et al., 2013; Siri et al., 2008). The majority of studies that compared the neural correlates of nouns and verbs employ single-word tasks (naming or lexical generation). To date, one study has explored the neural correlates of nouns and verbs during picture description in post-stroke aphasia using network lesion-symptom mapping. This study suggested that object word production was supported by posterior networks across the occipital, parietal and posterior inferior temporal regions, while action word production was supported by frontal networks (Gleichgerrcht et al., 2016). The study, however, did not account for variations in the lexical-semantic properties of the object and action words.

There is a large body of evidence on the effect of lexical-semantic properties and particularly word imageability on single-word production (Alyahya, Halai, Conroy, & Lambon Ralph, 2018b, 2020a; Bird, Howard, & Franklin, 2003) and comprehension in post-stroke aphasia (e.g., Alyahya et al., 2018a; Sandberg & Kiran, 2014) with increased accuracy and efficient in processing concrete over abstract words, including nouns and verbs. Word frequency strongly influences single word and connected speech production in patients with semantic dementia (Bird et al., 2000; Hoffman, Meteyard, & Patterson, 2014). However, the effect of imageability and frequency during discourse production in aphasia remains relatively unexplored. Studies have reported that people with fluent and non-fluent aphasia mainly use light (high-frequency and semantically empty) verbs (e.g., *'go, do')* during sentence production (Berndt, Haendiges, Mitchum, & Sandson, 1997b), and they tend to produce more high-frequency and high-imageability verbs compared to neuro-typical adults during storytelling (Bird & Franklin, 1996).

To date, the differences in lexical-semantic properties (word imageability and frequency) of nouns and verbs during discourse production have not been examined in post-stroke aphasia, and across different discourse genres. This might be due to a methodological challenge with discourse data, which is how to quantify and analyse the pattern of vocabulary elicited. Perhaps most commonly, studies have measured deficits in fluency and diversity by coding the number and type of tokens (e.g., Alyahya et al., 2020b; Bastiaanse & Jonkers, 1998; Crepaldi et al., 2011). Other investigations in semantic dementia have adopted a more distributional analysis approach in which the production rate of words at different points along a psycholinguistic continuum are evaluated in controls and patients. This is a method that can be much more sensitive to subtle changes (Bird et al., 2000; Hoffman et al., 2014).

In this study, we adopted this distributional analysis approach to create a 'lexical-semantic landscape' of the vocabulary produced during discourse (with respect to word imageability and frequency). We examined how this landscape differ between a large cohort of people with a wide range of post-stroke aphasia severity/classifications, and a group of neuro-typical adults. We also explored differences between sub-groups of aphasia: fluent versus non-fluent aphasia, and high versus low performers on semantic and phonology domains. Additionally, we assessed the influence of different discourse stimuli (composite picture description, storytelling narrative, and procedural discourse) on the lexical-semantic

Europe PMC Funders Author Manuscripts

properties of the words produced during connected speech. By comparing the ‘lexical-semantic landscape’ (i.e., production rate across different levels of word frequency and imageability) it is possible to evaluate the pattern of changes in discourse production in people with aphasia and neuro-typical controls. Specifically, whether there is a reduction in the overall number of words produced and/or change in the shape of the lexical-semantic landscape. This approach has a second useful feature in that it is possible to see how different types of word (e.g., nouns and verbs) are located with respect to different levels of frequency and imageability within the ‘lexical-semantic landscape’. Finally, we mapped the neural correlates of word retrieval during connected speech within the lexical-semantic landscape using univariate and multivariate lesion-symptom mapping approaches. Although there are strong advocates for both univariate and multivariate lesion-symptom mapping approaches (Bates et al., 2003; DeMarco & Turkeltuab, 2018; Mah, Husain, Rees, & Nachev, 2014; Sperber, Wiesen, & Karnath, 2019; Tyler, Marslen-Wilson, & Stamatakis, 2005), these alterative analyses tackle different fundamental questions (localisation versus prediction modelling), and have opposite strengths and weaknesses (Haufe et al., 2014; Hebart & Baker, 2018). Previous studies have shown that (a) multivariate approaches do not always perform better than univariate approaches, (b) different multivariate approaches are not equal, and (c) using both univariate and multivariate approaches is more likely to be complementary as they can be used to answer different questions (Schumacher, Halai, & Lambon Ralph, 2019; Sperber et al., 2019; Zhao, Halai, & Lambon Ralph, 2020). Therefore, we present results from a univariate and two multivariate approaches.

Methods

Participants

Forty-six participants (32 males) who had developed aphasia following a single left haemorrhagic or ischaemic stroke were tested in the chronic stage (> 12 months post-stroke, mean = 69.43 months, SD = 48.86, range = 16 – 280), with a mean age of 63.21 years (SD = 11.93, range = 44 – 87), and mean years of education of 12.65 years (SD = 2.59, range = 9 – 19). The Boston Diagnostic Aphasia Examination (BDAE: Goodglass & Kaplan, 1983) was administered to each participant, and they were diagnosed and classified using the BDAE standard aphasia classification criteria. No restrictions were placed according to aphasia type or severity (spanning from mild anomia to global aphasia), with a mean severity of 2.8/5 (SD = 1.2, range = 1 – 5). The exclusion criteria included having suffered more than one stroke or any other neurological conditions, severe motor-speech disorders as described in the participant’s clinical workup, participants who did not produce any response in any of the discourse samples, any contraindications for MRI scanning, and being pre-morbidly left-handed. All participants were native English-speakers with normal or corrected-to-normal vision and/or hearing. Detailed demographic information is presented in Supplementary Table 1. Discourse samples were also collected from 20 age/education matched healthy/neuro-typical right-handed native English-speaking adults (9 males, mean age = 68.85 years, SD = 8.47, range = 57 – 84; mean years of education = 14 years, SD = 2.8, range = 9 – 19). All participants reported no abnormal neurological conditions or history of brain-injury. Informed consent was obtained prior to participation under approval from a local ethics committee.

Europe PMC Funders Author Manuscripts

Discourse samples: elicitation, transcription and coding

Three elicitation stimuli were employed with no time limit. First, a relatively simple and commonly used composite picture description was utilised to elicit a descriptive discourse using the ‘Cookie Theft’ picture from the Boston Diagnostic Aphasia Examination (BDAE: Goodglass & Kaplan, 1983). This was included because it is widely used in clinical assessments and research. Second, a storytelling narrative using the ‘Dinner Party’ pictorial script (Fletcher & Birt, 1983) was employed to elicit a more natural mode of connected speech with reduced working memory load compared (due to the presence of picture stimuli) compared to other forms of storytelling narratives. Participants were presented with this series of eight black-and-white picture sequences and were asked to look through them and then describe in detail what was going on in these pictures. Finally, participants were asked to provide a procedural discourse ‘how they prepare a cup of tea’, which is another natural form of connected speech but without the use of picture stimuli. No prompts or questions were provided throughout the testing by the examiner, except for nonverbal encouragement.

Discourse samples were digitally recorded and then transcribed verbatim (orthographically), and checked against the recording to correct for any discrepancies. This was followed by content analyses carried out on each transcript. Transcription, coding and analyses were conducted by the first author (RSWA), a qualified and experienced Speech and Language Pathologist. The following measures were computed from each transcript: (1) Content word counts, which included all words that were intelligible, informative and relevant to the discourse (adapted from Nicholas & Brookshire, 1993). Contractions (e.g., it’s or haven’t) were counted as two separate words; (2) number of nouns and lexical verbs (excluding auxiliary such as ‘is’ in ‘is going’ and modal verbs such as ‘should’ in ‘should go’) were extracted from the content word count, and used as a measure of quantity. The present tense form of the verb ‘to be’ was accepted as a lexical verb (i.e., ‘is’ in ‘she is happy’); (3) type counts (i.e., the number of distinct words) for nouns and lexical verbs, which were used as a measure of lexical-semantic diversity. Type counts were selected over ratios, because unlike ratios, they do not inflate estimates of lexical diversity (particularly in non-fluent aphasia); as shorter samples produced by persons with severe aphasia can often appear to be richer in diversity due to their higher ratios compared to longer samples. The use of ratios would restrict comparisons between participants and across discourses of different lengths.

The second step was to obtain the lexical-semantic properties (word frequency and imageability) for each noun and verb presented in each transcript (details are provided in Table 2). Frequency values represents lemma frequency per million words (combined written and spoken counts) obtained from the British National Corpus Consortium (2007). Imageability ratings were drawn from the MRC Psycholinguistic Database (Coltheart, 1981) and a corpus of published norms (Bird, Franklin, & Howard, 2001), as these databases include ratings for words within specified word class to disambiguates any noun-verb ambiguous words (e.g., ‘brush’). Frequency and imageability values were obtained from the respective database using an automated approach (LOOKUP function) implemented in MATLAB (2018a). Frequency values were obtained for all words, and imageability values were successfully obtained for 96.3% of the words in all transcripts (see details in

Europe PMC Funders Author Manuscripts

Supplementary Table 2). Only words with a defined imageability value were entered into the imageability analyses.

Statistical analyses

Preliminary analyses were conducted to inspect the nature of the discourse data in which the properties of the words produced during discourse samples was compared across the three discourse stimuli and between neuro-typical adults and persons with aphasia. The comparisons were also conducted within the aphasia group by splitting the group into participants with fluent aphasia (N = 25: anomia, conduction and transcortical sensory aphasia) versus non-fluent aphasia (N = 21: global, mixed non-fluent, Broca's and transcortical mixed aphasia). We expected participants with aphasia to produce fewer words than neuro-typical controls, and participants with non-fluent aphasia to produce fewer words that those with fluent aphasia. Specifically, we examined differences in the quantity, diversity and lexical-semantic properties of nouns and verbs, using mixed 2 × 3 × 2 ANOVAs on word count (to measure quantity of production), type count (to measure diversity of production), and imageability and frequency values (measures of lexical-semantic properties). In all ANOVAs, group (neuro-typical vs. persons with aphasia) was entered in the models as the between-subject factor, whereas discourse (storytelling narrative vs. picture description vs. procedural) and word class (nouns vs. verbs) were entered as the within-subject factors. All significant interactions were explored using t-tests post hoc analyses and Bonferroni corrected for multiple comparisons. Similar analyses were then conducted but the between-subject factor was set to persons with fluent versus non-fluent aphasia.

