Investigating language processing in stroke-affected brains

A pilot study on the key role of stimulus design to capture interface domains

Alice Todesco[1] Domenica Romagno[2]
Michelangelo Mancuso[2] and Claudia Marzi[1]
[1] Italian National Research Council | [2] University of Pisa

Stroke represents one of the leading causes of acquired cognitive and communicative impairments, limiting the ability to use language appropriately and profoundly impacting quality of life and social participation. Post-stroke aphasia offers a crucial pathway for the understanding of the neuropsychological correlates of language processing and, despite advances in acute stroke management, linguistic deficits continue to represent a significant challenge for rehabilitation and long-term care. Additionally, one of the major challenges in investigating post-stroke language deficits lies in the marked variability across patients. In this pilot study, we investigated whether stimulus design may enhance the detection of linguistic processing deficits in stroke-affected individuals, targeting lexical class distinctions and morphosyntax/lexical semantics interface-level processing. Results from twelve Italian-speaking patients with unilateral stroke (6 left-hemisphere, 6 right-hemisphere) revealed differential effects of lesion lateralisation and lexical class on task performance. Task accuracy was shaped by individual and lesion-related factors, revealing task-specific dissociations between left- and right-hemisphere patients. These findings underscore the value of using targeted linguistic stimuli to detect subtle deficits that may go undetected by standard screening tools, contributing to a more comprehensive understanding of post-stroke aphasia.

Keywords: lexical-semantic processing, morphosyntax/lexical semantics interface processing, linguistic assessment, aphasia, post-stroke

https://doi.org/10.1075/ml.25010.tod | Published online: 7 May 2026
The Mental Lexicon ISSN 1871-1340 | E-ISSN 1871-1375

1. Introduction

Stroke represents one of the leading causes of acquired cognitive and communicative impairments worldwide, with language deficits being a common and often persistent consequence. According to the World Stroke Organization, over 12 million people suffer from stroke annually (Feigin et al., 2022); of these, it is estimated that between 21% and 40% will sustain permanent damage to linguistic functions (Cichon et al., 2021; Engelter et al., 2006; Flowers et al., 2016). However, language impairments resulting from stroke exhibit considerable heterogeneity, reflecting the complex neuroanatomical and functional organisation of language in the brain. Classical models, such as the Wernicke-Lichtheim-Geschwind model (Geschwind, 1965), distinguish between expressive and receptive deficits based on lesion location. Broca's aphasia, typically associated with damage to the left inferior frontal gyrus, manifests as non-fluent speech with preserved comprehension. In contrast, Wernicke's aphasia, resulting from lesions in the left posterior superior temporal gyrus, is characterised by fluent speech and impaired comprehension (Dronkers, Wilkins, Van Valin Jr, Redfern, & Jaeger, 2004). Contemporary dual-stream models of language processing (Hickok & Poeppel, 2007) offer a more nuanced view, proposing two interacting pathways: a dorsal stream supporting sensorimotor integration and articulation, and a ventral stream responsible for semantic and syntactic processing. Stroke-induced damage may disrupt one or both streams, leading to overlapping and complex patterns of impairment that challenge straightforward classification.

Neuroimaging studies have significantly advanced our understanding of the brain regions involved in language processing and their disruption following stroke. Lesions to the perisylvian cortex in the left hemisphere are most frequently associated with aphasia, including damage to Broca's and Wernicke's areas, the arcuate fasciculus, and surrounding regions (Kümmerer et al., 2013). Beyond these classical areas, recent evidence highlights the role of subcortical structures (e.g., basal ganglia and thalamus) and the importance of white matter connectivity in supporting language functions (Fridriksson et al., 2018). Furthermore, factors such as lesion size, location, and individual differences in cognitive reserve influence both the nature of language impairment and the likelihood of recovery (Marantz, 2013). Besides, language deficits rarely occur in isolation and often coexist with other cognitive impairments, including deficits in working memory, attention, and executive function (Brownsett et al., 2014).

