Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Neuropsychologia 148 (2020) 107651

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: http://www.elsevier.com/locate/neuropsychologia

Naming errors and dysfunctional tissue metrics predict language recovery after acute left hemisphere stroke

Erin L. Meier [a,*], Shannon M. Sheppard [a,b], Emily B. Goldberg [a], Catherine R. Head [a], Delaney M. Ubellacker [a], Alexandra Walker [a], Argye E. Hillis [a,c,d]

[a] Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, MD, USA
[b] Department of Communication Sciences and Disorders, Chapman University, Irvine, CA, USA
[c] Department of Physical Medicine and Rehabilitation, Johns Hopkins University School of Medicine, Baltimore, MD, USA
[d] Department of Cognitive Science, Johns Hopkins University, Baltimore, MD, USA

ARTICLE INFO

Keywords:
Left hemisphere stroke
Language
Naming errors
Longitudinal
Recovery
Structural imaging

ABSTRACT

Language recovery following acute left hemisphere (LH) stroke is notoriously difficult to predict. Global language measures (e.g., overall aphasia severity) and gross lesion metrics (e.g., size) provide incomplete recovery predictions. In this study, we test the hypothesis that the types of naming errors patients produce, combined with dysfunctional brain tissue metrics, can provide additional insight into recovery following acute LH stroke. One hundred forty-eight individuals who were hospitalized with a new LH stroke completed clinical neuroimaging and assessments of naming and global language skills. A subset of participants again completed language testing at subacute, early (5–7 months post-stroke), and late (≥11 months post-stroke) chronic phases. At each time point, we coded naming errors into four types (*semantic*, *phonological*, *mixed* and *unrelated*) and determined error type totals and proportions. Dysfunctional tissue measures included the percentage of damage to language network regions and hypoperfusion in vascular territories. A higher proportion of *semantic* errors was associated with better acute naming, but higher proportions of other error types was related to poorer accuracy. Naming and global language skills significantly improved over time , but naming error profiles did not change. Fewer acute *unrelated* errors and less damage to left angular gyrus resulted in optimal naming and language recovery by the final testing time point, yet patients with more acute errors and damage to left middle temporal gyrus demonstrated the greatest increases in language over time. These results illustrate that naming error profiles, particularly *unrelated* errors, add power to predictions of language recovery after stroke.

1. Introduction

In the first days following stroke, approximately 30% of survivors experience aphasia (Flowers et al., 2016; Laska et al., 2001; Pedersen et al., 2003; Wade et al., 1986), a language disorder characterized by receptive and expressive language deficits. Deficits persist into the chronic phase of recovery (around six months post-stroke and onwards) in upwards of 30% of these individuals (Engelter et al., 2006; Flowers et al., 2016). Despite endeavors to identify factors that determine aphasia prognosis (Laska et al., 2001; Lazar et al., 2008; Lazar and Boehme, 2017; Pedersen et al., 2003; Plowman et al., 2012), the ability to predict whether a patient will recover following acute left hemisphere (LH) stroke remains elusive. Lazar et al. (2008) found that the combined variables of age, lesion size, and initial aphasia severity at the acute phase explain only a small percentage of variance in language recovery by the 90-day mark (Lazar et al., 2008). Furthermore, recovery mechanisms—specifically reorganization of functional networks (Hillis and Heidler, 2002)—continue beyond the subacute phase, and as such, identifying specific neural and behavioral factors that correspond to optimal acute to chronic longitudinal outcomes is crucial.

1.1. Anomia and naming errors as markers of severity

Anomia, or impaired word retrieval, is the most persistent deficit in long-term aphasia (Goodglass and Wingfield, 1997). While psycholinguistic models of lexical access (Dell et al., 1997; Dell & O'Seaghdha, 1992; Foygel and Dell, 2000; Indefrey and Levelt, 2004; Levelt et al., 1999; Rapp and Goldrick, 2000; Schwartz et al., 2006) differ in terms of

* Corresponding author.
E-mail address: emeier5@jhmi.edu (E.L. Meier).

https://doi.org/10.1016/j.neuropsychologia.2020.107651
Received 8 July 2020; Received in revised form 29 September 2020; Accepted 2 October 2020
Available online 9 October 2020
0028-3932/

information flow and levels of representation, most models propose separate semantic and phonological processing stages (cf. theories more aligned with parallel distributed processing accounts, e.g., Lambon Ralph et al., 2002; Plaut, 1995). It is believed that the types of errors patients make during naming attempts (along with patterns of performance across different tasks) reflect impairments within distinct semantic or phonological systems.

For example, semantic errors during object naming are those that are conceptually related to the target item and often manifest as category coordinate (e.g., *cat* for "rabbit"), category superordinate (e.g., *vegetable* for "asparagus") or semantic associate (e.g., *milk* for "cow") errors. In post-stroke aphasia, these errors are attributed to deficits in access or retrieval of semantic information, weak connections between semantic and lexical levels, and/or degraded lexical representations (Caramazza and Hillis, 1990; Dell et al., 1997; Foygel and Dell, 2000; Hillis and Caramazza, 1995; Jefferies et al., 2007, 2008; Jefferies and Lambon Ralph, 2006; Mirman and Britt, 2013; Schwartz et al., 2006). Phonological paraphasias most prototypically are non-word errors that either have some degree of phonemic overlap with target items (e.g., *kittel* for "kitten") or are neologisms with little to no phonemic overlap with the target (e.g., *flugger* for "kitten"). These errors most often occur due to impaired access to phonological codes or errors in phonological assembly (Dell et al., 1997; Dell & O'Seaghdha, 1992; Schwartz et al., 2006). Mixed paraphasias (e.g., *cat* for "rat") have semantic and phonemic similarity with the target item and occur due to breakdown at the semantic-word level (Dell et al., 1997; Foygel and Dell, 2000; Schwartz et al., 2006) or due to cascading activation and/or feedback connectivity between lexical and phonological layers (Rapp and Goldrick, 2000). Dell et al. (2004) proposed that omissions, or failures at attempting naming altogether, can occur due to breakdown at any point in the processing stream, yet some evidence exists for a semantic locus for such errors (Chen et al., 2019; Lambon Ralph et al., 2002).

Prior evidence indicates that certain types of object naming errors are more sensitive to anomia severity and are more likely to decrease with recovery than other error types. Specifically, low naming accuracy has been linked to a greater number of non-single word responses (omissions and circumlocutions) and words unrelated to the target item (Le Dorze and Nespoulous, 1989; Mitchum et al., 1990; Moerman et al., 1983; Schuell and Jenkins, 1961; Schwartz and Brecher, 2000). Based on the weight and decay computational model by Dell et al. (1997), Schwartz and Brecher (2000) proposed—and found evidence suggesting—that formal errors (i.e., phonologically-related real words), nonwords and unrelated word errors occur most often in severe anomia but that semantic and mixed errors span the range of anomia severity. The literature (e.g., Basso et al., 1996; Mitchum et al., 1990; Schwartz and Brecher, 2000) generally suggests that severe anomia and production of predominantly semantic errors do not go hand in hand (cf. Rapp and Goldrick, 2000). As naming recovers over time, patients produce higher proportions of responses that are either semantically or phonologically related to target items (Basso et al., 1996; Crary and Kertesz, 1988; Kohn and Smith, 1994; Mitchum et al., 1990; Schwartz and Brecher, 2000), although error profile evolution varies between patients (Crary and Kertesz, 1988; Schwartz and Brecher, 2000). Naming deficit profiles identified early in recovery, such as the types of errors patients produce, also may be informative in predicting future language gains, although no study has explicitly tested this hypothesis.

1.2. Brain structure integrity and post-stroke language recovery

Although gross lesion metrics (e.g., size, stroke laterality) have been implicated in post-stroke recovery in general, the integrity of specific brain structures is critical for longitudinal recovery of specific language abilities—like object naming—in post-stroke aphasia. A key mechanism of recovery in acute stroke is reperfusion of hypoperfused tissue (Hillis and Heidler, 2002). With regards to language, Hillis et al. (2001b) found that reperfusion of tissue within Brodmann Area (BA) 22 (Wernicke's area, including most of the left superior temporal gyrus (LSTG)) caused improvement in lexical-semantics for production and comprehension tasks in hyperacute stroke. In a subsequent study, Hillis et al. (2006) reported that early reperfusion of BA 22, BA 37 (i.e., left fusiform (LFuG), posterior middle (LpMTG) and inferior temporal (LpITG) gyri) and BA 44/45 (i.e., inferior frontal gyrus, pars opercularis (LIFGop) and pars triangularis (LIFGtri)) resulted in improved confrontation naming specifically. At a further level of granularity, DeLeon and coworkers (2007) reported that impaired naming due to deficits in amodal semantic processing correlated with dysfunctional tissue (damage and/or hypoperfusion) in BA 22, whereas impaired naming due to lexical access deficits was related to damage/hypoperfusion in BA 37; overall impaired oral naming corresponded to dysfunctional tissue in BA 22, 37 and 39 (i.e., angular gyrus (AG)).

Reorganization of systems-level brain structure-function relationships and the forging of new pathways and compensatory mechanisms mediate continued recovery in the weeks into years following stroke (Hillis and Heidler, 2002). With regards to long-term recovery, Hillis et al. (2018) reported that improvement in overall naming from acute to chronic post-stroke stages was linked to less damage in posterior LSTG (LpSTG) and to the left arcuate fasciculus (LAF). The integrity of LAF also has been implicated in acute to chronic longitudinal recovery from aphasia in general (Hosomi et al., 2009; Jang and Lee, 2014; Kim and Jang, 2013) and in response to a variety of therapies in individuals with chronic aphasia (Breier et al., 2011; Holland et al., 2016; Schlaug et al., 2009; van Hees et al., 2014). Other studies in chronic aphasia have highlighted the importance of ventral white matter tracts, such as the left inferior longitudinal fasciculus (LILF) (Bonilha et al., 2016; McKinnon et al., 2017; Meier et al., 2019) in response to naming therapy. Cortical necrosis also impacts naming therapy outcomes: Fridriksson (2010) found that the strongest predictor of poor response to naming treatment was damage at the junction of BA 37, 39 and 19 in LpMTG.