To allow for a more sensitive measure in capturing systematic variations between the groups and across the discourse genres, imageability and frequency distribution patterns were created (Fig.3/4). Specifically, the range of word imageability (and frequency) was divided into bands (presented on the x-axis of Fig.3/4) and the number of words produced within each of these bands was computed for each participant followed by computing the group mean. This distributional analysis approach indicates how often the participants produced words of different imageability and frequency range. Subsequently, the imageability and frequency bands were used to generate contour maps (Fig.5/6), i.e., a 'two-dimensional lexical-semantic landscape' of the words produced during discourse. The number of words produced within each of these frequency × imageability cell of the contour maps was computed for each participant and then we calculated the group mean (for both groups of persons with aphasia and neuro-typical adults). This was done for the discourse that best captured the imageability and frequency distributions as per the findings of the distributional analysis described above. Further analyses were conducted in order to examine whether the two-dimensional lexical-semantic landscapes are consistent across different aphasia sub-groups defined in accordance to three orthogonal aphasia-related domains: fluency, semantic and phonological ability. Specifically, we employed a multivariate decomposition algorithm (principal component analysis) on a detailed neuropsychological/aphasiological battery that assessed different aspects of language and cognitive skills, inducing repetition, naming, word retrieval, fluency, phonemic discrimination, semantic processing, single-word and sentence comprehension, working memory, and executive function. For more details on the

Europe PMC Funders Author Manuscripts

neuropsychological battery and our previous work on conceptualising aphasia as a graded multidimensional space, see Alyahya et al. (2018a). This data-driven approach maximises the amount of shared variance in a heterogeneous sample and accounts for systematic variations across tests. The principal component analysis generated components related to semantic, phonology and fluency. We used participants' components scores along these three components, and the whole aphasia group was then split into high versus low performers sub-groups based on the median value of the scores on each of these components.

Acquisition, processing and analyses of neuroimaging data

High-resolution structural T1-weighted MRI scans were acquired for each patient on a 3.0 T Philips Achieva scanner (Philips Healthcare, Best, The Netherlands) using an eight-element SENSE head coil. A T1-weighted inversion recovery sequence with 3D acquisition was utilised with the following parameters: repetition time = 9.0 millisecond (ms), echo time = 3.93 ms, acquired voxel size = $1.0 \times 1.0 \times 1.0$ mm^3, slice thickness = 1 mm, matrix size = 256_256, 150 contiguous slices, flip angle = 8, field of view = 256 mm, inversion time = 1150 ms, SENSE acceleration factor 2.5, total scan acquisition time = 575 seconds.

Participants' structural T1-weighted MRI scans were pre-processed and analysed with Statistical Parametric Mapping software (SPM12: Wellcome Trust Centre for Neuroimaging, http://www.fil.ion.ucl.ac.uk/spm/) running under MATLAB (2018a). The images were normalised into standard Montreal Neurological Institute (MNI) space using a modified unified segmentation-normalisation tool optimised for focal brain lesions (Seghier, Ramlackhansingh, Crinion, Leff, & Price, 2008). Structural scans from an age and education matched control group (18 male and 4 female; mean age = 69.13 years, SD = 5.85; and mean years of education = 13 years, SD = 2.66) were used as reference to identify abnormal tissue in the stroke group using a fuzzy clustering fixed prototypes (FCP) approach. This produces a whole brain map where each voxel is a probability of abnormality compared to the control group. We applied a binary threshold to this image to obtain a binary lesion image (i.e., U-threshold = 0.5). The images generated for each patient were visually inspected with respect to the original scan and manually corrected if necessary and were used to generate a lesion overlap map (Fig.1). Images were then smoothed with an 8mm full-width-half-maximum Gaussian kernel, in order to account for the global intra-individual shape differences.

To identify the neural correlates associated with word retrieval during connected speech within the lexical-semantic landscape, we conducted Voxel-Based Correlational Methodology (VBCM: Tyler et al., 2005), a variant of Voxel-Lesion Symptom Mapping (Bates et al., 2003), which identifies statistical relationships between brain and behaviour by correlating the value per voxel (as a continuous variable) with the behavioural performance. We created multiple regression models using the FCP whole brain images (% abnormality) with the behavioural measure of interest (noun and verb retrieval within the lexical-semantic landscape) entered as regressors. Lesion volume (estimated using the automated lesion identification tool; Seghier et al., 2008) and demographic variables (age, months post stroke onset, and years of formal education) were entered in the models as covariates to control for participant's variabilities on these variables, and due to the importance of these demographic

Europe PMC Funders Author Manuscripts

variables for stroke population. Specifically, age and education are related to brain health and general executive/language performance, and time post stroke can predict language performance in people with post-stroke aphasia (Hope et al., 2017). We used the statistical non-parametric mapping (SnPM) toolbox (version 13.1.08; http://warwick.ac.uk/tenichols/snpm) to create each multi-subject model and 5000 permutations where implemented to determine statistical inference. The results were thresholded at p<0.005 voxel-level and cluster corrected using family-wise error (FWE) at p<0.05.

To supplement the univariate analysis, we conducted multivariate analyses using two approaches. First, we used support-vector regression lesion symptom mapping (SVR-LSM) toolbox (DeMarco & Turkeltuab, 2018). In this approach, we loaded the binary lesion images as the features and created a separate model for noun and verb retrieval within the lexical-semantic landscape, and included with demographic variables (age, education, time post-onset) in all models as covariates. The following settings were used: MATLAB SVM implementation, hyper-parameter optimization (Bayes optimization with default settings) and lesion threshold = 3. The resulting beta weights were evaluated by permutation testing (n=10000) and thresholded at voxel-wise p<0.005 and cluster-wise p<0.05. We ran each behavioural model twice, with and without correction for lesion volume using the 'regress on both' option as recommended by DeMarco and Turkeltuab (2018). Second, we used the pattern recognition of neuroimaging toolbox (PRoNTo V2.1, http://www.mlnl.cs.ucl.ac.uk/pronto) (Schrouff et al., 2013) as an alternative multivariate approach because it formally evaluates model predictions, and does not truncate beta weights post hoc. For this approach, we entered the FCP (% abnormality) images as continuous values, which quantifies the amount of abnormality at each voxel across the whole brain including both the lesioned and intact hemispheres. We created separate models for noun versus verb retrieval within the lexical-semantic landscape, with demographic variables (age, education, time post-onset) entered in all models as covariates. We followed the pipeline through in two pathways: (i) restricted to lesion territory (similar to SVR-LSM); and (ii) using the whole brain as input (similar to the VBCM). Data from the whole brain was included in case there was any meaningful variations that might otherwise be missed when using a lesion mask. The models were estimated using the relevance vector regression (RVR: Tipping, 2001) implementation as this method is quick and does not require hyper-parameter optimisation. PRoNTo relies on kernel methods to overcome the high dimensionality problem in neuroimaging (using $n \times n$ pair-wise similarity matrix) and features were mean centred. A 10-fold cross-validation scheme was used to determine model performance (i.e., trained on 90% of the data and tested on 10% held out set). We also report r values that represent the cross validated correlation value of the models. The correlation is between the observed (true) scores and the model estimated (predicted) scores in the left out cases. Model inference, specifically, is evaluated using permutation testing whereby the observed values are shuffled randomly and the prediction model is re-calculated to obtain a distribution of model performance by chance (n=5iterations). The real cross-validated correlation is then compared to the null distribution to determine significance with a p < 0.05 alpha threshold. As with the SVR-LSM, we ran each model with and without lesion volume as a covariate. The anatomical labels were obtained using Harvard-Oxford atlas (Desikan et al., 2006).

Results

Table 1 provides information on the discourse data produced by the neuro-typical and aphasia groups.

The properties of the words produce during discourse

Descriptive statistics on the quantity and diversity of nouns and verbs produced during discourse are reported in Table 1 and illustrated in Fig.2. Descriptive statistics on the frequency and imageability values are reported in Table 2 for the groups of neuro-typical adults, all persons with aphasia, persons with fluent aphasia, and persons with non-fluent aphasia across the three discourse genres. Results from the 2 × 3 × 2 ANOVAs are reported in Table 3. In summary, the findings indicated, as expected, that: (1) persons with aphasia produced less quantity and diversity of nouns and verbs compared to neuro-typical adults across all three discourse genres (storytelling narrative, descriptive, and procedural). Furthermore, there were more words (especially nouns) in terms of quantity and diversity produced during the storytelling narrative compared to the other two discourse genres. The group of persons with aphasia produced more verbs than nouns during the procedural discourse; (2) the average imageability values for nouns were significantly higher than those for verb, and this was consistent between the neuro-typical and aphasia groups and across the three discourse genres; (3) the average frequency values did not differ for the two word classes, and between the neuro-typical and aphasia groups across the three discourse genres; and (4) the group of persons with fluent aphasia produced greater quantity and diversity of nouns and verbs compared to the group of persons with non-fluent aphasia across the three discourse genres (storytelling narrative, descriptive and procedural). Furthermore, there was greater production of words (especially nouns) in terms of quantity and diversity during storytelling narratives compared to the other two discourse genres. The persons with fluent aphasia produced more verbs than nouns during the procedural discourse. Details are provided in Supplementary Results.

Europe PMC Funders Author Manuscripts

Impact of lexical-semantic properties on content words produced during discourse

A Imageability effects

Neuro-typical adults versus aphasia groups: The imageability distribution pattern used by the neuro-typical group (Fig.3A) is fairly similar across the three discourse genres, with more words produced with high-imageability (501 - 600), and followed by low-imageability (< 300, and 300 - 400). It is apparent that this distribution was affected by the word class, in which there was: (i) increased production of concrete frequently used nouns (501 - 600 and 601 - 700, e.g. *'door, plate, milk'*), but no production of nouns with low imageability in all three discourse genres; and (ii) increased production of semantically 'light' and abstract verbs with lower imageability (< 300 and 301 - 400, e.g. *'think, decide, go, get, put')*, followed by mid-imageability verbs (401 - 500 and 501 - 600, e.g. *'lay, ask, fill')*, and very few high-imageability verbs that were produced in the storytelling narrative only (601 - 700, e.g., *'ring'*). Results from a 3 discourse genre × 2 word class × 5 imageability band repeated measure ANOVA are reported in Supplementary Results.

The imageability distribution used by persons with aphasia (Fig.3B) was similar in pattern to the distribution by the neuro-typical group for the three discourse genres, and with the same word class distinction. The only difference being that the aphasia group produced fewer words than the neuro-typical group in all bands across the three discourse genres (note the scale in Fig.3). Results from a 3 discourse genre × 2 word class × 5 imageability band repeated measure ANOVA are reported in Supplementary Results. These results are similar to those of the neuro-typical adults. A direct comparison between the neuro-typical and aphasia groups was conducted using a mixed ANOVA and reported in the Supplementary Results.