Therefore, the assessment and treatment of post-stroke aphasia represent a crucial issue in neuro-rehabilitation, while also offering a significant opportunity for studying the functional organisation of the language network in the human brain. Despite advances in acute stroke management, the study of post-stroke aphasia poses significant challenges for the investigation of the neuropsychological correlates of language processing, as well as for rehabilitation and long-term care, since current clinical practice tends to rely on a limited number of standardised test batteries that primarily evaluate broad linguistic functions such as comprehension and production, offering limited insight into the cognitive mechanisms underlying language processing. However, to disentangle the different dimensions involved in linguistic deficits represents a pivotal issue for the study, assessment, and rehabilitation of language disruptions. In fact, effective investigation of the neural correlates of behaviourally manifested deficits demands the utilisation of batteries of stimuli that can specifically target different linguistic phenomena (Battista et al., 2017). Consequently, precision in the design of tasks and stimuli is essential to advance understanding in this field, highlighting the need for linguistic tests that capture multiple variables while enabling a clear separation of the underlying dimensions.

1.1 Diagnostic batteries

Currently, several batteries of linguistic stimuli are used for the study, assessment and rehabilitation of Italian-speaking brain-damaged patients presenting with aphasic manifestations. The first group of batteries, which can be referred to as *diagnostic batteries*, aims to provide clinical evaluation tools that identify specific clinical syndromes or variants of deficits. The *Battery for the Analysis of Aphasic Deficits* (*Batteria per l'Analisi dei Deficit Afasici*, BADA (Miceli, Laudanna, & Capasso, 2006)) includes sub-tests that assess phonological processing, reading, writing, and broad comprehension and production abilities. The *Neuropsychological Aphasia Examination* (*Esame Neuropsicologico per l'Afasia*, ENPA (Capasso & Miceli, 2001)) – designed to assess aphasia in Italian speakers – evaluates different aspects of language, including spontaneous speech, auditory comprehension, repetition, naming, reading, and writing. It also includes tests for non-verbal skills such as memory and visuospatial abilities, to allow for the assessment of potential correlations between aphasia and cognitive impairments involving other domains. The *Aachen Aphasia Test* (AAT), also known as the *Aachener Aphasietest* (Huber, Poeck, Weniger, & Willmes, 1983), is a battery originally developed for German patients and subsequently adapted for Italian (Luzzatti et al., 1994). IT-AAT allows for an assessment of spontaneous speech, auditory comprehension, written language, repetition, and naming, and it has been recently revised using an enlarged participant sample and updated statistical methods (Luzzatti, De Bleser, Scola, Frustaci, & Willmes, 2023). The *Screening for Aphasia in Neurodegeneration* (SAND), was developed for English-speaking patients (Garrard & Ahmed, 2012) and then adapted into Italian (Catricalà et al., 2017). The battery consists of nine subtests designed to assess naming, comprehension, repetition, reading, semantic association, writing and picture description. Lastly, a prominent example of the use of diagnostic batteries can be found in the assessment and classification of different variants of Primary Progressive Aphasia (PPA) (Gorno-Tempini et al., 2011), (Battista et al., 2017). In fact, numerous batteries have been designed that specifically aim to provide tools for the classification of the different variants of PPA (namely, non-fluent/agrammatic, semantic and logopenic).

1.2 Domain-specific batteries

A second category of batteries, here defined as *domain-specific batteries*, focuses on analysing specific dimensions of language. The most widespread of these batteries is the *Boston Naming Test* (BNT (Kaplan, Goodglass, & Weintraub, 1983)), a widely used neuropsychological assessment tool designed to evaluate naming abilities. It consists of 60 line-drawings of objects, arranged for increasing order of difficulty, that the participant is asked to name. The *CaGi Test* (Catricalà et al., 2013), specifically designed to assess semantic memory disorders, picture naming, word-picture matching, picture sorting, and feature generation. Concerning the study of pragmatic abilities, the *Test for the Assessment of Pragmatic Abilities and Cognitive Substrates* (APACS) (Arcara & Bambini, 2016)) has been developed to monitor language use in context, the recognition of non-literal meanings, and adherence to social conventions in conversations. The Italian version of the updated Bishop's (Bishop, 2009) Test for Reception of Grammar (TROG-2) is used to assess the comprehension of grammatical structures, originally developed for children but also adapted for adults.

While each of the aforementioned batteries offers specific advantages in the study of linguistic deficits, none of these tools specifically take into account the relationship between cognitive and linguistic categories, thus leaving unanswered the question of which broader cognitive properties contribute to an observed deficit pattern. In other words, existing batteries fail to fully account for the role of *interface semantic properties*.