These collective studies indicate that the integrity of specific brain structures, particularly the LAF and many temporal lobe areas, is crucial for longitudinal naming recovery. Naming errors provide a different—but also valuable—method to index anomia severity and recovery. Yet, the relative utility of brain structure and naming error metrics in capturing acute to chronic language recovery following LH stroke is unknown.

1.3. The current study

In this study, we propose that the types of object naming errors patients produce, combined with dysfunctional brain tissue metrics, can provide insight into recovery after a new LH stroke. Our primary goals pertain to stroke recovery, but we also investigated relationships between various language and demographic variables at the acute post-stroke stage to situate the longitudinal results. We addressed the following questions: 1) What are the relationships between naming error profiles, demographic variables (sex, age, and education), and naming accuracy in acute stroke? 2) Do global language abilities, overall naming accuracy, and naming error profiles change from early (acute/subacute) to later post-stroke recovery stages? 3) To what extent do error totals, demographic variables, and dysfunctional tissue metrics from an earlier stroke recovery time point predict future language skills and longitudinal change?

We hypothesized that a high proportion of errors that are related to the target (particularly in semantics) would be associated with better acute naming and predict future recovery and language gains, whereas a high proportion of unrelated errors would be associated with poorer acute and later-stage language recovery and less improvement. We also hypothesized that naming accuracy and global language skills would improve and that the proportion of unrelated errors would decrease from early to later post-stroke stages. Finally, we predicted that when considered together, acute naming errors and dysfunctional tissue

metrics (i.e., percentage of damage to regions of interest, hypoperfusion in vascular territories, and total lesion volume) would independently predict later-stage language skills and change in language abilities over time.

2. Materials and methods

2.1. Participants

As part of an ongoing longitudinal stroke recovery project, we studied patients who presented at Johns Hopkins Hospital or Johns Hopkins Bayview Medical Center within 48 h of onset of a new LH stroke. For the current study, exclusionary criteria were lack of premorbid proficiency in English, uncorrected vision and/or hearing, factors preventing participation in testing procedures (e.g., reduced level of consciousness, intubation), and a history of dementia or other neurological disease affecting the brain (excluding stroke). Our final sample included 148 individuals (69 women; mean age: 60.74 ± 13.23 years), seen between September 2012 and June 2020.[1] Of note, 51 participants had a history of prior stroke per MRI report; these participants were included to achieve a sample representative of the larger stroke population, as more than 20% of Americans who experience a stroke each year have a history of prior infarct (Benjamin et al., 2019). In subsequent sections, we describe the steps we took to account for multiple strokes in our analyses. Demographic information, including age, handedness, and years of education, is summarized in Table 1.

Study protocols adhere to the Declaration of Helsinki and were approved by the Institutional Review Board of Johns Hopkins University. Participants or their healthcare proxies (for patients with impaired language comprehension) provided written informed consent to all study procedures.

2.2. Language assessment

Participants completed a language battery at acute (≤11 days, n = 140), subacute (2 weeks-5 months, n = 40), early chronic (5–7 months, n = 41) and/or late chronic (≥11 months, n = 33) post-stroke stages. At each study time point, we administered either the Western Aphasia Battery-Revised (WAB-R; Kertesz, 2007) or Boston Diagnostic Aphasia Examination-3 (BDAE-3; Goodglass et al., 2001) to capture patients' global language deficits. We used scores from similar subtests of verbal expression and auditory comprehension from each measure[2] to generate within-measure language z-scores at early (acute/subacute) and late (early/late chronic) time points. We calculated z-scores according to the standard formula ($[x - \mu]/\sigma$), where x was the patient's subtest score at a given time point and μ and σ were the sample mean and standard deviation, respectively, for a given subtest across time points. We averaged subtest z-scores to generate a single global language z-score for each participant at each time point. If a participant had multiple timepoints (e.g., acute and subacute or both chronic time points), we incorporated the earliest (time point 1) and latest (time point 2) data points into the calculations.

We used the 30-item version of the Boston Naming Test (BNT) (Fisher et al., 1999) to measure confrontation naming abilities. Each item was scored as either correct (1) or incorrect (0). For incorrect items, a trained speech-language pathologist (CRH) coded errors using the 10 classifications from Goodglass et al. (2001) and conferred with other study authors (ELM and SMS) regarding challenging cases. In the event that a patient provided multiple responses, the final, most complete response was coded. We then condensed errors into four main error types. *Semantic* errors included verbal paraphasias semantically-related to the target such as category coordinates (e.g., "hammer" for *saw*, "squirrel" for *beaver*) and semantic associates (e.g., "explosion" for *volcano*, "pyramid" for *sphinx*) as well as circumlocutions that described a semantic feature of the target (e.g., "used in math" for *protractor*). *Phonological* errors included nonword and real-word phonemically--based paraphasias with greater than 50% preservation of target phonology (including phonemic transpositions and phonemes within one position of its target position) and neologisms with less than 50% overlap with the phonology of the target word. *Mixed* errors included single-word responses that were semantically related to the target and retained approximately 50% of the target word phonology (e.g., "elevator" for *escalator*) and multi-word responses that included a semantic and phonological error, as described previously. *Unrelated* errors included verbal (real word) paraphasias unrelated to the target, multi-word paraphasic errors, off-topic utterances including omissions (e.g., "I don't know"), perseverations, and perceptual errors (e.g., "ghost" for *octopus*).

Another trained speech-language pathologist (EBG) who was blinded to the original error codes also completed error-coding using the four-code system. Interrater reliability according to Cohen's weighted kappa was excellent, ranging from 0.911 to 0.979 ($p < 0.0001$) across time points. For each participant, we calculated error totals as well as the proportion of each error type (sum of each error type divided by the total number of errors) at every time point. In certain analyses, we used error type proportions (rather than error totals) in order to mitigate the confound of anomia severity on the results. Table 1 provides a summary of language performance at each time point.

2.3. Neuroimaging data

During their acute hospitalization, participants underwent a clinical imaging protocol that included diffusion-weighted imaging (DWI), T1-weighted, and fluid attenuated inversion recovery (FLAIR) sequences.[3] Trained technicians manually delineated acute lesions slice-by-slice on DWIs for patients in MRIcron (https://www.nitrc.org/projects/mricron). We traced prior strokes on FLAIR scans for participants with a history of previous infarct per the clinical MRI report.

Images were normalized to MNI space using SPM12 (Statistical Parameter Mapping; https://www.fil.ion.ucl.ac.uk/spm/software/spm12/). For acute lesions, the DWI b0 image was warped to a template from older healthy adults (Rorden et al., 2012), and normalization parameters were applied to the corresponding acute lesion maps. For prior strokes, we used the MR segment-normalize routine in Clinical Toolbox (Rorden et al., 2012). Within this pipeline, FLAIR images and the corresponding lesion maps were coregistered to T1-weighted images; the lesion was masked out via enantiomorphic warping; and segmentation and bias correction routines from SPM12 were applied to the anatomical image. The corrected T1-weighted image and lesion map were warped to the same template used in normalization of the DWIs.

For patients with prior stroke, we combined the normalized acute and chronic lesion maps into one total lesion map in MarsBaR (Brett et al., 2002). We used NiiStat (https://www.nitrc.org/projects/niistat/) to determine acute and total stroke volumes. Most prior strokes were small (mean size: 1.944 ± 2.766 cc) in the participants who completed follow-up testing and were included in brain-behavior prediction

[1] We remotely administered follow-up assessments to one participant because we were unable to complete in-person testing during the COVID-19 pandemic.

[2] WAB-R subtests included Information Content; Fluency, Grammatical Competence, and Paraphasias; Yes/No Questions; Auditory Word Recognition; Sequential Commands; Repetition; Object Naming; Word Fluency; Sentence Completion; and Responsive Speech. BDAE-3 subtests included Simple Social Responses; Word Comprehension; Complex Ideational Material; Automatized Sequences; Repetition of Single Words; Repetition of Sentences; and Responsive Naming.

[3] Four participants had contraindications for MRI and instead received a CT scan. Due to differences between CT and MR-DWI in detecting acute stroke, these patients were excluded from imaging analyses.

Table 1
Demographic and language information.

	Time 1			Time 2		
	n	Range	Mean (SD)	*n*	Range	Mean (SD)
Age	148	28.05–91.43	60.74 (13.23)	–	–	–
Years of education	145	5–21	13.89 (2.88)	–	–	–
Handedness (R:L)	136:12	–	–	–	–	–
BNT (total correct)	148	0–30	19.32 (8.99)	63	0–30	22.48 (8.02)
Language z-scores	130	−3.27–0.63	−0.100 (0.938)	61	−2.14–0.62	0.199 (0.660)

Notes: *n* reflects the number of participants for whom data were available at each time point. Time 1 corresponds to the earliest data point, equivalent to the acute time point for 140 patients and the subacute time point for 8 patients. Time 2 corresponds to the latest data point in participants with follow-up testing, either the subacute (n = 11), early chronic (n = 19) or late chronic (n = 33) time point. R = right, L = left.

analyses. Acute and total lesion volumes did not significantly differ in the subgroup of participants with multiple strokes who completed follow-up testing ($t = -0.168$, $p = 0.868$). Nonetheless, to ensure that prior stroke did not unduly influence the results, we extracted the percentage of damaged tissue in regions of interest (ROIs) from NiiStat based on the intersection of combined (acute plus prior stroke) lesion maps with template regions for patients with prior stroke. We also controlled for total lesion volume in all lesion analyses. ROIs included 21 cortical regions implicated in aphasia recovery (Fridriksson et al., 2018), including left middle frontal gyrus (LMFG), LIFGtri, LIFGop, LIFG pars orbitalis (LIFGorb), anterior insula (LaInsula), posterior insula (LpInsula), portions of the basal ganglia (putamen, globus pallidus), precentral gyrus (PrCG), postcentral gyrus (PoCG), LSTG, LMTG, LITG, the superior and middle temporal poles (LSTpole, LMTpole, respectively), LpSTG, LpMTG, LpITG, supramarginal gyrus (LSMG), LAG, and middle occipital gyrus (LMOG) extracted from the Johns Hopkins University (JHU) atlas (Faria et al., 2012; Mori et al., 2008) as well as four white matter tracts, i.e., LAF, LILF, inferior fronto-occipital (LIFOF), and uncinate (LUF) fasciculi, extracted from the atlas by Catani and Thiebaut de Schotten (2008).