Fluent versus non-fluent aphasia sub-groups

Fig.3C/D illustrates the imageability distribution related to the content words produced during the three discourse genres by the fluent versus non-fluent aphasia sub-groups. Details of the analyses comparing the two groups are reported in Supplementary Results and indicated that the non-fluent aphasia group produced fewer words in all imageability bands across the three discourse genres than the fluent group while maintaining similar pattern of distribution.

B Frequency effects

Neuro-typical adults versus aphasia groups: Fig.4A illustrates the frequency distribution of the neuro-typical group in all three discourse genres, and it indicated that more words were produced within mid-frequency values (1 - 2.5) for the storytelling narrative and picture description, whereas more words were produced within a lower frequency band (< 1) for procedural discourse. Specifically, there were more nouns produced within mid-frequency (2 - 2.5, e.g., *'fish, floor, minutes')* in storytelling narrative and picture description, and more nouns within lower frequency (< 1) in procedural discourse, which represent prompt-specific compound vocabulary (e.g., *'teabag, teapot')* and thus this might not be generalisable to all procedural discourses. There were very few nouns produced in the high-frequency bands (2.5 - 3 and 3 - 3.5) and no nouns at the highest frequency band (> 3.5) across all discourse genres. On the other hand, there were more verbs produced within higher frequency bands (> 3, e.g., *'see, can, do')* and a limited number of verbs produced within low and mid-frequency bands. Results from a 3 discourse genre × 2 word class × 7 frequency band repeated measure ANOVA are reported in Supplementary Results.

The frequency distribution for the persons with aphasia group (Fig.4B) was similar in shape to the frequency distribution used by the neuro-typical group across the three discourse genres, with the same word class distinction, while producing fewer words in all bands. Results from a 3 discourse genre × 2 word class × 7 frequency band repeated measure ANOVA are reported in Supplementary Results. The data between neuro-typical and aphasia groups was directly compared using a mixed ANOVA and reported in the Supplementary Results.

Fluent versus non-fluent aphasia sub-groups

Fig.4C/D illustrated the frequency distribution related to the content words produced during the three discourse genres by the fluent versus non-fluent aphasia sub-groups. Details of the analyses between the two groups are reported in Supplementary Results, and indicated that the non-fluent aphasia group produced fewer words in all bands across the three discourse genres than the fluent group while maintaining relatively similar pattern of distribution.

C Lexical-semantic landscape

Neuro-typical adults versus aphasia groups: The constructed contour maps (Fig.5) demonstrate that the lexical-semantic landscape of the words produced during discourse by the aphasia group was broadly similar to the one of the neuro-typical adult group. Both maps contain two peaks: (i) in the top of the map and represents words with high-imageability and mid-frequency, and (ii) in the bottom right side corner of the map and represents words with low-imageability and high-frequency. Interestingly, by constructing contour maps for each word class, it becomes clear that each peak predominantly reflects different word classes for both groups: the space for nouns consists of one peak within the high-imageability and mid-frequency range, which contains commonly used everyday nouns (e.g., *'chair, cat, bedroom, restaurant'*) with a lack of abstract nouns. Conversely, the contour maps for the verbs, in both groups, consists of one main peak for low-imageability and high-frequency words, which contains regularly used light verbs (e.g., *'go, have, get, make, do, want, be')* and abstract emotional and cognitive verbs (e.g., *'think, assume, like '*). The verb contour map also includes another smaller peak for words with high-imageability and mid-frequency (i.e., overlapping with the peak for nouns). This part of the space contains concrete, frequently used verbs (e.g., *'ring, eat')*, and is might be influenced by task demands. Overall, there was very low use of low-frequency abstract words. The low-frequency abstract verb peak become broader in the neuro-typical group, extending towards concrete verbs with mid-frequency (e.g., *'concentrate, surprise, admire '*), although the number of verbs produced in these ranges was very small. In order to visualise the similarities/differences in the lexical-semantic landscapes and to determine statistical differences, we generated subtraction maps between the neuro-typical and aphasia groups. This subtraction map was constructed by performing t-tests between the lexical-semantic landscapes of the two groups on each imageability × frequency cell (Fig.5) with a Bonferroni correction ($p \leq 0.001$). The subtraction maps indicated areas of the lexical-semantic landscape that were used significantly more often by the neuro-typical group compared to the aphasia group (coloured green on Fig.5). The results indicated that the difference between groups reflected a global reduction in the number of words produced, rather than a reshaping of the space.

Aphasia sub-groups

The resultant contour maps from the aphasia sub-groups (Fig.6) demonstrated two main characteristics that are consistent with the other results. First, irrespective of how the persons with aphasia were subdivided (by fluency, semantic or phonological domains), each sub-group still utilised the same parts of the lexical-semantic landscape and differences between the maps reflected a generalised reduction in the number of words produced by the low performer groups rather than a change in the shape of the lexical-semantic landscape. Secondly, the largest differences between the maps resulted from splitting the groups of persons with aphasia according to fluency, whereas there were only modest differences between the groups split on the two other orthogonal aphasia domains (semantics and phonology).

Neuroimaging results

To identify the neural correlates associated with word retrieval during connected speech accounting for the lexical-semantic property of the word, we identified the two peaks within the neuro-typical adults' lexical-semantic landscape (representing a noun peak and a verb peak). Then, we computed the number of words produced by each person with aphasia at these peaks and used these values as the measure of interest in the univariate and multivariate lesion-symptom mapping analyses. Correlation analyses conducted between the number of words produced at these peaks and word count revealed significant strong positive correlations (noun: $r = 0.76$ and verb: $r = 0.89$, $p < 0.0001$). The VBCM results controlled for lesion volume and demographic variables (time post-stroke onset, age, education) (Fig.7A) revealed overlapping left frontal regions for verb (extend = 1579 voxels) and noun (extend = 950 voxels) retrieval within the lexical-semantic landscape. The overlap includes the left inferior frontal gyrus (pars opercularis), pre-central gyrus, central opercular cortex and insular cortex. The cluster related to verbs extended to the post-central gyrus; however, a direct contrast between noun versus verb retrieval within the lexical-semantic landscape revealed no differences in either direction.

The multivariate analyses yielded relatively similar results. The SVR-LSM results controlled for lesion volume and demographic variables (Fig.7B) identified significant clusters for both noun and verb retrieval within the lexical-semantic landscape that were only voxel-wise significant. The clusters involved the central opercular cortex. The cluster associated with verbs further extends to include the insular cortex and the inferior frontal gyrus (pars opercularis). When lesion volume was removed from the model, the SVR-LSM produced significant clusters that were voxel-wise and cluster-wise significant for both nouns and verbs. The significant clusters associated with verb and noun retrieval within the lexical-semantic landscape were overlapping in the left inferior frontal gyrus (pars opercularis and pars triangularis), frontal and central opercular cortices, insular cortex and pre- and post-central gyri. For noun retrieval, the cluster further extended to include the frontal orbital cortex, anterior superior temporal gyrus and planum polare. Fig.7 shows a high degree of similarity between the VBCM and SVR-LSM results.

The PRoNTo approach revealed significant brain-behaviour relationships for all models except for noun retrieval within the lexical-semantic landscape when controlled for lesion volume correction. There were no differences between models using the whole brain or restricted lesion space approaches. Specifically, for the models controlled for lesion volume and demographic variables, the cross-validated correlation coefficient for verb retrieval within the lexical-semantic landscape using the whole brain and restricted lesion territory was $r = 0.37$ and $r = 0.46$, respectively. On the other hand, the cross-validated correlation coefficient for the models without lesion volume correction using the whole brain was $r = 0.48$ and $r = 0.48$ for nouns and verbs, respectively; and using the restricted lesion territory it was $r = 0.51$ and $r = 0.54$ for nouns and verbs, respectively. For all significant models, we projected the beta weights to brain space and displayed the results in Fig.7C/D. The negative beta values from the PRoNTo results, in general, converged with results from both VBCM and SVR-LSM.

Europe PMC Funders Author Manuscripts

In conclusion, the convergent findings from the different lesion-symptom mapping approaches indicated that left frontal regions covering the inferior frontal gyrus (pars opercularis and pars triangularis), frontal and central opercular cortex, insular cortex and pre-central gyrus is associated with word retrieval during connected speech within the lexical-semantic landscape.

Discussion

Amongst the collection of variable symptoms in post-stroke aphasia, at least some degree of reduced connected speech production is a very common feature. Accordingly, it is important to be able to assess and understand the nature of this core aphasiological symptom. The majority of language production assessment tools used in research and clinical examinations rely heavily on single-word naming and picture description, which is not an everyday form of communication and connected speech production. In this detailed investigation on a large number of persons spanning the full range of severity of chronic post-stroke aphasia, we assessed the lexical-semantic properties of content words produced during connected speech. We explored this across different discourse genres to examine if the elicitation stimuli change the obtained results. We also compared the performance of the aphasia group to an age/education matched group of neuro-typical controls.

Across all three discourse genres, we found that although, as expected, persons with aphasia produced lower number of words than controls, they still sampled the same range of word frequency and imageability as that used by the neuro-typical adults. This pattern was consistent across the entire range of persons with aphasia, including fluent and non-fluent sub-groups, and also sub-groups defined by different levels of semantic and phonological abilities. This provides evidence that person with aphasia can produce words with the same lexical-semantic complexity as neural-typical adults, albeit at reduced word quantity. Persons with aphasia also showed reduced quantity and diversity of retrieval for different word classes (nouns and verbs) compared to neuro-typical adults across different discourse genres. Below we discuss these results.

The nature of vocabulary produced during discourse in post-stroke aphasia

It is striking that, although the quantity of vocabulary was reduced in fluent aphasia and even further in non-fluent aphasia, the pattern and types of words being produced were similar to that observed in neuro-typical controls. This is not the case in all clinical populations; previous examinations of connected speech in semantic dementia has found that the sampling of the two-dimensional frequency × imageability landscape was not only lower than neuro-typical adults but was changed in configuration with considerable under-sampling of low frequency words and a partially-compensatory over-sampling of high frequency substitutes (Bird et al., 2000; Hoffman et al., 2014). How can these contrastive clinical profiles be explained? There is a large body of evidence demonstrating that semantic dementia patients suffer from progressive degradation of core semantic representations (Hodges, Patterson, Oxbury, & Funnell, 1992; Snowden, Goulding, & Neary, 1989). This degraded semantic input to the language production system leads to profound anomia (Lambon Ralph, McClelland, Patterson, Galton, & Hodges, 2001) and semantic degradation

(Hoffman et al., 2014; Lambon Ralph, Graham, Ellis, & Hodges, 1998). Accordingly, it seems very likely that the low frequency words used in connected speech simply have too little semantic input to drive their retrieval. By contrast, post-stroke aphasia does not appear to involve the same type of core semantic degradation as that found in semantic dementia (Jefferies & Lambon Ralph, 2006), but instead reflects damage to other language components (Alyahya et al., 2020b). In terms of classical models of language production, semantic dementia would reflect impairment at the conceptualisation phase, whereas the breakdown in post-stroke aphasia is either after the semantic stage of the language production pathway and/or in other aspects related to fluent connected speech (e.g., in the mechanisms that control phrasal construction and the interface between lexical retrieval and sentence formation (Dell, Schwartz, Martin, Saffran, & Gagnon, 1997; Garrett, 1988). Preservation of conceptual input presumably means that persons with aphasia attempt to sample the entire breadth of the lexical-semantic landscape but the likelihood of any of these target words being produced is made (equally) less likely by post-semantic production deficits. This conclusion would seem to align with an observation that many persons with aphasia state – namely, that they know what they want to say but cannot produce it.