1.3 Interface semantic properties

It has been shown that semantic regularities in the composition of classes of words with the same morphosyntactic behaviour do not depend directly on the denotational components of word meaning but rather are determined by shared interface semantic properties, independently of language-specific features (Bertinetto, 1986; Croft, 2010; Levin & Hovav, 1994; Van Valin Jr, 1990, 2005).

These properties correspond to those semantic components that surface in morphosyntactic structure, and they include: *telicity*, which distinguishes events that entail a specified endpoint from those without temporal and spatial delimitation and guides the distribution of temporal adverbials such as *for X time* (compatible with atelic constructions) and *in X time* (occurring only in telic constructions); *agentivity*, which refers to the degree to which a referent is capable of initiating or controlling an action or event and determines the possibility (or lack thereof) of passive constructions; *dynamicity*, which indicates whether a given situation involves change or movement over time and determines the possibility (or lack thereof) of progressive constructions; and *individuation*, which refers to the degree to which a noun or entity is perceived as a separable unit and represents the semantic basis for the distinction between countable and uncountable nouns (Bertinetto, 1986; Levin & Hovav, 1994; Tenny, 1994; Van Valin & LaPolla, 1997; Vendler, 1967). Contemporary neurolinguistic studies have shown increasing interest in the interplay between language and cognition, and in this context interface semantics has proven to be a privileged access point, allowing the observation of this interaction through empirically testable linguistic phenomena (Peelen, Romagno, & Caramazza, 2012; Romagno, 2012, 2019; Romagno, Rota, Ricciardi, & Pietrini, 2012; Shetreet, Friedmann, & Hadar, 2010).

Therefore, the specific targeting of interface phenomena in the construction of batteries of linguistic stimuli may offer a unique opportunity to gain a deeper understanding of the interaction between linguistic categories and the broader cognitive principles that govern their organisation. However, experimental batteries standardised for Italian do not support a clear analysis of such properties, as they fail to disentangle semantic and morphosyntactic features effectively.

1.4 Perspectives on word-classes

According to *prototype theory*, cognitive categories are organised around central, prototypical members, with other items varying in representativeness based on their proximity to the prototype (Rosch, 1973). Prototype effects appear in various linguistic phenomena, including the way in which linguistic categories are structured (Lakoff, 2008). A key example of this is the association between word classes such as *nouns* and *verbs*, and the cognitive categories they prototypically correspond to across languages, namely *entities* and *events*. The issue of categorisation is a crucial one in the field of linguistics, and one that has been extensively discussed in literature. From a theoretical standpoint, the central debate concerns whether word classes correspond to universal cognitive categories (Baker, 2003), (Panagiotidis, 2015) or are shaped by language-specific features (Croft, 2001; Haspelmath, 2012). Another major discussion concerns whether word-class distinctions arise from semantic features inherent in lexical roots (Lakoff, 2008; Panagiotidis, 2005), or emerge through syntactic structures (Barner & Bale, 2002, 2005; Chomsky, 1970). Experimental research into word-class distinctions, particularly the noun-verb dichotomy, has provided valuable insights into language processing, while also revealing the multifactorial nature of the issue (Romagno, 2016, 2024). Studies have explored various factors, including semantic features, neural correlates, neural activation patterns, dissociation patterns, and word-class deficits in different pathologies (Handjaras et al., 2016; Law, Yum, & Cheung, 2017; Papagno, Capasso, & Miceli, 2009; Peelen et al., 2012; Shapiro & Caramazza, 2003; Xia & Peng, 2022; Zhang et al., 2018, among many others). A consistent observation across languages, cohorts, materials, and methods resides in the observation that verbs consistently demand greater processing resources than nouns, though the precise factors contributing to this phenomenon remain unclear. Furthermore, findings suggest that the representation of word-classes in the brain is not a simple, localised phenomenon but rather involves a complex network of brain regions. In particular, the left posterior middle temporal gyrus appears to be a crucial hub for verb processing (Peelen et al., 2012; Romagno et al., 2012), but other regions like the left inferior parietal lobe and left mid-superior temporal cortex also seem to contribute to processing semantic features relevant to both nouns and verbs (Williams, Reddigari, & Pylkkänen, 2017). These findings point towards a multi-faceted and distributed representation of word-class information in the brain, where semantic features and grammatical properties, as well as task design and individual differences in brain organisation, all interact to shape how different types of words are processed in the brain. In light of this complexity, the precise selection of linguistic stimuli that allow for a disentanglement of the semantic and morphosyntactic components of the world-class distinction is essential for advancing our understanding of how linguistic categories interact with cognitive structures. However, existing standardised batteries for Italian often fall short in this regard, as they do not adequately differentiate between overlapping semantic and morphosyntactic features. This further highlights the importance of developing experimental tools capable of isolating distinct linguistic dimensions.