In addition to infarct visible on DWI, hypoperfusion surrounding the infarct (the ischemic penumbra) contributes to deficit profiles in acute stroke (Beaulieu et al., 1999; Hillis et al., 2006, 2008; Hillis et al., 2001a, b; Sorensen et al., 1996). Although perfusion-weighted imaging (PWI) is often included in clinical scanning protocols in suspected infarcts, it is not always performed due to time constraints and/or contraindication for intravenous contrast. Due to lack of flow void, vessels that supply hypoperfused regions appear hyperintense on T2-weighted sequences such as FLAIR. As such, as an alternative to PWI, we estimated hypoperfusion in six vascular territories within the left hemisphere (anterior cerebral artery, posterior cerebral artery, middle cerebral artery [MCA]-frontal, MCA-temporal, MCA-insular, MCA-parietal) using the NIH-FLAIR Hyperintense Vessel (FHV) scoring system (Reyes et al., 2017). Within this system, raters provide a score for each vascular territory that reflects the number of FHVs, where 0 = 0 FVHs, 1 = 1–2 FHVs and 2 = 3 or more FHVs.

2.4. Statistical analyses

All analyses were conducted in R (R Core Team, 2020). To address aim #1, we determined relationships between acute language variables (i.e., error type proportions and total correct on the BNT) and demographic variables using non-parametric statistics, either Wilcoxon rank sum tests (for sex) or Spearman correlations (for age and years of education). To further explore acute naming error profiles, we used Spearman correlations to determine the relationships between the different types of acute error proportions as well as associations between error type proportions and naming accuracy.

To address aim #2, we determined change in naming over time using a binomial linear mixed effects model predicting item accuracy (coded as 1s and 0s) from fixed effects of time point (acute, subacute, early chronic, and late chronic), lexical frequency of items (from Brookshire and Nicholas, 1995), age, and education with random intercepts for participants and items. We determined change in global skills by conducting a Welch's two-sided *t*-test. We used analysis of covariance (ANCOVA) to determine if error type proportions—reflecting different naming error profiles—changed from early (acute/subacute) to later time points, controlling for the percentage of coded errors, days between testing time points, age, and education.

To address aim #3, we first conducted four linear regression analyses to determine if the total number of each error type produced at an early recovery stage (acute/subacute) predicted either: 1) later-stage naming accuracy (i.e., total correct items on the BNT at the latest testing time point), 2) global language skills (i.e., z-scores at the latest testing time point), 3) change in naming accuracy (i.e., the difference in BNT total correct at the latest versus the earliest time points), or 4) change in global language skills (i.e., the difference in z-scores at latest versus the earliest testing time points). Within each of these models, we controlled for days between testing time points, age, years of education, and the percentage of coded errors.

To gain insight into neurobehavioral profiles of recovery, we conducted four Least Absolute Shrinkage and Selection Operator (LASSO) regression analyses, predicting either later-stage language skills (naming accuracy or global language skills) or change in language (naming or global language skills) from first to later time points. Predictors included error totals and demographic variables significant in prior analyses, dysfunctional tissue metrics (i.e., percentage of damaged tissue in ROIs, FHV ratings and total lesion volume) and nuisance variables (days between testing time points, percentage of coded errors). LASSO regression is suitable for models with predictors that exhibit high levels of multicollinearity, which was likely in the present study due to the fact that adjacent ROIs often are damaged in conjunction. LASSO regression adds a penalty that shrinks coefficients towards zero, resulting in a simple model with maximal prediction capacity. We ran 5000 permutations and a cross-fold validation equal to *n* per LASSO model. We excluded brain metrics (percent damage to ROIs, FHV ratings) for which <10% of the sample had values indicating damage or dysfunction.

For results significant at an uncorrected threshold, we performed multiple comparison correction using the false discovery rate ($q < 0.05$) at the model level across correlation, ANCOVA and linear regression analyses.

3. Results

3.1. Relationships between error type proportions, demographic variables and naming in acute stroke

At the acute post-stroke time point, naming errors were recorded for

115 participants.[4] On average, participants produced a greater proportion of *semantic* (62.15 ± 36.02%) and *unrelated* (25.25 ± 31.28%) errors compared to *phonological* (10.97 ± 18.04%) and *mixed* (1.57 ± 7.09%) errors. Error type proportions did not differ between men and women ($q \geq 0.572$) and were not related to either age ($q \geq 0.120$) or years of education ($q \geq 0.359$) (see Supplemental Table 1 for complete results). Acute BNT accuracy did not vary by sex ($W = 1538.5$, $p = 0.534$) or age ($r = 0.009$, $p = 0.926$), but there was a trending, weak association between total correctly-named items and years of education ($r = 0.167$, $p = 0.078$).

A higher proportion of *semantic* errors was strongly associated with lower proportions of *phonological* ($r = -0.572$, $q < 0.001$) and *unrelated* ($r = -0.824$, $q < 0.001$) errors and weakly associated with a lower proportion of *mixed* errors ($r = -0.261$, $q = 0.009$). A greater proportion of phonological errors weakly correlated with a greater proportion of *mixed* errors ($r = 0.241$, $q = 0.014$). We found an insignificant relationship between *unrelated* and *mixed* error proportions ($r = 0.037$, $q = 0.698$) but a trending association between *unrelated* and *phonological* error rates ($r = 0.175$, $q = 0.073$). A higher proportion of *semantic* errors was associated with better naming accuracy ($r = 0.550$, $q < 0.001$), whereas higher proportions of *phonological* ($r = -0.367$, $q < 0.001$), *mixed* ($r = -0.219$, $q = 0.019$), and *unrelated* ($r = -0.554$, $q < 0.001$) errors were associated with poorer naming.

3.2. Change in language over time

The linear mixed effects model revealed overall naming accuracy significantly improved from early to later recovery stages ($\beta = 0.597$, $SE = 0.046$, $t = 13.003$, $p < 0.001$), controlling for lexical frequency of target items ($\beta = 0.007$, $SE = 0.003$, $t = 2.607$, $p = 0.009$), age ($\beta = 0.013$, $SE = 0.022$, $t = 0.577$, $p = 0.564$), and education ($\beta = 0.334$, $SE = 0.104$, $t = 3.222$, $p = 0.001$). Global language skills (according to summary z-scores) also significantly improved from early to later recovery stages ($t_{56} = -4.859$, $p < 0.001$). In contrast, controlling for the percentage of coded errors per time point, days post-stroke onset, age and years of education, the proportion of *semantic* ($F_{1,92} = 1.176$, $p = 0.281$), *phonological* ($F_{1,92} = 0.012$, $p = 0.914$), *mixed* ($F_{1,92} = 1.604$, $p = 0.209$), and *unrelated* ($F_{1,92} = 0.782$, $p = 0.379$) errors that patients produced did not change significantly over time (see Supplemental Table 2 for full ANCOVA models). The null findings may have been caused by patients with mild anomia with higher proportions of *semantic* and lower proportions of other error types, leading to little room for group-level change. As such, we re-ran the analysis only with patients whose early-stage naming accuracy fell within the lowest quartile, i.e., fewer than 15 items correct ($n = 18$). However, there was still no significant effect of time for any error type ($p \geq 0.086$) in the subsample. Supplemental Table 3 presents summary data for overall language skills, naming, and error types for each discrete time point (i.e., acute, subacute, 6-months, and 12-months post-stroke) for the longitudinal errors sample.

3.3. Predictions of language recovery and gains by error types, demographic variables and dysfunctional tissue measures

In the remainder of analyses, we used the total number of each error type—rather than error type proportions—in order to capture the severity of underlying impairments. *Mixed* errors were excluded from the analyses as no patients produced more than one *mixed* error. Therefore, predictors within the multivariable regression analyses included the total number of *semantic*, *phonological*, and *unrelated* errors, in addition to demographic variables previously-cited as critical for aphasia recovery (age and education), days between testing time points, and percentage of coded errors.

The regression model predicting later-stage naming accuracy (i.e., total correct on the BNT at the latest time point for each patient) was significant ($F_{7,49} = 9.608$, $q < 0.001$, adjusted $R^2 = 0.518$). Within this model, controlling for other factors, fewer total *unrelated* errors produced early in recovery ($\beta = -0.762$, $SE = 0.124$, $t = -6.155$, $p < 0.001$) and more years of education ($\beta = 0.831$, $SE = 0.293$, $t = 2.834$, $p = 0.007$) predicted better future naming. The multivariable model predicting recovery of global language skills (i.e., z-scores at the latest testing time point) also was significant ($F_{7,46} = 5.642$, $q < 0.001$, adjusted $R^2 = 0.380$), and within this model, fewer early-stage *unrelated* errors ($\beta = -0.055$, $SE = 0.011$, $t = -4.865$, $p < 0.001$) predicted better later-stage overall language abilities. See Table 2 for the full results of regression models predicting recovery of naming and global language skills.