The relationship between imageability, frequency and word class

The pattern of results with regards to the lexical-semantic properties of the words produced during discourse by the neuro-typical group replicated previous findings from a different sample of neuro-typical controls (Bird et al., 2000). In this study, persons with aphasia produced fewer verbs and nouns in terms of quantity and diversity than neuro-typical adults, consistent with previous studies based on smaller sample sizes (Bastiaanse & Jonkers, 1998; Berndt et al., 1997b). The present results also complement another study that indicated intact verb production with respect to semantic category and weight during connected speech in aphasia (Cruice, Pritchard, & Dipper, 2014). As with previous studies on a different patient group (i.e. semantic dementia) (Bird et al., 2000; Hoffman et al., 2014), the distribution-based analysis provides a very clear and unique way to explore the results obtained from clinical populations. As well as revealing striking differences between post-stroke aphasia and semantic dementia (see above), these analyses also reveal the inescapable relationship between word classes and lexical-semantic properties. Specifically, the lexical-semantic landscape contained many high-imageability nouns with mid-frequency (e.g., *'phone, girl')* and high-frequency verbs with low-imageability (e.g., *'cook, change')* across all discourse genres. This is consistent with small-sample studies showing that patients with fluent and non-fluent aphasia used high-frequency light verbs in their narratives (Berndt et al., 1997b), and fluent aphasia sampling word frequency in a similar manner to controls (Bastiaanse, 2011). The corollary of these differential distributions for nouns and verbs is that any apparent word-class effect needs to be treated with caution, as they could reflect the inherent differences driven by word imageability and/or frequency. This resonates with previous documentations on the influence of semantic distinctions between nouns and verbs on single-word production in aphasia (e.g., Alyahya et al., 2018a; Bird et al., 2000).

The influence of different types of discourse elicitation stimuli

The picture-supported storytelling narrative, which is not commonly used in research and clinical practice, boosted connected speech production including eliciting a greater quantity

Europe PMC Funders Author Manuscripts

and diversity of words, both nouns and verbs, and the full range of word frequency and imageability. Indeed, the entire frequency and imageability ranges (including the word class distinction) were considerably higher for this elicitation stimulus. This production boost was found in both the neuro-typical controls and person with aphasia (both fluent and non-fluent). These results are consistent with recent studies indicating the main effect of discourse elicitation stimuli in which the storytelling narrative had an advantage over other stimuli in eliciting greater quantity, diversity, words-per-minutes and propositional density (Alyahya et al., 2020b; Stark, 2019). We suspect that there are multiple factors that contributed to the success of storytelling narrative. First, in this study persons with aphasia were not given a time limit for discourse production, and they took longer to narrate the story compared to the time spent on the other discourse stimuli. Presumably this extended time led to more overt production (note that quantity does not necessarily follow duration: relative to neuro-typical controls, the persons with aphasia produced less output in their storytelling narratives but spent much longer doing so). Secondly, the stimuli used in this study to elicit storytelling narrative had many characters and events, and thus offers many more opportunities and topics for language production. Thirdly, the storytelling narrative was picture supported, which could help as the story-board provides direct prompts about items and actions to describe including temporal and spatial shifts between events and characters, and does so for all stages of the story providing opportunity for more language production. Further studies are needed to examine these possible factors. Nevertheless, the three discourse genres elicited very similar profiles across frequency and imageability, albeit picture descriptions and procedure discourse generating less words overall. There was a slight tendency for the picture description to elicit relatively more nouns and, conversely, the procedural discourse to generate relatively more verbs. The utilisation of other elicitation stimuli might be useful to elicit low-frequency and low-imageability words (e.g., *'nationalist, provender, naiad')*, a part of the lexical-semantic landscape that was not probed using the stimuli used in this study. This part of the space seems to consist of specialist technical vocabulary, which might be probed using personal questions. It is unclear, however, how informative this part of the space is in real speech contexts and whether we would gain insights by focusing on them, given that patients and healthy adults rarely use this part of the space.

Neural correlates of word retrieval during connected speech

The neural correlates associated with word retrieval (irrespective of the word class or lexical-semantic complexity) during connected speech mainly encompassed left frontal regions covering the inferior frontal gyrus, insular cortex, frontal and central operculum, and pre-central gyrus. These regions were identified using both univariate and multivariate lesion-symptom mapping approaches. These findings are highly consistent with previous lesion studies and fMRI experiments in healthy controls, which associated these left frontal regions to fluency in aphasia, propositional speech production in neuro-typical adults and motor-speech planning (Alyahya et al., 2020b; Basilakos, Rorden, Bonilha, Moser, & Fridriksson, 2015; Blank, Scott, Murphy, Warburton, & Wise, 2002; Dronkers, 1996). A study that utilised network lesion-symptom mapping to explore the neural correlates of nouns and verbs during picture description in aphasia also found action word production to be supported by frontal networks (Gleichgerrcht et al., 2016).

From a methodological point of view, it is important to note the complementary differences between the interpretation of univariate and multivariate analyses (Hebart & Baker, 2018). Generally, univariate analyses assign beta values to voxels in a relatively transparent way, in which the strength and sign of these values indicate meaningful brain-behaviour relationships. Therefore, it is easier to make inferences about localisation of functions. However, there are practical challenges with univariate methods that must be accounted for, such as correcting for multiple comparisons. There are also theoretical concerns, including assumption of voxel independence and mislocalisation of effects (DeMarco & Turkeltuab, 2018; Karnath, Sperber, & Rorden, 2018; Mah et al., 2014). On the other hand, multivariate approaches are inherently different. Specifically, they can be used to make formal behavioural predictions via mapping brain status to behavioural performance, and they are commonly used for encoding or decoding (Hebart & Baker, 2018; Naselaris, Kay, Nishimoto, & Gallant, 2011). One important limitation with multivariate approaches is how to interpret the feature weights (Haufe et al., 2014; Hebart & Baker, 2018). There are potential strategies that might aid improving interpretability, which include encoding methods (such as partial least squares and canonical correlation analysis) or bootstrap analyses (Kuceyeski et al., 2016) but this is non-trivial. Finally, multivariate decoding approaches typically require a large dataset, as data are partitioned into training versus test sets for cross validation. This can be practically challenging, particular for patient studies where recruitment is difficult. In a recent simulation study (Sperber et al., 2019), it was suggested that a sample size of 100 participants is required to produce stable and reproducible beta parameter mapping, whereas the sample size for prediction of clinical outcomes peaked at 40 and was relatively stable from this point up to 100 participants. Given the differences between various brain-mapping approaches, it is striking that the multivariate analyses (both SVR-LSM and PRoNTo) converged with the VBCM findings in this study.

Conclusion

We explored the properties of content words produced during connected speech across different discourse genres in persons with post-stroke aphasia, and compared their performance to an age/education matched group of neuro-typical controls. The results provided evidence that person with aphasia can produce words with the same lexical-semantic complexity as neural-typical adults, albeit at reduced word quantity. Across all three discourse genres, we found that although, the quantity of words produced by persons with aphasia was lower than neuro-typical controls; they still sampled the same range of word frequency and imageability as that used by the neuro-typical adults. This pattern was consistent across the entire ranges of the aphasia group, including sub-groups of fluent and non-fluent aphasia and those with high versus low semantic and phonological abilities. Findings from this study also indicated that cautions should be taken when selecting discourse elicitation stimuli, as different stimuli lead to differences in the lexical-semantic properties of the produced discourse. Picture-supported storytelling narrative, which is not commonly used in research and clinical practice, can boost connected speech production including eliciting a greater quantity and diversity of words across all classes and the full range of word frequency and imageability. For clinical assessments, this is important as the

Europe PMC Funders Author Manuscripts

other commonly utilised elicitation stimuli might sometimes under-estimate the production abilities and the range of vocabulary used by patients. For research investigations, maximising discourse output will inevitably make the assessments more sensitive and better able to grade differences between patients. Neuroimaging findings from both univariate and multivariate approaches revealed shared left frontal regions in association with word retrieval, irrespective of their word class and lexical-semantic complexity, during connected speech. This would imply that interventions such as non-invasive brain stimulation can target these left frontal brain regions to enhance word retrieval regardless of the word's class or complexity.

References

Alyahya RSW, Halai AD, Conroy P, Lambon Ralph MA. Noun and verb processing in aphasia: Behavioural profiles and neural correlates. NeuroImage: Clinical. 2018a; 18: 215–230. DOI: 10.1016/j.nicl.2018.01.023 [PubMed: 29868446]

Alyahya RSW, Halai AD, Conroy P, Lambon Ralph MA. The relationship between the multidimensionality of aphasia and psycholinguistic features. Aphasiology. 2018b; 32: 5–6. DOI: 10.1080/02687038.2018.1486374

Alyahya RSW, Halai AD, Conroy P, Lambon Ralph MA. Mapping psycholinguistic features to the neuropsychological and lesion profiles in aphasia. Cortex. 2020a; 124: 260–273. DOI: 10.1016/j.cortex.2019.12.002 [PubMed: 31958653]

Alyahya RSW, Halai AD, Conroy P, Lambon Ralph MA. A unified model of post-stroke language deficits including discourse production and their neural correlates. Brain. 2020b; 143: 1541–1554. DOI: 10.1093/brain/awaa074 [PubMed: 32330940]

Basilakos A, Fillmore P, Rorden C, Rorden C, Guo D, Bonilha L, Fridriksson J. Regional white matter damage predicts speech fluency in chronic post-stroke aphasia. Frontiers in Human Neuroscience. 2014; 8: 1–9. DOI: 10.3389/fnhum.2014.00845 [PubMed: 24474914]

Basilakos A, Rorden C, Bonilha L, Moser D, Fridriksson J. Patterns of poststroke brain damage that predict speech production errors in apraxia of speech and aphasia dissociate. Stroke. 2015; 46: 1561–1566. DOI: 10.1161/STROKEAHA.115.009211 [PubMed: 25908457]

Bastiaanse R. The retrieval and inflection of verbs in the spontaneous speech of fluent aphasic speakers. Journal of Neurolinguistics. 2011; 24 (2) 163–172. DOI: 10.1016/j.jneuroling.2010.02.006

Bastiaanse R, Jonkers R. Verb retrieval in action naming and spontaneous speech in agrammatic and anomic aphasia. Aphasiology. 1998; 12: 951–969. DOI: 10.1080/02687039808249463

Bates E, Wilson SM, Saygin AP, Dick F, Sereno MI, Knight RT, et al. Voxelbased lesion-symptom mapping. Natural Neuroscience. 2003; 6: 448–450.