1.5 Objective of the study

The present study aimed to assess the clinical efficacy of the BISMoS battery, a battery of linguistic stimuli specifically designed to investigate the neurocognitive correlates of word-class distinctions and interface-level semantic processing (Romagno, Busso, & Elce, 2020). These dimensions are often underexplored in standard aphasia assessments, despite their central role in language organisation.

A representative subset of the battery was administered to twelve cerebrovascular patients with temporal-parietal brain damage resulting from an ischemic stroke involving the first segment (M1) of the Middle Cerebral Artery (MCA).

By targeting linguistic structures that tap into both lexical and interface-level phenomena, the study seeks to contribute to a finer-grained understanding of post-stroke language impairments and to evaluate the potential of linguistically grounded tools in complementing conventional diagnostic procedures. In particular, we selected stimuli that systematically manipulate morphosyntactic and semantic features, allowing for the identification of selective impairments that might remain undetected in broader clinical tests. In our results, we show that the interplay between lesion lateralisation and task type systematically affects performance across linguistic conditions, highlighting how different components of language processing rely on distinct neural resources.

2. Materials and methods

2.1 Participants

Following clinical evidence and neuro-anatomical knowledge, the M1 segment of the Middle Cerebral Artery (MCA-M1) was identified as the primary Region of Interest (ROI) for this study. The MCA is the largest cerebral artery and the most commonly affected vessel in cerebrovascular episodes (Kandel, 2021; Mendelow, 2015); focusing on this ROI, therefore, maximised the likelihood of recruiting a sufficient number of subjects with relevant damage. Furthermore, MCA stroke has been widely shown to represent a predictive factor in the development of language deficits (Kandel, 2021; Mendelow, 2015; Mtui, Gruener, & Dockery, 2020). In fact, the importance of the vessel in the genesis of aphasia is so well-established that, in a recent article, it has been dubbed *the artery of aphasia* (Briggs et al., 2019). In particular, the M1 (or sphenoidal) segment has been shown to play an important role in supplying blood to language-critical areas (Fujii et al., 2011; Kakinuma et al., 2022). In fact, due to its location, a vascular event involving the M1 tract of the MCA would typically result in lesions involving the perisylvian area and underlying structures – a network extensively shown to be crucial for linguistic processing (Catani, Jones, & Ffytche, 2005; Halai, Woollams, & Ralph, 2017; Hickok & Poeppel, 2004; Williams et al., 2017) and specifically suggested to function as a hub for interface semantic processing (Romagno et al., 2012).

Only patients with lesions resulting from ischemic stroke (IS) were included. In fact, IS is statistically more prevalent (Andersen, Olsen, Dehlendorff, & Kammersgaard, 2009), which enhanced the possibility of recruiting an adequate number of subjects. Additionally, hemorrhagic stroke (HS) tends to cause more severe damage (Andersen et al., 2009), which could have raised potential practical and ethical concerns. Focusing on a single lesion etiology also helped control for potential confounding factors stemming from variability in underlying pathology.

Because patients did not undergo a uniform diagnostic protocol at the time of stroke, lesion characterisation could only be qualitative. Lesions were therefore classified as cortical, subcortical, or cortico-subcortical based on radiological reports, whereas precise lesion-size estimates were not possible. Thus, lesions involving the M1 segment of the MCA were classified according to their predominant extent. Cortical lesions mainly affected the perisylvian cortex supplied by the superficial MCA branches (e.g., inferior frontal gyrus, temporal cortex, inferior parietal lobule), with only minimal subcortical involvement. Subcortical lesions were confined to the deep MCA territory, typically including the basal ganglia, internal capsule, and adjacent white-matter pathways while sparing the cortex. Cortico-subcortical lesions showed combined involvement of cortical perisylvian regions and deeper white-matter tracts, reflecting more extensive damage to both the superficial and deep branches originating from the M1 segment.