The multivariable model predicting change in naming over time (i.e., total correct on the BNT at the latest minus the earliest time points) was significant ($F_{7,49} = 7.694$, $q < 0.001$, adjusted $R^2 = 0.456$). Controlling for other model factors, more early-stage *semantic* ($\beta = 0.737$, $SE = 0.248$, $t = 2.976$, $p = 0.005$), *phonological* ($\beta = 1.265$, $SE = 0.361$, $t = 3.503$, $p < 0.001$) and *unrelated* ($\beta = 0.423$, $SE = 0.103$, $t = 4.106$, $p < 0.001$) errors—as well as more years of education ($\beta = 0.858$, $SE = 0.244$, $t = 3.520$, $p < 0.001$)—predicted greater naming improvement. The model predicting change in global language abilities (i.e., language z-scores from the latest minus the earliest time points) was significant ($F_{7,42} = 4.808$, $q < 0.001$, adjusted $R^2 = 0.352$) as well; controlling for other factors, the only significant independent predictor was *unrelated* errors, such that more *unrelated* errors at an early stage predicted greater gains in overall language skills ($\beta = 0.058$, $SE = 0.011$, $t = 5.225$, $p < 0.001$). Complete results for the longitudinal change prediction models are reported in Table 3.

Fig. 1 shows the lesion overlay for the patients included in the follow-up analyses including dysfunctional tissue measures. Of note, 12 participants had some degree of damage in the right hemisphere (in addition to left hemisphere damage), mostly within deep subcortical structures. Nine participants had infarct in either the cerebellum or brainstem in addition to cerebral involvement. To inform the LASSO regressions, we used ANOVA to determine if language scores varied by broad lesion classifications (i.e., LH-only, bilateral, or other [cerebellum or brainstem with cerebral involvement]). We found no significant relationships between lesion group and BNT total correct at the latest time point ($W = 221.5$, $p = 0.680$), language z-scores at the latest time point ($W = 223$, $p = 0.588$), change in BNT ($W = 259$, $p = 0.672$) or change in z-scores ($W = 135$, $p = 0.338$). As such, we did not include lesion classification in the LASSO regressions.

Table 4 includes the complete results from the LASSO models. Controlling for other model variables, we found that fewer acute *unrelated* errors, more years of education, and less damage to LAG significantly predicted better recovery of naming by the final testing time point. Less damage to LAG was the sole significant predictor of recovery of global language abilities, controlling for other factors. The model predicting change in naming abilities over time included many variables; controlling for these factors, more acute *unrelated* errors and less damage to LAG predicted greater naming gains. On the other hand, patients who produced a greater number of acute *unrelated* errors and had damage in LpMTG demonstrated the greatest improvement in global language skills from early to later stroke recovery stages.

4. Discussion

We investigated the relationship between types of naming errors, dysfunctional brain tissue, and acute to chronic post-stroke recovery of naming and language abilities. In acute stroke, a greater proportion of *semantic* errors was associated with higher naming accuracy, but the

[4] Of the remaining 25 participants who completed the BNT at the acute time point, four participants attained perfect scores (30/30). Errors were not recorded in their majority for the remainder of participants, and as such, these individuals were excluded from all subsequent error analyses.

Table 2
Error and demographic predictors of later-stage language abilities.

	Multivariable model							
Dependent variable	(Df) F-stat	*q-value*	Adj. R^2	Univariate predictor	β	SE	t-stat	*p-value*
BNT total at T2	(7,49) 9.608	<0.001	0.518	Semantic	−0.465	0.298	−1.561	0.125
				Phonological	0.151	0.434	0.347	0.730
				Unrelated	−0.762	0.124	−6.155	<0.001
				T2-T1days	0.001	0.004	0.392	0.697
				Age	0.041	0.063	0.647	0.521
				Education	0.831	0.293	2.834	0.007
				%coded	0.127	0.033	3.879	<0.001
Language z-score at T2	(7,46) 5.642	<0.001	0.380	Semantic	0.016	0.027	0.592	0.557
				Phonological	0.045	0.040	1.112	0.272
				Unrelated	−0.055	0.011	−4.865	<0.001
				T2-T1days	−0.0002	0.0003	−0.475	0.637
				Age	0.005	0.006	0.883	0.382
				Education	0.047	0.028	1.683	0.099
				%coded	0.007	0.003	2.397	0.021

Notes: BNT = Boston Naming Testing, T2 = second testing time point, T1 = first testing time point, T2-T1days = number of days between testing time points, %coded = percentage of errors coded at time point 1. Semantic/Phonological/Unrelated reflect error totals of each type at T1.

Table 3
Error and demographic predictors of longitudinal change in language abilities.

	Multivariable model							
Dependent variable	(Df) F-stat	*q-value*	Adj. R^2	Univariate predictor	β	SE	t-stat	*p-value*
Δ in BNT total (T2-T1)	(7,49) 7.694	<0.001	0.456	Semantic	0.737	0.248	2.976	0.005
				Phonological	1.265	0.361	3.503	<0.001
				Unrelated	0.423	0.103	4.106	<0.001
				T2-T1days	0.005	0.003	1.639	0.108
				Age	0.0007	0.052	0.014	0.989
				Education	0.858	0.244	3.520	<0.001
				%coded	−0.072	0.027	−2.658	0.011
Δ in language z-score (T2-T1)	(7,42) 4.808	<0.001	0.352	Semantic	0.031	0.027	1.168	0.249
				Phonological	0.023	0.039	0.596	0.554
				Unrelated	0.058	0.011	5.225	<0.001
				T2-T1days	0.0001	0.0003	0.276	0.784
				Age	0.004	0.006	0.723	0.474
				Education	0.007	0.027	0.269	0.790
				%coded	−0.004	0.003	−1.217	0.230

Notes: Δ = change, BNT = Boston Naming Testing, T2 = second testing time point, T1 = first testing time point, T2-T1days = number of days between testing time points, %coded = percentage of errors coded at time point 1. Semantic/Phonological/Unrelated reflect error totals of each type at T1.

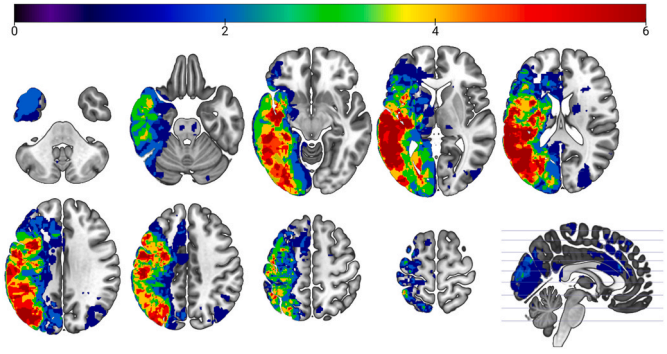

Fig. 1. Lesion overlap. Overlay of lesions across the sample of patients with error data at the acute stage and at least one later time point (n = 46).

opposite pattern (i.e., higher proportions, lower accuracy) was true of other error types. Naming and global language abilities improved from early to later post-stroke recovery stages, but naming error profiles (i.e., error type proportions) did not evolve over time. The strongest and most consistent predictors of language recovery and longitudinal change were the number of *unrelated* errors produced early in recovery, years of education, and the integrity of tissue within temporoparietal cortex.

4.1. Naming error profiles

In order to better understand the role of naming errors in recovery, we first investigated relationships between error type proportions and language and demographic variables in acute stroke. We found that a higher proportion of *semantic* errors produced at the acute time point was significantly associated with lower proportions of other error types and with better naming accuracy. On the other hand, a higher proportion of *phonological* errors coincided with higher proportions of *mixed* and *unrelated* errors, and poorer acute naming abilities correlated with higher proportions of these three error types. In other words, when patients made errors in the acute stage, they were better off if they produced more *semantic* errors relative to other error types.

These results partially align with prior work and also provide novel insights into the nature of semantic errors in acute stroke. In this study, *semantic* errors included verbal paraphasias related to target items (often category coordinates: e.g., *broccoli-asparagus*) and circumlocutions (often a semantic feature of the target). These errors indicate some degree of successful semantic access, which may reflect a stronger lexical-semantic system and a higher likelihood of successful lexical access in general. The strong positive relationship between *semantic* error proportions and naming accuracy supports this interpretation and coheres with data suggesting individuals with mild anomia—and even healthy controls, when they make errors—predominantly produce *semantic*

Table 4
Error type and dysfunctional tissue predictors of language recovery.