Berndt R, Haendiges A, Mitchum C, Sandson J. Verb retrieval in aphasia: 2. Relationship to sentence processing. Brain and Language. 1997b; 56: 107–137. [PubMed: 8994700]

Bird H, Franklin S. Cinderella revisited: A comparision of fluent and non-fluent aphasic speech. Journal of Neurolinguistics. 1996; 9: 187–206.

Bird H, Franklin S, Howard D. Age of acquisition and imageability ratings for a large set of words, including verbs and function words. Behavior Research Methods, Instruments, & Computers. 2001; 33: 73–79.

Bird H, Howard D, Franklin S. Why is a verb like an inanimate object? Grammatical category and semantic category deficits. Brain and Language. 2000; 72 (3) 246–309. DOI: 10.1006/brln.2000.2292 [PubMed: 10764520]

Bird H, Howard D, Franklin S. Verbs and nouns: the importance of being imageable. Journal of Neurolinguistics. 2003; 16: 113–149.

Blank SC, Scott SK, Murphy K, Warburton E, Wise RJS. Speech Production: Wernicke, Broca and beyond. Brain. 2002; 125: 1829–1838. [PubMed: 12135973]

Bryant L, Ferguson A, Spencer E. Linguistic analysis of discourse in aphasia: A review of the literature. Clinical Linguistics & Phonetics. 2016; 30: 489–518. DOI: 10.3109/02699206.2016.1145740 [PubMed: 27002416]

Coltheart M. The MRC psycholinguistic database. The Quarterly Journal of Experimental Psychology. 1981; 33A: 497–505.

Consortium B. N. C. British National Corpus version 3 (BNC *XML* edition). Oxford University Computing Services; Oxford: 2007.

Crepaldi D, Berlingeri M, Cattinelli I, Borghese NA, Luzzatti C, Paulesu E. Clustering the lexicon in the brain: A meta-analysis of the neurofunctional evidence on noun and verb processing. Frontiers in Human Neuroscience. 2013; 7: 1–15. DOI: 10.3389/fnhum.2013.00303 [PubMed: 23355817]

Crepaldi D, Ingignoli C, Verga R, Contardi A, Semenza C, Luzzatti C. On nouns, verbs, lexemes, and lemmas: Evidence from the spontaneous speech of seven aphasic patients. Aphasiology. 2011; 25: 71–92. DOI: 10.1080/02687031003637114

Cruice M, Pritchard M, Dipper L. Verb use in aphasic and non-aphasic personal discourse: What is normal? Journal of Neurolinguistics. 2014; 28: 31–47. DOI: 10.1016/j.jneuroling.2013.12.002

Dell GS, Schwartz MF, Martin N, Saffran EM, Gagnon DA. Lexical access in aphasic and nonaphasic speakers. Psychological Review. 1997; 104: 801–838. DOI: 10.1037/0033-295X.104.4.801 [PubMed: 9337631]

DeMarco AT, Turkeltuab PE. A multivariate lesion symptom mapping toolbox and examination of lesion-volume biases and correction methods in lesion-symptom mapping. Human Brain Mapping. 2018; 39: 4169–4182. DOI: 10.1002/hbm.24289 [PubMed: 29972618]

Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, et al. Killiany RJ. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage. 2006; 1: 968–680.

Dipper, LT, Pritchard, M. Advances in Speech-language Pathology:. Fernandes, F, editor. InTech; 2017.

Dronkers NF. A new brain region for coordinating speech artculation. Nature. 1996; 384: 159–161. [PubMed: 8906789]

Fletcher, M, Birt, D. Storylines: Picture sequences for language practice. Longman; London: 1983.

Garrett, MF. Linguistics: The Cambridge survey: III Language: Psychological and biological aspects. Newmeyer, FJ, editor. Cambridge University Press; Cambridge: 1988. 69–96.

Gleichgerrcht E, Fridriksson J, Rorden C, Nesland T, Desai R, Bonilha L. Separate neural systems support representations for actions and objects during narrative speech in post-stroke aphasia. NeuroImage: Clinical. 2016; 10: 140–145. DOI: 10.1016/j.nicl.2015.11.013 [PubMed: 26759789]

Goodglass, H, Kaplan, E. Boston Diagnostic Aphais Examination. 2 ed. Lea & Febiger; Philadelphia: 1983.

Haufe S, Meinecke F, Gorgen K, Dahne S, Haynes J-D, Blankertz B, Bießmann F. On the interpretation of weight vectors of linear models in multivariate neuroimaging. Neuroimage. 2014; 87: 96–110. [PubMed: 24239590]

Hebart MN, Baker CI. Deconstructing multivariate decoding for the study of brain function. Neuroimage. 2018; 180: 4–18. [PubMed: 28782682]

Hodges J, Patterson K, Oxbury S, Funnell E. Semantic dementia. Progressive fluent aphasia with temporal lobe atrophy. Brain. 1992; 115: 1783–1806. [PubMed: 1486461]

Hoffman P, Meteyard L, Patterson K. Broadly speaking: Vocabulary in semantic dementia shifts towards general, semantically diverse words. Cortex. 2014; 2014: 30–42. DOI: 10.1016/j.cortex.2012.11.004

Jefferies E, Lambon Ralph MA. Semantic impairment in stroke aphasia versus semantic dementia: a case-series comparison. Brain. 2006; 129: 2132–2147. [PubMed: 16815878]

Karnath HO, Sperber C, Rorden C. Mapping human brain lesions and their functional consequences. Neuroimage. 2018; 165: 180–189. [PubMed: 29042216]

Kuceyeski A, Navi BB, Kamel H, Raj A, Relkin N, Toglia J, O'Dell M. Structural connectome disruption at baseline predicts 6-months poststroke outcome. Human Brain Mapping. 2016; 37: 2587–2601. [PubMed: 27016287]

Lambon Ralph MA, Graham KS, Ellis A, Hodges J. Naming in semantic dementia-what matters? Neuropsychologia. 1998; 36: 775–784. [PubMed: 9751441]

Lambon Ralph MA, McClelland J, Patterson K, Galton CJ, Hodges J. No right to speak? The relationship between object naming and semnatic impairments: Neuropsychological evidence and a computational model. Jouranl of Cognitive Neuroscience. 2001; 13: 341–356.

Mah Y-H, Husain M, Rees G, Nachev P. Human brain lesion-deficit inference remapped. Brain. 2014; 127: 2522–2531.

Naselaris T, Kay KN, Nishimoto S, Gallant JL. Encoding and decoding in fMRI. Neuroimage. 2011; 56: 400–410. [PubMed: 20691790]

Nicholas L, Brookshire RH. A system for quantifying the informativeness and efficiency of the connected speech of adults with aphasia. Journal of speech and Hearing Research. 1993; 36: 338–350. [PubMed: 8487525]

Sandberg C, Kiran S. Analysis of abstract and concrete word processing in persons with aphasia and age-matched neurologically healthy adults using fMRI. Neurocase. 2014; 20: 361–388. DOI: 10.1080/13554794.2013.770881 [PubMed: 23548150]

Schrouff J, Rosa MJ, Rondina JM, Marquand AF, Ashburner J, Phillips C, Mourao-Miranda J. PRoNTo: Pattern Recognition for Neuroimaging Toolbox. Neuroinform. 2013; 11: 319–337. DOI: 10.1007/s12021-013-9178-1

Schumacher R, Halai AD, Lambon Ralph MA. Assessing and mapping language, attention and executive multidimensional deficits in stroke aphasia. Brain. 2019; 142: 3202–3216. DOI: 10.1093/brain/awz258 [PubMed: 31504247]

Seghier ML, Ramlackhansingh A, Crinion JT, Leff AP, Price C. Lesion identification using unified segmentation-normalisation models and fuzzy clustering. Neuroimage. 2008; 41: 1253–1266. [PubMed: 18482850]

Siri S, Tettamanti M, Cappa S, Rosa P, Saccuman C, Scifo P, Vigliocco G. The neural substrate of naming events: Effects of processing demands but not of grammatical class. Cerebral Cortex. 2008; 18: 171–177. [PubMed: 17507455]

Snowden JS, Goulding PJ, Neary D. Semantic dementia: a form of circumscribed cerebral atrophy. Behavioral Neurology. 1989; 2: 167–182.

Sperber C, Wiesen D, Karnath H-O. An empirical evaluation of multivariate lesion behaviour mapping using support vector regression. Human Brain Mapping. 2019; 40: 1381–1390. [PubMed: 30549154]

Stark B. A Comparison of three discourse elicitation methods in aphasia and age-matched adults: Implications for language assessment and outcome. American Journal of Speech-Language Pathology. 2019; 28: 1067–1083. DOI: 10.1044/2019_AJSLP-18-0265 [PubMed: 31181170]

Tipping MES. pare Bayesian learning and the relevance vector machine. J Mach Learn Res. 2001; 1: 211–244.