No *a priori* limitations were imposed on gender and age of participants, given that no significant differences related to these variables had emerged during the standardisation process of the BISMoS tasks. Education, on the other hand, represented a strong predictor of accuracy in the norm group. It is also important to note that cognitive reserve – commonly indexed by years of education (henceforth *YOE*) – is widely recognised as a factor influencing neural plasticity and thus playing a key role in post-stroke recovery (Stern, 2002, 2012). Thus, no specific selection criteria were applied regarding YOE and time since stroke onset (henceforth *TSO*), which are reported as potential modulators of task performance (Kiran, 2012; Stern, 2012).

Participant recruitment took place during routine clinic activities, when patients attended their scheduled appointments. Patients' participation was entirely voluntary. Inclusion criteria required participants to have a single, unilateral brain lesion (either left or right hemisphere), and the cognitive capacity to understand and follow instructions. Exclusion criteria included history of previous neurological or psychiatric disorders, severe aphasia preventing task comprehension, significant uncorrected sensory deficits, or any other medical condition likely to interfere with task performance. All participants were monolingual native speakers of Italian, with no reported history of multilingualism or extensive second-language exposure, to minimise potential confounds in language processing.

Tasks were administered in short blocks, including rest periods, depending on the patient's level of functioning, to ensure the patient's comfort. Special care was taken to explain that the subject may terminate participation in the study at any time, for any reason, without any consequences. Subjects were told that their participation may help scientists learn more about the normal functioning of the brain and the neurocognitive mechanisms for language processing, and how such functioning and mechanisms break down following brain damage. All the collected data were kept in locked cabinets and the only identifying information on these data was a unique alphanumeric code assigned to each participant. The final cohort consisted of twelve subjects, six for each group, as detailed in Table 1.

Table 1. Demographic and clinical information for all subjects. The table reports each subject's group designation (L = left hemisphere lesion; R = right hemisphere lesion), age, years of education (YOE), time since stroke onset (TSO), and lesion type (C = cortical; S = subcortical; CS = corticosubcortical)

Subject	Group	Age	YOE	TSO	Lesion Type
1	L	53	8	1y 9m	S
2	L	66	8	0y 4m	C
3	L	59	8	3y 1m	CS
4	L	54	8	0y 8m	CS
5	L	44	18	0y 2m	C
6	L	62	8	8y	CS
1	R	54	8	6y	CS
2	R	49	13	2y 4m	CS
3	R	57	5	3y	CS
4	R	56	8	0y 3m	S
5	R	74	18	0y 9m	CS
6	R	68	18	3y	S

2.2 Materials and methods

Each administration session began with a general assessment of the subject's overall cognitive abilities, achieved through the use of the Italian version of the Montreal Cognitive Assessment (MoCA). No cut-off score was applied, as some degree of impairment was expected due to participants' neurological condition; MoCA scores were thus considered for descriptive and correlation purposes only. Following the MoCA, patients completed a subset of the BISMoS battery. Given that the full BISMoS administration requires approximately 90 minutes, a reduced version lasting a maximum of 45 minutes was used to accommodate the cognitive and attentional constraints of brain-damaged individuals (see Appendix A for a selection of stimuli). The final subset preserved the balance and core assessment dimensions of the full battery while reducing the number of items in each task by approximately half.

All sessions were conducted in a quiet, well-lit clinic room and scheduled outside regular operating hours to minimise environmental noise. Patients were seated at a table with only the necessary materials (MoCA and BISMoS forms, pen, water), and caregivers waited outside unless clinically required. Each session lasted approximately 60 minutes and included two brief breaks. Tasks were administered orally by the examiner following the official instructions of the BISMoS battery and the Italian adaptation of the MoCA test. Although printed materials were available, all responses were collected orally and transcribed by the examiner. This uniform response modality was adopted to avoid inconsistencies and to accommodate patients with possible left-hemisphere lesions associated with contralateral hemiparesis, which may prevent written responses.