Metric	T2 BNT total		T2 Language z-scores		Δ in BNT T2-T1		Δ in z-scores T2-T1	
	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value
T2-T1days	–	–	–	–	0.152	0.502	–	–
Phonological errors	–	–	–	–	0.191	0.361	0.026	0.140
Unrelated errors	**−0.180**	**0.017**	–	–	**0.911**	**0.003**	**0.311**	**0.003**
% coded errors	–	–	–	–	−0.364	0.051	–	–
Years of education	**0.321**	**< 0.001**	0.059	0.257	0.338	0.093	–	–
% damage LIFGop	–	–	–	–	0.178	0.392	–	–
% damage LIFGorb	–	–	–	–	−0.225	0.240	–	–
% damage LPoCG	–	–	–	–	−0.219	0.362	–	–
% damage LPrCG	–	–	–	–	0.091	0.619	–	–
% damage LSMG	–	–	–	–	0.572	0.183	–	–
% damage LAG	**−0.395**	**< 0.001**	**−0.806**	**< 0.001**	**−1.117**	**0.024**	–	–
% damage LMOG	–	–	–	–	0.340	0.332	–	–
% damage LpSTG	–	–			−0.0001	0.527		
% damage L putamen	–	–	–	–	−0.261	0.399	–	–
% damage L globus pallidus	–	–	–	–	0.214	0.242	–	–
% damage LpMTG	–	–	–	–	–	–	**0.229**	**0.002**
% damage LUF	–	–	–	–	−0.423	0.092	–	–
FHV L MCA-frontal	–	–	–	–	0.109	0.519	–	–
FHV L MCA-insula	–	–	–	–	0.155	0.466	–	–
FHV L MCA-parietal	–	–	–	–	0.189	0.448	–	–
Total lesion volume	–	–	–	–	0.019	0.526	–	–

Notes: BNT = Boston Naming Test, T2-T1 days = number of days between first and final testing time points, L = left, p = posterior, IFGop = inferior frontal gyrus, pars opercularis, IFGorb = IFG, pars orbitalis, PoCG = postcentral gyrus, PrCG = precentral gyrus, SMG = supramarginal gyrus, AG = angular gyrus, STG = superior temporal gyrus, MTG = middle temporal gyrus, MOG = middle occipital gyrus, UF = uncinate fasciculus, FHV = FLAIR hyperintense vessel scores, MCA = middle cerebral artery territory. Δ denotes change from time point 1 (T1) to time point 2 (T2). Bold font indicates significant results at $p < 0.05$.

errors (Grima and Franklin, 2017; Nicholas et al., 1989; Tallberg, 2005). However, this result does not reflect that *semantic* errors (or *mixed* errors) are severity-insensitive like some researchers (e.g., Dell et al., 1997; Schwartz and Brecher, 2000) previously proposed; given the relationship we found, it is unlikely that we would find similar *semantic* error totals among patients with mild, moderate and severe anomia. Furthermore, we found a much higher proportion of *semantic* errors compared to other error types, unlike Basso et al. (1996), who reported only 3.6% of errors as being *semantic* in patients within the first four months of recovery. Similarly, only two (of 28) patients made predominantly *semantic* errors within one month of stroke in Mitchum et al. (1990). Reasons for this discrepancy could be differences between studies in stroke duration (i.e., on the order of days in our study versus months in the others), *semantic* error definitions (e.g., inclusion of circumlocutions in our study but not in others, e.g., Schwartz and Brecher, 2000) and anomia severity. The latter point is especially likely since we recruited patients who experienced a new left hemisphere stroke but did not necessitate that they present with persistent anomia and aphasia. A central conclusion that can be drawn from our acute findings is that the production of *semantic* errors alone does not necessarily reflect a severe semantic deficit. Instead, it is more likely that low performance on neuropsychological assessments of semantics, combined with a high total number of *semantic* errors, truly indicate an impaired semantic system.

Our results linking low naming correctness with high *phonological* and *unrelated* error proportions are more consistent with prior work (Le Dorze and Nespoulous, 1989; Mitchum et al., 1990; Moerman et al., 1983; Schuell and Jenkins, 1961; Schwartz and Brecher, 2000), with some important caveats. For example, Schwartz and Brecher (2000) provided more nuanced distinctions between types of *phonological* errors (e.g., close versus remote neologisms, formal (real-word) errors versus nonwords), whereas we grouped these error types together. Although outside the primary scope of this paper, future work delving into differences between *phonological* error types in terms of phonemic overlap and lexicality and their relationship with naming correctness could complement prior findings. Similarly, we classified *unrelated* errors as responses with essentially no semantic or phonological overlap with targets, including many different types of errors: off-target utterances (e.g., "couple of flowers" for *volcano*), omissions (e.g., "I don't know"), perseverations and perceptual errors. Although this error type somewhat served as a catch-all category, this classification method makes conceptual sense, given that patients who produce any of these errors likely have a weaker lexical system and/or demonstrate other deficits that interfere with lexical retrieval. Chen et al. (2019) proposed different probable causes for omission errors, including core deficits within the semantic system (either deficient semantic representations or disconnect between semantic representations and lexical items), impairments in correct selection of competing lexical representations, deficits in the speech motor system and deficits in other perceptual (e.g., visual) or cognitive systems. Based on our limited language battery, we cannot specify the locus of impairment for *unrelated* errors (or the subtypes therein), but doing so represents an important future direction of this work.

In terms of longitudinal change, we found that overall naming ability and global language skills significantly improved from early (acute/subacute) to later post-stroke stages. However, error type proportions did not change over time, even for patients with the most severe naming impairments. It may be that the total number of each error type changes over time (reflecting overall accuracy) but that the types of errors most individual patients produce does not (Capitani and Laiacona, 2004). That being said, prior studies (Basso et al., 1996; Crary and Kertesz, 1988; Kohn and Smith, 1994; Mitchum et al., 1990; Schwartz and Brecher, 2000) have reported evolution in error profiles. Using a different approach than ours, Crary and Kertesz (1988) reported that change in error types over the first year post-stroke mirrors the transition from one aphasia syndrome to another. Schwartz and Brecher (2000) found phonologically-related errors declined from the first to later testing points in patients with chronic aphasia and noted a general shift away from phonological and towards semantic error types. In contrast, Basso et al. (1996) found that the total numbers of no responses and neologisms patients produced during oral naming tasks decreased over time. One possible reason our results conflict with these latter two studies is that we investigated changes in error proportions, rather than error totals, to comment on error profiles without the added confound of anomia severity. Anecdotally, 17 patients did exhibit an increase of at least 20% in semantic error proportions (and a corresponding decrease

in other error types, especially *unrelated* errors) from early to later time points, but 8 patients exhibited the inverse, a shift from *semantic* to *unrelated* errors. The greater implication of our findings is that error profiles might appear resistant to change at a group level, especially in a sample of patients with varying levels of aphasia and anomia severity, like in the current study. On an *individual* level, however, the types of errors patients produce are amenable to change and are likely related to global language recovery to some extent. In future studies, taking a case series approach and including in-depth neuropsychological assessment can aid in determining locus of breakdown and potential underlying causes of different error patterns in individual patients.

4.2. Predictions of longitudinal naming and global language recovery and change

Regardless of the lack of longitudinal change, error types produced at an earlier time point did predict future language recovery by the final testing time point and change in language over time. Specifically, the number of early *unrelated* errors was a significant predictor within all linear regression models, and *semantic* and *phonological* error totals also predicted BNT change. When error totals were considered in conjunction with demographic variables and dysfunctional tissue metrics, the number of acute *unrelated* errors remained a significant predictor in all models excluding the global language recovery analysis. Of note, the direction of the effects varied between recovery (i.e., performance at the latest time point) and change (i.e., the difference between scores at the latest versus earliest time points) models. The naming and global language recovery models indicated that fewer *unrelated* errors produced at an earlier time point predicted better language skills at a later time point. In acute stroke, low naming accuracy was significantly associated with a high proportion of *unrelated* errors, and therefore, it stands to reason that patients with very impaired acute naming skills may still have exhibited the most severe impairments at later recovery stages. On the other hand, patients who produced *more* errors at an earlier time point exhibited the greatest change in language over time. This finding likely reflects that patients who produce more early errors have the greatest room for growth. Most importantly, these results highlight the importance of further investigation into the origins of *unrelated* errors, as identifying their underlying cause can help tailor naming treatment protocols for patients who produce many such errors.

Within the combined error type and dysfunctional tissue models, the integrity of temporoparietal cortex was the most consistent significant neural predictor of recovery. Specifically, less damage to LAG at the acute stage predicted higher naming accuracy and global language scores at a future recovery stage as well as greater naming gains. Many structural and functional connections converge in the left temporoparietal cortex, and this area is considered a critical cross-modal hub for many processes, semantics in particular (Binder et al., 2009; Binder and Desai, 2011; Davey et al., 2015; Humphreys and Lambon Ralph, 2015; Lambon Ralph et al., 2016; Noonan et al., 2013; Seghier, 2013; Whitney et al., 2012; Xu et al., 2016, 2017). As such, it stands to reason that the preservation of tissue within LAG is crucial for receptive and expressive language skills. However, somewhat counterintuitively, patients who had highly-damaged nearby LpMTG demonstrated the greatest improvement in global language skills over time. In a recent lesion symptom mapping study of a largely-overlapping acute cohort, we found that greater damage to LpMTG and higher FHV parietal ratings were associated with acute naming deficits. In this cohort, as expected, more severe acute anomia was significantly associated with more severe global language impairments ($r = 0.750$, $p < 0.001$). Consequently, the association between LpMTG damage and greater language gains could merely reflect a pattern of LpMTG damage and very low acute language performance in certain patients, and that these individuals had more room for improvement than patients with higher acute scores. Alternatively, given LpMTG's role as a network hub and the importance of temporoparietal cortex in normal lexical-semantic processing (Noonan et al., 2013; Xu et al., 2016, 2017)—and in people with chronic aphasia (Fridriksson, 2010; Griffis et al., 2017a, 2017b)—it may be that the brain prioritizes functional reorganization of temporoparietal-mediated processes in the event of acute dysfunction (whether restricted diffusion or hypoperfusion), ultimately leading to better long-term outcomes.