Tyler L, Marslen-Wilson W, Stamatakis EA. Dissociating neuro-cognitive component processes: Voxel-based correlational methodology. Neuropsychologia. 2005; 43 (5) 771–778. DOI: 10.1016/j.neuropsychologia.2004.07.020 [PubMed: 15721189]

Vigliocco G, Vinson D, Druks J, Cappa S. Nouns and verbs in the brain: A review of behavioural, electrophysiological, neuropsychological and imaging studies. Neuroscience and Biobehavioral Reviews. 2011; 35: 407–426. [PubMed: 20451552]

Zhao Y, Halai AD, Lambon Ralph MA. Evaluating the granularity and statistical structure of lesions and behaviour in post-stroke aphasia. Brain Communications. 2020; 2 doi: 10.1093/braincomms/fcaa062

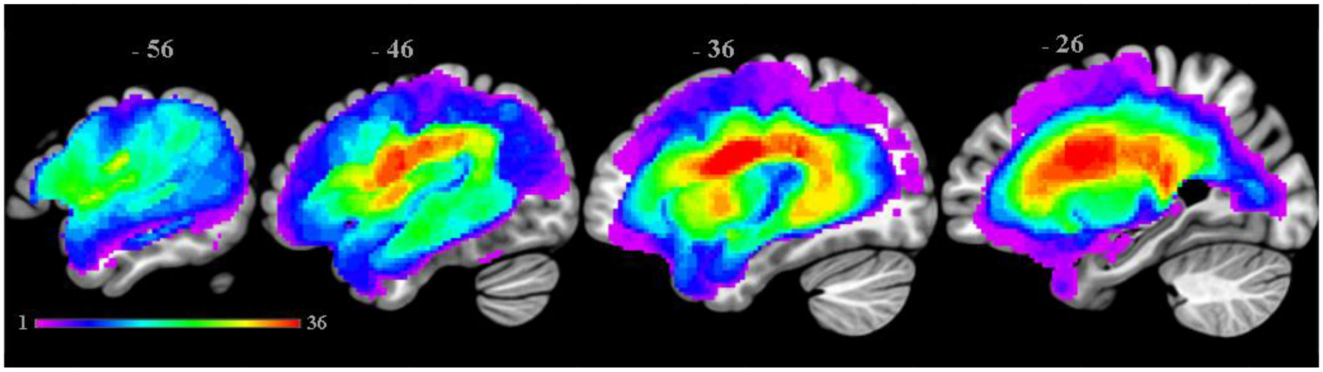

Fig. 1. Lesion overlap map across 46 participants with post-stroke aphasia.

Colour scale illustrates the distribution of the lesions and represents the number of participants with a lesion at that location. The maximum number of participants who had a lesion in one voxel was 36 (MNI coordinate: -38, -9, 24; central opercular cortex).

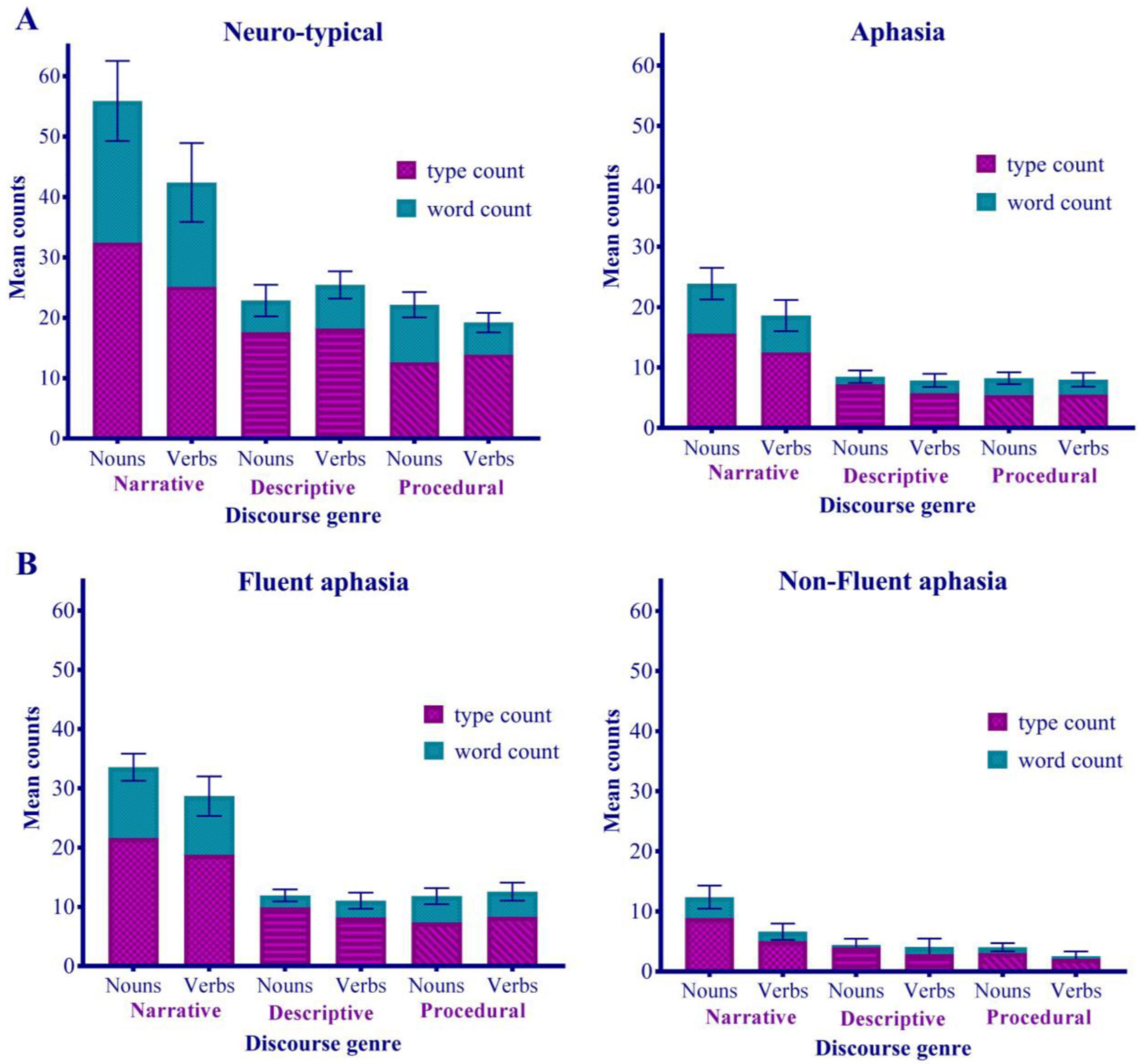

Fig. 2. The quantity and diversity of nouns and verbs produced during discourse.

Bar graphs showing the group mean and standards errors (errors bars) of the quantity of nouns and verbs (turquoise bars) and the diversity of nouns and verbs (purple bars) produced during three different discourse genres among the: (A) neuro-typical adults and all persons with aphasia, and (B) the fluent and non-fluent aphasia sub-groups.

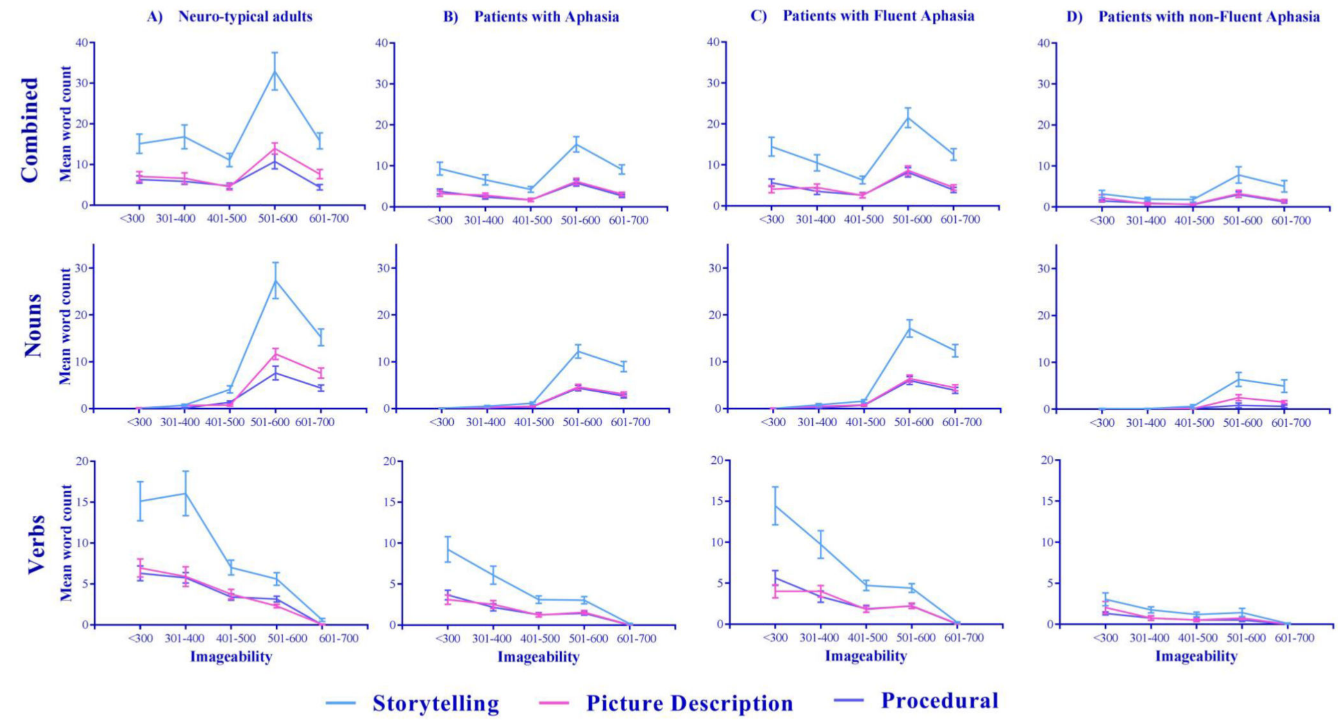

Fig. 3. Imageability distribution of the content words produced during different discourse genres.

The group mean of the number of words produced within each imageability band by: A) neuro-typical adult group, B) persons with aphasia group, C) persons with fluent aphasia sub-groups, and D) persons with non-fluent aphasia sub-group. Top row: nouns and verbs combined, middle row: nouns, bottom row: verbs. Error bars represent the standard error of the mean.