The regression models also yielded some interesting findings regarding the predictive power of other variables. We found no associations between education and global language skills, which aligns with the literature suggesting that education level is not a robust predictor of overall aphasia recovery (Lazar et al., 2008; Plowman et al., 2012; Watila and Balarabe, 2015). Although we found only a trend between years of education and acute naming performance, education level consistently predicted future naming skills. Higher levels of education have been implicated in preservation of lexical-semantics for picture naming in healthy older adults (Connor et al., 2004; Le Dorze and Durocher, 1992; Neils et al., 1995; Paolieri et al., 2018) and in cerebral reserve in the face of brain pathology (Staff et al., 2004). Consistent with the literature (see the review by Watila and Balarabe, 2015), smaller total lesion volume strongly correlated with better acute ($r = -0.721$, $p < 0.001$) and later-stage ($r = -0.530$, $p < 0.001$) global language skills, yet lesion volume was retained only in the BNT change LASSO regression model as an insignificant predictor. The LASSO procedure heightens predictive power by creating a sparse model containing the most optimal predictor variables; therefore, it can be inferred that the other variable combinations—without total lesion volume—provided the best predictions of language recovery and change over time. Lesion size has long been considered a key determinant of longitudinal stroke outcomes. Yet the retention of error variables over lesion volume in the LASSO models indicates that the types of errors patients produced early in recovery—particularly *unrelated* errors—are even better predictors of post-stroke language recovery outcomes. These findings further emphasize the importance of early evaluation of naming and characterization of naming errors in patients following left hemisphere stroke.

4.3. Study limitations

The current study has several limitations. At each time point, certain participants had incomplete language datasets. For some participants, naming errors were not recorded for the entire BNT or were not recorded at all. To compensate for this issue, we controlled for the percentage of coded errors and excluded patients without any recorded errors in analyses. We also were unable to compute language summary z-scores for participants who completed only a handful of subtests from the WAB-R or the BDAE-3, and such participants were excluded from global language recovery analyses.

A related issue is that we were unable to collect data for all participants for all time points. Attrition at the follow-up time points was high (only 57 of 148 patients had longitudinal error data), mostly due to unsuccessful repeated contact attempts or participant transportation difficulties. Additional reasons for attrition after the acute post-stroke stage included subsequent stroke, heart attack, relocation to an assistive living facility and subsequent loss of contact, move to a location out of state to live with family, and conflicts due to rehabilitation schedules. We acknowledge that the high rate of attrition from the acute to subsequent time points is a major limitation of the current study; as such, future work is necessary to validate the current findings.

We were also unable to collect acute behavioral data for eight participants (due to medical issues, early discharge, or other time constraints), and as such, the earliest time point for these patients was the subacute phase. Given that language recovery is most rapid in the acute to subacute post-stroke stages (Hillis and Heidler, 2002), including participants without acute data may have influenced the results. Therefore, we replicated all analyses from aims #2 and #3 that used early stage data (i.e., combined acute and subacute data) and excluded participants without acute behavioral data. The main findings remained unchanged (see Supplementary Material, Tables 4–7).

One key difference between our study and prior studies is that we included participants with a history of prior stroke and participants with brain damage to regions outside of the left cerebral hemisphere. We included these participants because our primary goal was to further understanding of post-stroke language recovery (rather than informing the neurobiology of language or psycholinguistic models, per se) in a sample more representative of the stroke population than those of prior studies. Recent work by our group indicates that history of prior stroke on its own is not a significant determinant of language outcomes but that accounting for total brain damage and lesion location–as we have done in the present study–is crucial.

5. Conclusions

In this study, we found that compared to other error types, a high proportion of *semantic* errors produced at the acute post-stroke stage was associated with more intact naming abilities. While lexical retrieval and global language skills (specifically auditory comprehension and verbal expression) improved from early to later post-stroke stages, the proportions of error types did not change over time. Errors classified as *unrelated*, including omissions, off-topic utterances, perseverations, and perceptual errors, had the best predictive utility of any error type. Indeed, the most predictive factors of longitudinal improvement in naming abilities and global language skills were fewer *unrelated* errors produced early in recovery, in addition to the integrity of LAG. On the other hand, patients with more preserved LAG who produced *more unrelated* errors improved most in naming over time. Patients who demonstrated the greatest gains in global language skills from early to later post-stroke stages also produced a higher number of *unrelated* errors but had damage to the posterior middle temporal cortex, perhaps because they had lower baseline scores and greatest room for improvement. Considering the detrimental impact of *unrelated* errors on language, future studies aimed at identifying the underlying cause of *unrelated* errors in individual patients and determining treatment protocols to shift persistent naming errors away from *unrelated* and towards the *semantic* type are warranted.

Author contributions

Erin L. Meier, Conceptualization, Methodology, Validation, Formal analysis, Investigation, Data curation, Writing - original draft. Shannon M. Sheppard, Conceptualization, Methodology, Validation, Writing - review & editing. Emily B. Goldberg, Validation, Formal analysis, Investigation, Data curation, Writing - review & editing. Catherine R. Head, Methodology, Validation, Data curation, Writing - review & editing. Delaney M. Ubellacker, Investigation, Data curation, Writing - review & editing. Alexandra. Walker, Investigation, Data curation, Writing - review & editing. Argye E. Hillis, Conceptualization, Resources, Writing - review & editing, Supervision, Funding acquisition

Declaration of competing interest

None.

Acknowledgments

This work was supported by the National Institutes of Health, National Institute on Deafness and Other Communication Disorders through grant R01DC005375.

We extend our thanks to members of the Stroke Cognitive Outcomes and REcovery (S.C.O.R.E.) Lab who assisted with participant recruitment and data collection. We also gratefully acknowledge our participants who made this work possible.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neuropsychologia.2020.107651.

References

Basso, A., Corno, M., Marangolo, P., 1996. Evolution of oral and written confrontation naming errors in aphasia. A retrospective study on vascular patients. J. Clin. Exp. Neuropsychol. 18 (1), 77–87. https://doi.org/10.1080/01688639608408264.

Beaulieu, C., De Crespigny, A., Tong, D.C., Moseley, M.E., Albers, G.W., Marks, M.P., 1999. Longitudinal magnetic resonance imaging study of perfusion and diffusion in stroke: evolution of lesion volume and correlation with clinical outcome. Ann. Neurol. 46 (4), 568–578. https://doi.org/10.1002/1531-8249(199910)46:4<568::AID-ANA4>3.0.CO;2-R.

Benjamin, E.J., Muntner, P., Alonso, A., Bittencourt, M.S., Callaway, C.W., Carson, A.P., Chamberlain, A.M., Chang, A.R., Cheng, S., Das, S.R., Delling, F.N., Djousse, L., Elkind, M.S.V., Ferguson, J.F., Fornage, M., Jordan, L.C., Khan, S.S., Kissela, B.M., Knutson, K.L., 2019. … on behalf of the American heart association council on epidemiology and prevention statistics committee and stroke statistics subcommittee. Heart disease and stroke statistics—2019 update: A report from the American Heart Association. Circulation 139 (10). https://doi.org/10.1161/CIR.0000000000000659.

Binder, J.R., Desai, R.H., 2011. The neurobiology of semantic memory. Trends Cognit. Sci. 15 (11), 527–536. https://doi.org/10.1016/j.tics.2011.10.001.

Binder, J.R., Desai, R.H., Graves, W.W., Conant, L.L., 2009. Where Is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. Cerebr. Cortex 19 (12), 2767–2796. https://doi.org/10.1093/cercor/bhp055.

Bonilha, L., Gleichgerrcht, E., Nesland, T., Rorden, C., Fridriksson, J., 2016. Success of anomia treatment in aphasia is associated with preserved architecture of global and left temporal lobe structural networks. Neurorehabilitation Neural Repair 30 (3), 266–279.

Breier, J.I., Juranek, J., Papanicolaou, A.C., 2011. Changes in maps of language function and the integrity of the arcuate fasciculus after therapy for chronic aphasia. Neurocase 17 (6), 506–517. https://doi.org/10.1080/13554794.2010.547505.

Brett, M., Anton, J., Valabregue, R., Poline, J.B., 2002. Region of interest analysis using the MarsBar toolbox for. SPM 99 (2), S497. *16*.

Brookshire, R.H., Nicholas, L.E., 1995. Relationship of word frequency in printed materials and judgements of word frequency in daily life to Boston Naming Test of performance of aphasic adults. Clinical Aphasiology 23, 107–119.

Capitani, E., Laiacona, M., 2004. A method for studying the evolution of naming error types in the recovery of acute aphasia: a single-patient and single-stimulus approach. Neuropsychologia 42 (5), 613–623. https://doi.org/10.1016/j.neuropsychologia.2003.10.006.

Caramazza, A., Hillis, A.E., 1990. Where do semantic errors come from? Cortex 26 (1), 95–122. https://doi.org/10.1016/S0010-9452(13)80077-9.

Catani, M., Thiebaut de Schotten, M., 2008. A diffusion tensor imaging tractography atlas for virtual in vivo dissections. Cortex 44 (8), 1105–1132. https://doi.org/10.1016/j.cortex.2008.05.004.

Chen, Q., Middleton, E., Mirman, D., 2019. Words fail: lesion-symptom mapping of errors of omission in post-stroke aphasia. J. Neuropsychol. 13 (2), 183–197. https://doi.org/10.1111/jnp.12148.

Connor, L.T., Spiro, A., Obler, L.K., Albert, M.L., 2004. Change in object naming ability during adulthood. J. Gerontol. B Psychol. Sci. Soc. Sci. 59 (5), P203–P209. https://doi.org/10.1093/geronb/59.5.P203.

Crary, M.A., Kertesz, A., 1988. Evolving error profiles during aphasia syndrome remission. Aphasiology 2 (1), 67–77. https://doi.org/10.1080/02687038808248888.

Davey, J., Cornelissen, P.L., Thompson, H.E., Sonkusare, S., Hallam, G., Smallwood, J., Jefferies, E., 2015. Automatic and controlled semantic retrieval: TMS reveals distinct contributions of posterior middle temporal gyrus and angular gyrus. J. Neurosci. 35 (46), 15230–15239. https://doi.org/10.1523/JNEUROSCI.4705-14.2015.

DeLeon, J., Gottesman, R.F., Kleinman, J.T., Newhart, M., Davis, C., Heidler-Gary, J., Lee, A., Hillis, A.E., 2007. Neural regions essential for distinct cognitive processes underlying picture naming. Brain 130 (5), 1408–1422. https://doi.org/10.1093/brain/awm011.

Dell, G.S., O'Seaghdha, P.G., 1992. Stages of lexical access in language production. Cognition 42 (1–3), 287–314.

Dell, G.S., Schwartz, M.F., Martin, N., Saffran, E.M., Gagnon, D.A., 1997. Lexical access in aphasic and nonaphasic speakers. Psychol. Rev. 104 (4), 801–838.

Dell, Gary S., Lawler, E.N., Harris, H.D., Gordon, J.K., 2004. Models of errors of omission in aphasic naming. Cogn. Neuropsychol. 21 (2–4), 125–145. https://doi.org/10.1080/02643290342000320.

Engelter, S.T., Gostynski, M., Papa, S., Frei, M., Born, C., Ajdacic-Gross, V., Gutzwiller, F., Lyrer, P.A., 2006. Epidemiology of aphasia attributable to first ischemic stroke: incidence, severity, fluency, etiology, and thrombolysis. Stroke 37 (6), 1379–1384. https://doi.org/10.1161/01.STR.0000221815.64093.8c.

Faria, A.V., Joel, S.E., Zhang, Y., Oishi, K., van Zjil, P.C.M., Miller, M.I., Pekar, J.J., Mori, S., 2012. Atlas-based analysis of resting-state functional connectivity: evaluation for reproducibility and multi-modal anatomy–function correlation studies. Neuroimage 61 (3), 613–621. https://doi.org/10.1016/j.neuroimage.2012.03.078.

Fisher, N.J., Tierney, M.C., Snow, G.W., Szalai, J.P., 1999. Odd/even short forms of the Boston Naming Test: preliminary geriatric norms. Clin. Neuropsychol. 13 (3), 359–364. https://doi.org/10.1076/clin.13.3.359.1742.

Flowers, H.L., Skoretz, S.A., Silver, F.L., Rochon, E., Fang, J., Flamand-Roze, C., Martino, R., 2016. Poststroke aphasia frequency, recovery, and outcomes: a systematic review and meta-analysis. Arch. Phys. Med. Rehabil. 97 (12), 2188–2201. https://doi.org/10.1016/j.apmr.2016.03.006 e8.

Foygel, D., Dell, G.S., 2000. Models of impaired lexical access in speech production. J. Mem. Lang. 43 (2), 182–216. https://doi.org/10.1006/jmla.2000.2716.

Fridriksson, J., 2010. Preservation and modulation of specific left hemisphere regions is vital for treated recovery from anomia in stroke. J. Neurosci. 30 (35), 11558–11564. https://doi.org/10.1523/JNEUROSCI.2227-10.2010.

Fridriksson, J., den Ouden, D.-B., Hillis, A.E., Hickok, G., Rorden, C., Basilakos, A., Yourganov, G., Bonilha, L., 2018. Anatomy of aphasia revisited. Brain 141 (3), 848–862. https://doi.org/10.1093/brain/awx363.

Goodglass, H., Kaplan, E., Barresi, B., 2001. BDAE-3: Boston Diagnostic Aphasia Examination, third ed. Lippincott Williams & Wilkins.

Goodglass, H., Wingfield, A. (Eds.), 1997. Anomia: Neuroanatomical and Cognitive Correlates. Academic Press.

Griffis, J.C., Nenert, R., Allendorfer, J.B., Szaflarski, J.P., 2017a. Damage to white matter bottlenecks contributes to language impairments after left hemispheric stroke. Neuroimage: Clinical 14, 552–565. https://doi.org/10.1016/j.nicl.2017.02.019.

Griffis, J.C., Nenert, R., Allendorfer, J.B., Szaflarski, J.P., 2017b. Linking left hemispheric tissue preservation to fMRI language task activation in chronic stroke patients. Cortex 96, 1–18. https://doi.org/10.1016/j.cortex.2017.08.031.

Grima, R., Franklin, S., 2017. Usefulness of investigating error profiles in diagnosis of naming impairments. Int. J. Lang. Commun. Disord 52 (2), 214–226. https://doi.org/10.1111/1460-6984.12266.

Hillis, A.E., Gold, L., Kannan, V., Cloutman, L., Kleinman, J.T., Newhart, M., Heidler-Gary, J., Davis, C., Aldrich, E., Llinas, R., Gottesman, R.F., 2008. Site of the ischemic penumbra as a predictor of potential for recovery of functions. Neurology 71 (3), 184–189. https://doi.org/10.1212/01.wnl.0000317091.17339.98.

Hillis, A.E., Kleinman, J.T., Newhart, M., Heidler-Gary, J., Gottesman, R., Barker, P.B., Aldrich, E., Llinas, R., Wityk, R., Chaudhry, P., 2006. Restoring cerebral blood flow reveals neural regions critical for naming. J. Neurosci. 26 (31), 8069–8073. https://doi.org/10.1523/JNEUROSCI.2088-06.2006.

Hillis, A.E., Beh, Y.Y., Sebastian, R., Breining, B., Tippett, D.C., Wright, A., Saxena, S., Rorden, C., Bonilha, L., Basilakos, A., Yourganov, G., Fridriksson, J., 2018. Predicting recovery in acute poststroke aphasia. Ann. Neurol. 83 (3), 612–622. https://doi.org/10.1002/ana.25184.

Hillis, A.E., Caramazza, A., 1995. The compositionality of lexical semantic representations: clues from semantic errors in object naming. Memory 3 (3–4), 333–358. https://doi.org/10.1080/09658219508253156.

Hillis, A.E., Heidler, J., 2002. Mechanisms of early aphasia recovery. Aphasiology 16 (9), 885–895. https://doi.org/10.1080/0268703.

Hillis, A.E., Kane, A., Tuffiash, E., Ulatowski, J.A., Barker, P.B., Beauchamp, N.J., Wityk, R.J., 2001a. Reperfusion of specific brain regions by raising blood pressure restores selective language functions in subacute stroke. Brain Lang. 79 (3), 495–510. https://doi.org/10.1006/brln.2001.2563.

Hillis, A.E., Wityk, R.J., Tuffiash, E., Beauchamp, N.J., Jacobs, M.A., Barker, P.B., Selnes, O.A., 2001b. Hypoperfusion of Wernicke's area predicts severity of semantic deficit in acute stroke. Ann. Neurol. 50 (5), 561–566. https://doi.org/10.1002/ana.1265.

Holland, R., Johns, S.L., Woollams, A.M., 2016. The impact of phonological versus semantic repetition training on generalisation in chronic stroke aphasia reflects differences in dorsal pathway connectivity. Neuropsychol. Rehabil. 1–20. https://doi.org/10.1080/09602011.2016.1190384.

Hosomi, A., Nagakane, Y., Yamada, K., Kuriyama, N., Mizuno, T., Nishimura, T., Nakagawa, M., 2009. Assessment of arcuate fasciculus with diffusion-tensor tractography may predict the prognosis of aphasia in patients with left middle cerebral artery infarcts. Neuroradiology 51 (9), 549–555. https://doi.org/10.1007/s00234-009-0534-7.

Humphreys, G.F., Lambon Ralph, M.A., 2015. Fusion and fission of cognitive functions in the human parietal cortex. Cerebr. Cortex 25 (10), 3547–3560. https://doi.org/10.1093/cercor/bhu198.

Indefrey, P., Levelt, W.J.M., 2004. The spatial and temporal signatures of word production components. Cognition 92 (1–2), 101–144. https://doi.org/10.1016/j.cognition.2002.06.001.

Jang, S.H., Lee, H.D., 2014. Recovery of injured arcuate fasciculus in the dominant hemisphere in a patient with an intracerebral hemorrhage. Am. J. Phys. Med. Rehabil. 93 (12), e15–e18. https://doi.org/10.1097/PHM.0000000000000202.

Jefferies, E., Baker, S.S., Doran, M., Ralph, M.A.L., 2007. Refractory effects in stroke aphasia: a consequence of poor semantic control. Neuropsychologia 45 (5), 1065–1079. https://doi.org/10.1016/j.neuropsychologia.2006.09.009.

Jefferies, E., Lambon Ralph, M.A., 2006. Semantic impairment in stroke aphasia versus semantic dementia: a case-series comparison. Brain 129 (8), 2132–2147. https://doi.org/10.1093/brain/awl153.

Jefferies, E., Patterson, K., Ralph, M.A.L., 2008. Deficits of knowledge versus executive control in semantic cognition: insights from cued naming. Neuropsychologia 46 (2), 649–658. https://doi.org/10.1016/j.neuropsychologia.2007.09.007.

Kertesz, A., 2007. Western Aphasia Battery: Revised. Pearson.

Kim, S.H., Jang, S.H., 2013. Prediction of aphasia outcome using diffusion tensor tractography for arcuate fasciculus in stroke. Am. J. Neuroradiol. 34 (4), 785–790. https://doi.org/10.3174/ajnr.A3259.

Kohn, S.E., Smith, K.L., 1994. Evolution of impaired access to the phonological lexicon. J. Neurolinguistics 8 (4), 267–288. https://doi.org/10.1016/0911-6044(94)90013-2.

Lambon Ralph, M.A., Jefferies, E., Patterson, K., Rogers, T.T., 2016. The neural and computational bases of semantic cognition. Nat. Rev. Neurosci. 18 (1), 42–55. https://doi.org/10.1038/nrn.2016.150.

Lambon Ralph, M.A., Moriarty, L., Sage, K., 2002. Anomia is simply a reflection of semantic and phonological impairments: evidence from a case-series study. Aphasiology 16 (1–2), 56–82. https://doi.org/10.1080/02687040143000448.