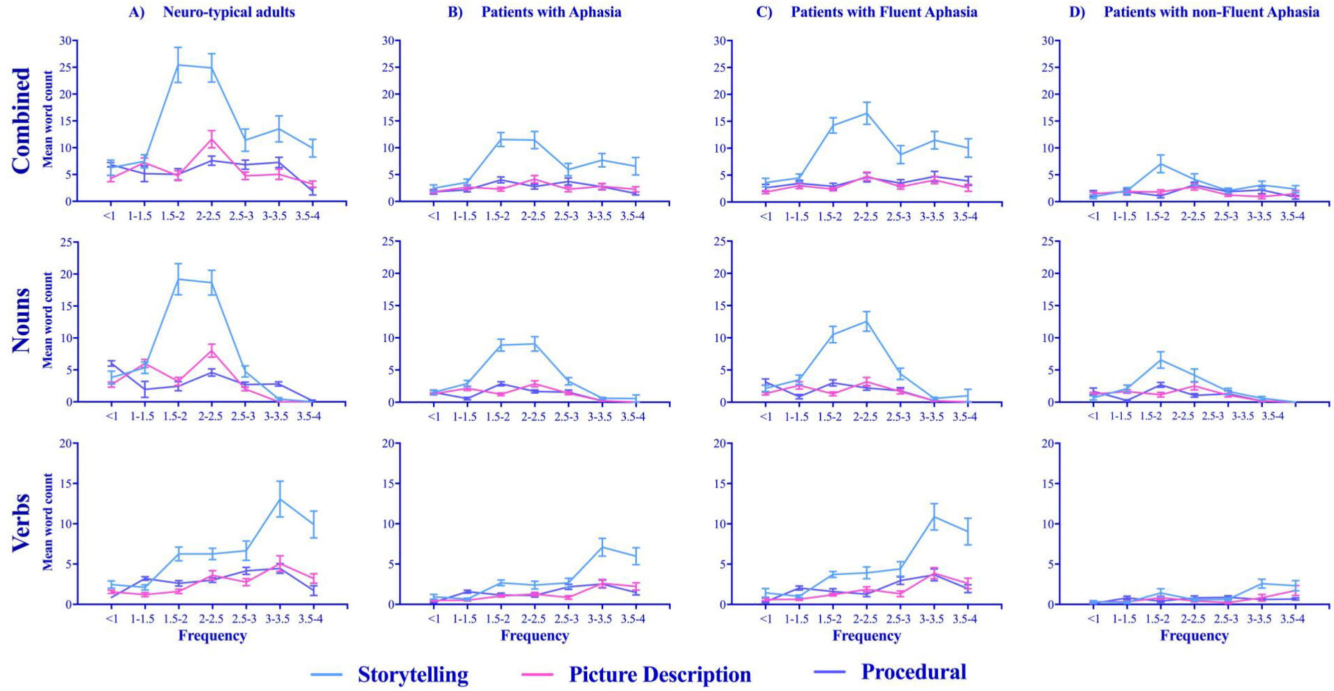

Fig. 4. Frequency distribution of the content words produced during different discourse genres.

The group mean of the number of words produced within each frequency band by: A) neuro-typical adult group, B) persons with aphasia group, C) persons with fluent aphasia sub-group, and D) persons with non-fluent aphasia sub-group. Top row: nouns and verbs combined, middle row: nouns, bottom row: verbs. Error bars represent the standard error of the mean.

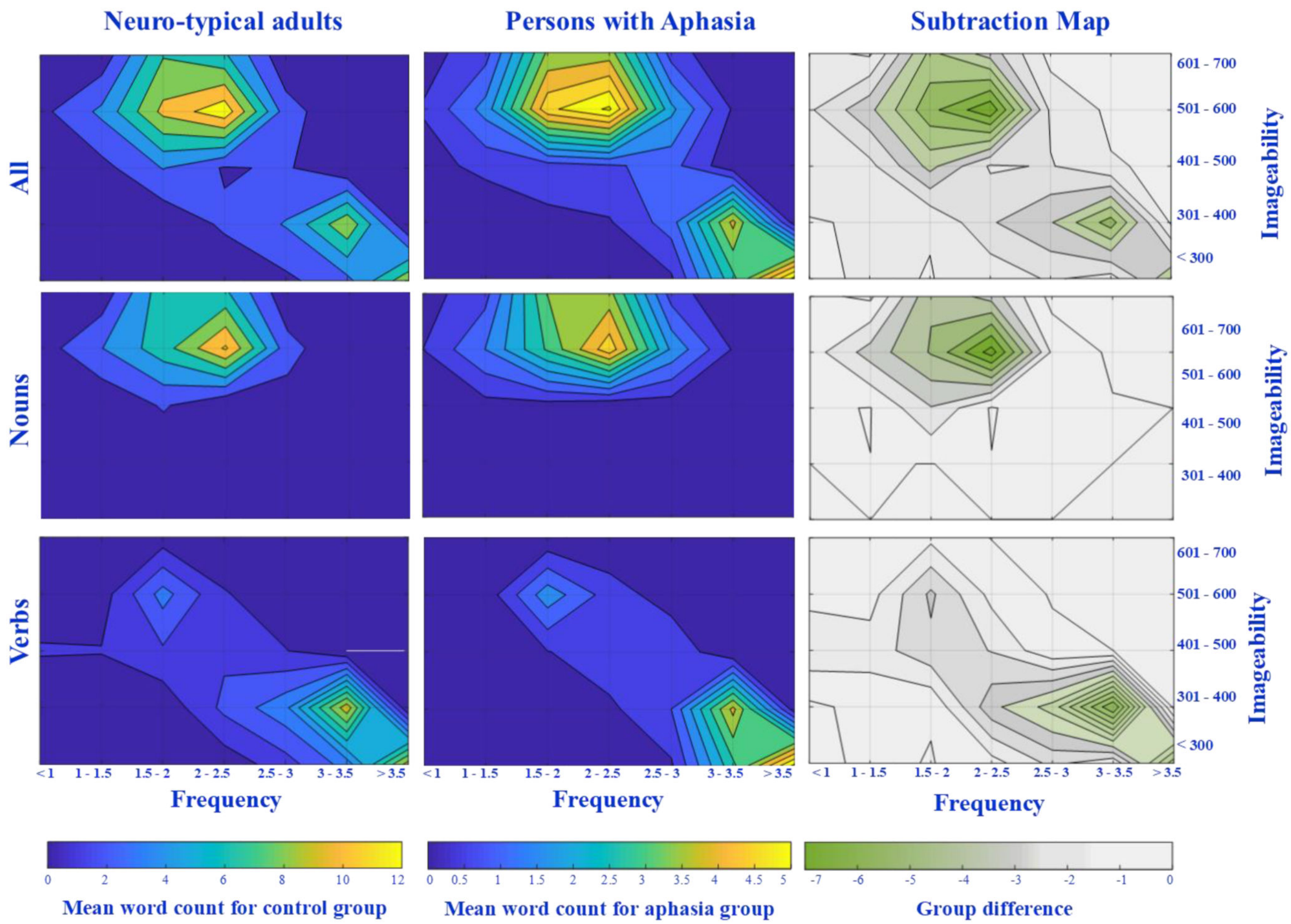

Fig. 5. Contour maps of neuro-typical adults and persons with aphasia representing lexical-semantic landscapes.

Two-dimensional frequency × imageability landscapes representing the mean word count produced during storytelling narrative. T-tests were used to show differences between the two groups in the third column and significantly different parts of the spaces are shown in green ($p \leq 0.001$).

Europe PMC Funders Author Manuscripts

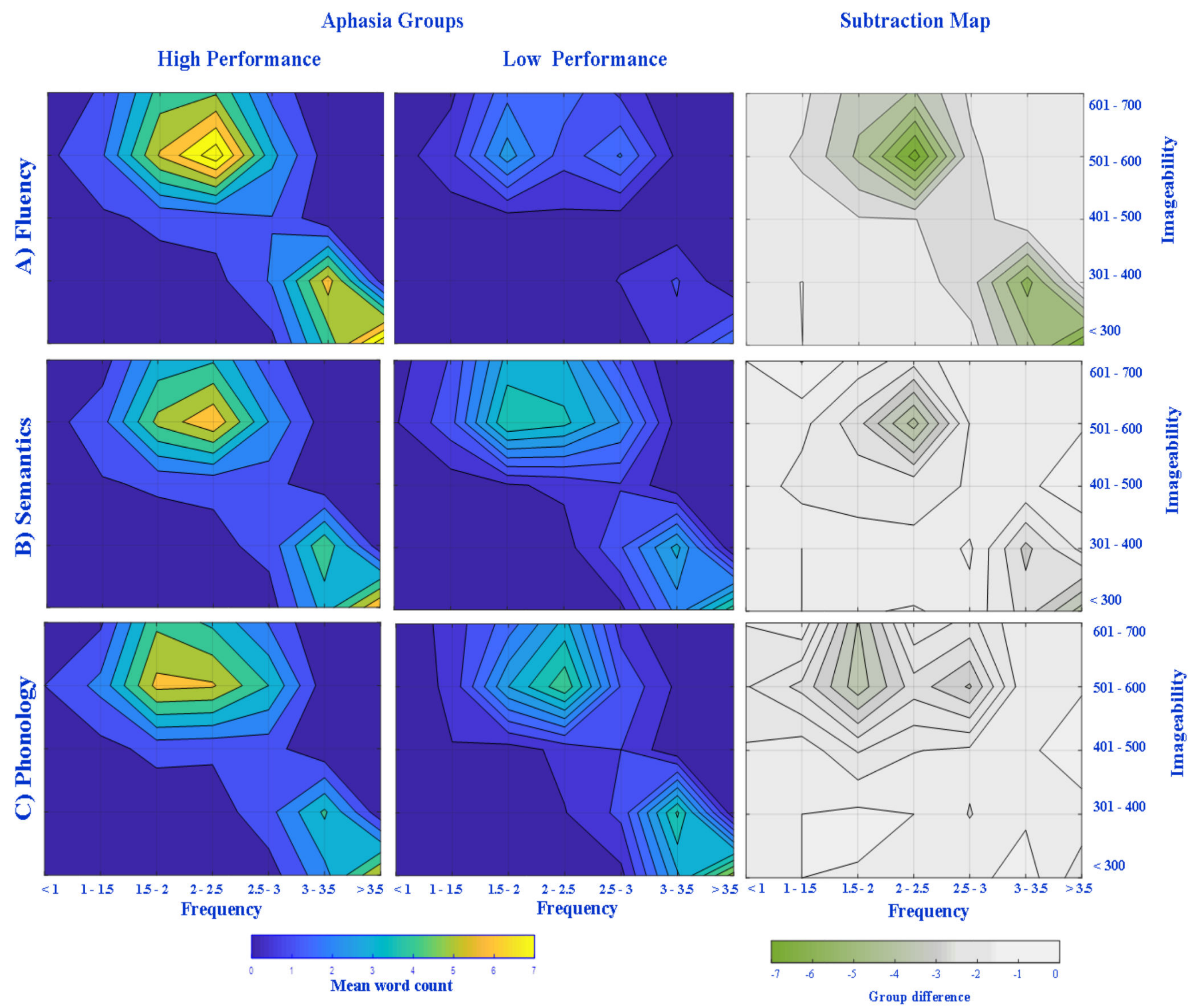

Fig. 6. Contour maps for different aphasia sub-groups representing lexical-semantic landscapes.
Two-dimensional frequency × imageability landscapes showing the mean word count produced during storytelling narrative by: A) persons with fluent versus non-fluent aphasia, B) high versus low performers on semantic domain, and C) high versus low performers on phonological domain. T-tests were used to show differences between the two groups in the third column and significantly different parts of the spaces are shown in green ($p \leq 0.001$).