Laska, A.C., Hellblom, A., Murray, V., Kahan, T., Von Arbin, M., 2001. Aphasia in acute stroke and relation to outcome. J. Intern. Med. 249 (5), 413–422.

Lazar, R.M., Speizer, A.E., Festa, J.R., Krakauer, J.W., Marshall, R.S., 2008. Variability in language recovery after first-time stroke. J. Neurol. Neurosurg. Psychiatr. 79 (5), 530–534. https://doi.org/10.1136/jnnp.2007.122457.

Lazar, R.M., Boehme, A.K., 2017. Aphasia as a predictor of stroke outcome. Curr. Neurol. Neurosci. Rep. 17 (11) https://doi.org/10.1007/s11910-017-0797-z.

Le Dorze, G., Durocher, J., 1992. The effects of age, educational level, and stimulus length on naming in normal subjects. J. Speech Lang. Pathol. Audiol. 16 (1), 21–29.

Le Dorze, G., Nespoulous, J.-L., 1989. Anomia in moderate aphasia: problems in accessing the lexical representation. Brain Lang. 37 (3), 381–400. https://doi.org/10.1016/0093-934X(89)90026-6.

Levelt, W.J., Roelofs, A., Meyer, A.S., 1999. A theory of lexical access in speech production. Behavioral and Brain Sciences 22 (1), 1–38.

McKinnon, E.T., Fridriksson, J., Glenn, G.R., Jensen, J.H., Helpern, J.A., Basilakos, A., Rorden, C., Shih, A.Y., Spampinato, M.V., Bonilha, L., 2017. Structural plasticity of the ventral stream and aphasia recovery. Ann. Neurol. 82 (1), 147–151. https://doi.org/10.1002/ana.24983.

Meier, E.L., Johnson, J.P., Pan, Y., Kiran, S., 2019. The utility of lesion classification in predicting language and treatment outcomes in chronic stroke-induced aphasia. Brain Imaging and Behavior. https://doi.org/10.1007/s11682-019-00118-3.

Mirman, D., Britt, A.E., 2013. What we talk about when we talk about access deficits. Phil. Trans. Biol. Sci. 369 (1634), 20120388. https://doi.org/10.1098/rstb.2012.0388, 20120388.

Mitchum, C.C., Ritgert, B.A., Sandson, J., Berndt, R.S., 1990. The use of response analysis in confrontation naming. Aphasiology 4 (3), 261–279. https://doi.org/10.1080/02687039008249079.

Moerman, C., Corluy, R., Meersman, W., 1983. Exploring the aphasiac's naming disturbances; A new approach using the neighbourhood limited classification method. Cortex 19 (4), 529–543. https://doi.org/10.1016/S0010-9452(83)80034-3.

Mori, S., Oishi, K., Jiang, H., Jiang, L., Li, X., Akhter, K., Hua, K., Faria, A.V., Mahmood, A., Woods, R., Toga, A.W., Pike, G.B., Neto, P.R., Evans, A., Zhang, J., Huang, H., Miller, M.I., van Zijl, P., Mazziotta, J., 2008. Stereotaxic white matter atlas based on diffusion tensor imaging in an ICBM template. Neuroimage 40 (2), 570–582. https://doi.org/10.1016/j.neuroimage.2007.12.035.

Neils, J., Baris, J.M., Carter, C., Dell'aira, A.L., Nordloh, S.J., Weiler, E., Weisiger, B., 1995. Effects of age, education, and living environment on Boston Naming Test performance. J. Speech Hear. Res. 38 (5), 1143–1149. https://doi.org/10.1044/jshr.3805.1143.

Nicholas, L.E., Brookshire, R.H., Maclennan, D.L., Schumacher, J.G., Porrazzo, S.A., 1989. Revised administration and scoring procedures for the Boston Naming test and norms for non-brain-damaged adults. Aphasiology 3 (6), 569–580. https://doi.org/10.1080/02687038908249023.

Noonan, K.A., Jefferies, E., Visser, M., Lambon Ralph, M.A., 2013. Going beyond inferior prefrontal involvement in semantic control: evidence for the additional contribution of dorsal angular gyrus and posterior middle temporal cortex. J. Cognit. Neurosci. 25 (11), 1824–1850. https://doi.org/10.1162/jocn_a_00442.

Paolieri, D., Marful, A., Morales, L., Bajo, M.T., 2018. The modulating effect of education on semantic interference during healthy aging. PloS One 13 (1), e0191656. https://doi.org/10.1371/journal.pone.0191656.

Pedersen, P.M., Vinter, K., Olsen, T.S., 2003. Aphasia after stroke: type, severity and prognosis. Cerebrovasc. Dis. 17 (1), 35–43. https://doi.org/10.1159/000073896.

Plaut, D.C., 1995. Double dissociation without modularity: evidence from connectionist neuropsychology. J. Clin. Exp. Neuropsychol. 17 (2), 291–321. https://doi.org/10.1080/01688639508405124.

Plowman, E., Hentz, B., Ellis, C., 2012. Post-stroke aphasia prognosis: a review of patient-related and stroke-related factors: aphasia prognosis. J. Eval. Clin. Pract. 18 (3), 689–694. https://doi.org/10.1111/j.1365-2753.2011.01650.x.

R Core Team, 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. https://www.R-project.org/.

Rapp, B., Goldrick, M., 2000. Discreteness and interactivity in spoken word production. Psychol. Rev. 107 (3), 460–499. https://doi.org/10.1037/0033-295X.107.3.460.

Reyes, D., Hitomi, E., Simpkins, A., Lynch, J., Hsia, A., Benson, R., Nadareishvili, Z., Luby, M., Latour, L., Leigh, R., 2017. Detection of perfusion deficits using FLAIR and GRE based vessel signs. Stroke 48 (Suppl. l_1), ATP63. https://doi.org/10.1161/str.48.suppl_1.tp63. ATP63.

Rorden, C., Bonilha, L., Fridriksson, J., Bender, B., Karnath, H.-O., 2012. Age-specific CT and MRI templates for spatial normalization. Neuroimage 61 (4), 957–965. https://doi.org/10.1016/j.neuroimage.2012.03.020.

Schlaug, G., Marchina, S., Norton, A., 2009. Evidence for plasticity in white-matter tracts of patients with chronic Broca's aphasia undergoing intense intonation-based speech therapy. Ann. N. Y. Acad. Sci. 1169 (1), 385–394. https://doi.org/10.1111/j.1749-6632.2009.04587.x.

Schuell, H., Jenkins, J.J., 1961. Reduction of vocabulary in aphasia. Brain 84 (2), 243–261. https://doi.org/10.1093/brain/84.2.243.

Schwartz, M., Dell, G., Martin, N., Gahl, S., Sobel, P., 2006. A case-series test of the interactive two-step model of lexical access: evidence from picture naming. J. Mem. Lang. 54 (2), 228–264. https://doi.org/10.1016/j.jml.2005.10.001.

Schwartz, M.F., Brecher, A., 2000. A model-driven analysis of severity, response characteristics, and partial recovery in aphasics' picture naming. Brain Lang. 73 (1), 62–91. https://doi.org/10.1006/brln.2000.2310.

Seghier, M.L., 2013. The angular gyrus multiple functions and multiple subdivisions. Neuroscientist 19 (1), 43–61.

Sorensen, A.G., Buonanno, F.S., Gonzalez, R.G., Schwamm, L.H., Lev, M.H., Huang-Hellinger, F.R., Reese, T.G., Weisskoff, R.M., Davis, T.L., Suwanwela, N., Can, U., Moreira, J.A., Copen, W.A., Look, R.B., Finklestein, S.P., Rosen, B.R., Koroshetz, W. J., 1996. Hyperacute stroke: evaluation with combined multisection diffusion-weighted and hemodynamically weighted echo-planar MR imaging. Radiology 199 (2), 391–401. https://doi.org/10.1148/radiology.199.2.8668784.

Staff, R.T., Murray, A.D., Deary, I.J., Whalley, L.J., 2004. What provides cerebral reserve? Brain 127 (5), 1191–1199. https://doi.org/10.1093/brain/awh144.

Tallberg, I.M., 2005. The Boston naming test in Swedish: normative data. Brain Lang. 94 (1), 19–31. https://doi.org/10.1016/j.bandl.2004.11.004.

van Hees, S., McMahon, K., Angwin, A., de Zubicaray, G., Read, S., Copland, D.A., 2014. Changes in white matter connectivity following therapy for anomia post stroke. Neurorehabilitation Neural Repair 28 (4), 325–334.

Wade, D.T., Hewer, R.L., David, R.M., Enderby, P.M., 1986. Aphasia after stroke: natural history and associated deficits. J. Neurol. Neurosurg. Psychiatry 49 (1), 11–16.

Watila, M.M., Balarabe, S.A., 2015. Factors predicting post-stroke aphasia recovery. J. Neurol. Sci. 352 (1–2), 12–18. https://doi.org/10.1016/j.jns.2015.03.020.

Whitney, C., Kirk, M., O'Sullivan, J., Lambon Ralph, M.A., Jefferies, E., 2012. Executive semantic processing Is underpinned by a large-scale neural network: revealing the contribution of left prefrontal, posterior temporal, and parietal cortex to controlled retrieval and selection using TMS. J. Cognit. Neurosci. 24 (1), 133–147. https://doi.org/10.1162/jocn_a_00123.

Xu, Y., He, Y., Bi, Y., 2017. A tri-network model of human semantic processing. Front. Psychol. 8 https://doi.org/10.3389/fpsyg.2017.01538.

Xu, Y., Lin, Q., Han, Z., He, Y., Bi, Y., 2016. Intrinsic functional network architecture of human semantic processing: modules and hubs. Neuroimage 132, 542–555. https://doi.org/10.1016/j.neuroimage.2016.03.004.