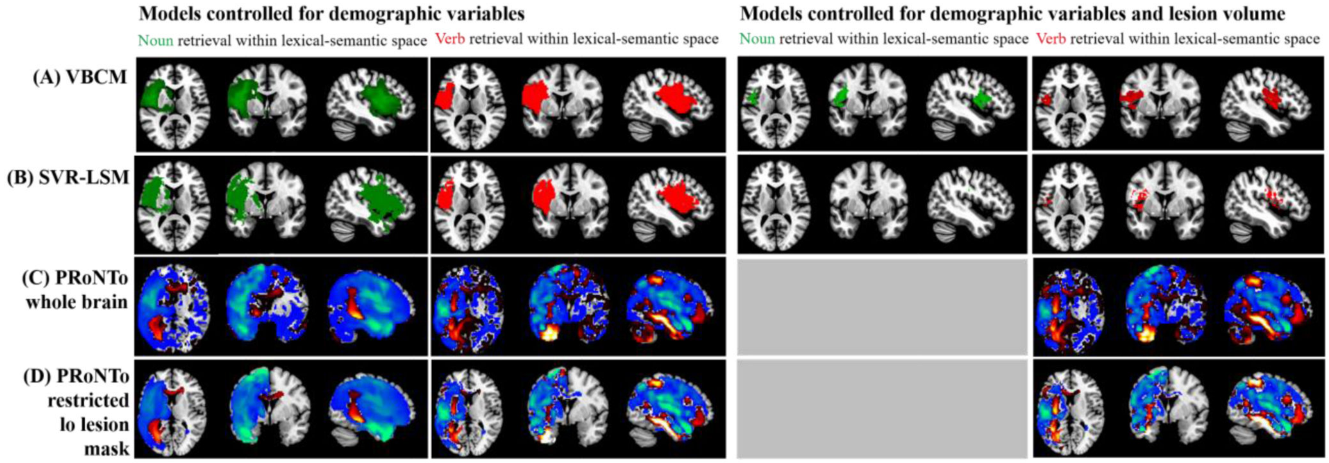

Fig. 7. Neuroimaging results from different lesion-symptom mapping approaches showing the neural correlates associated with word retrieval during connected speech accounting for the lexical-semantic properties of the words.

Left panels: models controlled for demographic variables (months post stroke onset, age, and education); right panels: models also controlled for lesion volume. MNI coordinates of slices from left to right: Z = 11, Y = 1, X = -45. (A) VBCM indicating the neural correlates associated with verb retrieval (red) and noun retrieval (green) within the lexical-semantic landscape thresholded at $p<0.005$ voxel-wise and FWE cluster corrected at $p<0.05$. (B) SVR-LSM showing voxels with significant beta weights after 10000 permutations testing, $p<0.005$ voxel-wise and $p<0.005$ cluster-wise for the models without lesion volume correction; and $p<0.005$ voxel-wise for the models with lesion volume correction. (C) PRoNTo depicting the neural weights back-projected on to 3D brain for significant model of noun and verb retrieval within the lexical-semantic landscape (permutation $p<0.05$) on the whole brain space. (D) PRoNTo depicting the neural weights back-projected on to 3D brain for significant model of noun and verb retrieval within the lexical-semantic landscape (permutation $p<0.05$) restricted to lesion territory. PRoNTo results thresholded from -0.0001 to -0.01 (blue-green colours) and 0.0001 to 0.01 (red-yellow couloirs), and the negative weights are considered as stronger in this approach. A grey surface indicates that no significant results were found for the respective measure and methodological approach.

Table 1

Descriptive statistics of the measures extracted from the three discourses produced by the neuro-typical and aphasia groups

Discourse genre	Storytelling Narrative			Descriptive			Procedural		
Measure	**Range**	**mean**	**SD**	**Range**	**mean**	**SD**	**Range**	**mean**	**SD**
Neuro-typical group									
Number of tokens	101 - 706	265.3	140.4	56 - 252	107.65	49.1	69 - 278	118.2	47.75
Duration (seconds)	43 - 322	128	61.4	27 - 118	48.2	25.1	24 - 121	48.50	22.00
Quantity of verb	16 - 152	42.4	28.4	9 - 47	18.2	9.90	11 - 44	19.20	07.11
Diversity of verbs	15 - 47	25.1	8.6	7 - 27	12.9	4.62	8 - 21	13.90	03.55
Quantity of nouns	20 - 144	56	28.9	10 - 57	21.6	11.4	14 - 56	21.75	09.04
Diversity of nouns	13 - 74	32.4	15.2	8 - 44	17.15	8.27	8 - 21	12.60	03.6
Aphasia group									
Number of tokens	8 - 454	156.7	114.53	6 - 315	66.57	59.70	5 - 262	59.02	57.74
Duration (seconds)	42 - 620	190.04	109.78	11 - 310	95.09	66.22	10 - 219	61.95	42.66
Quantity of verb	0 - 114	18.80	17.58	0 - 52	07.86	07.41	0 - 38	08.00	07.82
Diversity of verbs	0 - 49	12.52	09.99	0 - 23	05.76	04.49	0 - 21	06.09	05.42
Quantity of nouns	0 - 66	23.89	17.85	0 - 40	08.5	07.02	0 - 30	08.24	06.60
Diversity of nous	0 - 43	15.58	10.97	0 - 24	07.23	05.17	0 - 17	05.41	03.39

Europe PMC Funders Author Manuscripts

Table 2

Mean (SD) of frequency and imageability of the words produced by the neuro-typical adults and persons with aphasia across the three discourse genres

Word type	Discourse Group	Frequency[1]			Imageability[2]		
		Storytelling Narrative	Descriptive	Procedural	Storytelling Narrative	Descriptive	Procedural
All	Neuro-	2.34	2.34	1.88	469.10	469.64	481.90
	typical	(0.13)	(0.19)	(0.21)	(11.26)	(17.26)	(19.58)
	All persons	2.48	2.17	1.99	465.19	484.73	473.92
	with aphasia	(0.49)	(0.70)	(0.48)	(45.23)	(60.15)	(67.76)
	Fluent	2.44	2.14	2.04	473.43	487.83	477.44
	aphasia	(0.22)	(0.07)	(0.04)	(16.61)	(35.43)	(23.99)
	Non-fluent	2.56	2.26	2.00	457.25	475.31	492.26
	aphasia	(0.65)	(0.84)	(0.60)	(60.86)	(95.25)	(95.96)
Nouns	Neuro-	1.89	1.64	1.39	573.57	579.35	583.29
	typical	(0.08)	(0.11)	(0.19)	(18.19)	(13.73)	(16.85)
	All persons	1.92	1.69	1.64	579.09	579.28	589.73
	with aphasia	(0.27)	(0.49)	(0.29)	(34.18)	(21.95)	(21.95)
	Fluent	1.96	1.69	1.60	578.26	577.44	588.83
	aphasia	(0.19)	(0.36)	(0.3)	(13.45)	(18.71)	(22.26)
	Non-fluent	1.86	1.69	1.70	580.08	591.76	590.97
	aphasia	(0.35)	(0.65)	(0.28)	(23.68)	(25.35)	(22.09)
Verbs	Neuro-	2.79	3.04	2.36	364.62	359.93	380.5
	typical	(0.25)	(0.32)	(0.37)	(19.4)	(31.04)	(33.26)
	All persons	2.92	2.71	2.53	367.87	393.44	357.45
	with aphasia	(0.59)	(0.8)	(0.55)	(47.85)	(65.74)	(51.99)
	Fluent	2.91	2.71	2.49	368.60	389.87	366.04
	aphasia	(0.38)	(0.69)	(0.43)	(30.66)	(48.72)	(36.38)
	Non-fluent	2.93	2.81	2.48	366.99	397.96	343.14
	aphasia	(0.65)	(0.8)	(0.87)	(63.43)	(83.74)	(66.95)

[1]Lemma frequency that were log-transformed.

[2]Imageability ratings on a scale from 100 to 700.

Europe PMC Funders Author Manuscripts

Table 3

Findings from ANOVAs on the effects of group, discourse genre, word class and their interactions on the quantity, diversity and psycholinguistic properties (imageability and frequency) of the words produced during discourse

Conditions	Dependent variable			
	Quantity (word count)	Diversity (type count)	Mean imageability values	Mean frequency values
Group	($F(1,64) = 36.46$, $p < 0.0001$, $\eta^2 = 0.36$): neuro-typical > persons with aphasia	$F(1,64) = 53.69$, $p < 0.0001$, $\eta^2 = 0.46$): neuro-typical > persons with aphasia	NS	NS
Discourse	($F(2,128) = 80.32$, $p < 0.0001$, $\eta^2 = 0.56$): narrative > picture description and procedural	($F(2,128) = 59.43$, $p < 0.0001$, $\eta^2 = 0.48$): narrative > picture description and procedural	NS	NS
Word class	($F(1,64) = 26.3$, $p < 0.0001$, $\eta^2 = 0.29$): nouns > verbs	($F(1,64) = 9.82$, $p = 0.003$, $\eta^2 = 0.13$): nouns > verbs	($F(1,53) = 1083.63$, $p < 0.0001$, $\eta^2 = 0.95$): nouns > verbs	NS
Group × discourse	($F(2,128) = 11.4$, $p < 0.0001$, $\eta^2 = 0.15$): narrative by neuro-typical group > aphasia group	NS	NS	NS
Group × word class	NS	NS	NS	NS
Discourse × word class	($F(2,128) = 20.06$, $p < 0.0001$, $\eta^2 = 0.24$): nouns > verbs during narrative	($F(2,128) = 13.99$, $p < 0.0001$, $\eta^2 = 0.18$): nouns > verbs during narrative	NS	NS
Group × discourse × word class	NS	($F(2,128) = 7.1$, $p = 0.001$, $\eta^2 = 0.1$): verbs > nouns in procedural discourse by the aphasia group	NS	NS

Significant effects ($p < 0.01$) of 2 (group: neuro-typical vs. aphasia) × 3 (discourse: narrative vs. picture description vs. procedural) × 2 (word class: nouns vs. verbs) mixed ANOVA. NS = Not significant.

All significant main and interaction effects were further explored using post hoc t-tests and Bonferroni corrected for multiple comparisons ($p < 0.01$).