The full BISMoS was standardised using a between-subjects design on a sample of 98 healthy adults, balanced for age, sex, and YOE (Romagno et al., 2020, 41–62). All lexical materials were selected from among the 1000 most frequent Italian words, as selected from the ItTenTen corpus (Jakubíček, Kilgarriff, Kovář, Rychlỳ, & Suchomel, 2013). Our selected subset included 197 linguistic stimuli (out of the 426 in the full battery) organised in four categories: *morphosyntax/semantics interface*, *non semantically determined morphology and syntax*, *production*, and *comprehension*. The first category targets semantic and morphosyntactic interface phenomena, examining the effects of telicity, agentivity, dynamicity, and individuation through sentence completion tasks and acceptability judgments. The second category includes two subdomains, *verbal morphology* and *syntax*, both assessed via acceptability judgments, and focused on structures not determined by semantic constraints. Lastly, the *production* and *comprehension* sections assess general linguistic abilities and help rule out that observed deficits stem from basic lexcial or semantic impairments. Stimuli in these two sections involve sentence completion (production), semantic similarity judgments, and definition matching (comprehension).

Given the known association between different types of linguistic processing and distinct neural networks, examining contrasts across these categories provides a basis for identifying neurocognitive dissociations. For instance, comparing performance in the *comprehension* and *interface* categories allows for distinguishing between referential semantics — typically associated with anterior temporal regions — and interface-level semantics, linked to more posterior temporo-parietal areas (Dronkers et al., 2004; Romagno, 2012). Similarly, contrasting performance in the *non-semantic morphosyntax* and *interface* categories may reveal dissociations between morphosyntactic processes associated with the inferior frontal gyrus versus those involving temporo-parietal regions (Matchin & Hickok, 2020).

Data analysis focused on inter- and intra-group comparisons across categories, with the aim of identifying selective dissociation patterns in patients with left- versus right-hemisphere temporo-parietal damage.

3. Results

All statistical analyses were performed with R (R Core-Team, 2024). Response accuracy was scored at the single item level as correct (1) or incorrect (0), reflecting either acceptability judgments, guided phrase completions, or full phrase completions. To account for inter-individual variability and the nested structure of our data, a mixed-effects modelling approach was adopted. Trial-level accuracy was also modelled using logistic regression (detailed summaries of all models are reported in Appendix B), although these analyses explained limited variance. Such limitation likely reflects the reduced sample size and the sparse distribution of outcomes in this pilot context.

3.1 Preliminary analyses

We first examined whether overall accuracy differed between patients with left vs. right hemisphere lesions. Non-parametric and parametric tests yielded converging results, indicating no significant difference between the two hemisphere groups. Specifically, a *Kruskal-Wallis* rank sum test revealed no significant effect of hemisphere group on response accuracy ($\chi^2 = 0.32$, $p = 0.57$), and a *Welch* two-sample t-test also failed to detect a significant difference ($t = 0.56$, $p = 0.57$).

To further assess whether lesion hemisphere had a significant effect on accuracy, a generalised linear mixed-effects model (*glmer*) was fitted with response accuracy as the dependent variable, lesion group as fixed effect, and patients as a random intercept. The model confirmed the absence of a significant effect of lesion hemisphere ($p = 0.926$), suggesting no group-level differences in overall accuracy, even when accounting for inter-patient variability (see Table B1 in Appendix B).

We then explored the effect of individual variables that may modulate performance across both groups. For all patients, both years of education (YOE) and time since stroke onset (TSO) significantly affected response accuracy to all stimuli, as confirmed by Kruskal-Wallis rank sum tests for both variables (YOE: $\chi^2(3) = 35.5$, $p < 0.0001$; TSO: $\chi^2(10) = 108.28$, $p < 2.2\text{e-}16$). Post-hoc pairwise comparisons with Dunn's tests (*Holm* adjustment) confirmed that, for years of education (YOE), there were no significant differences between the medium and high education levels (13 vs. 18 years of education, $p = 1.00$), whereas both showed significantly better accuracy when compared to the low education groups (5 and 8 years, all $p < 0.05$). Regarding time since onset (TSO), pairwise comparisons (with *Bonferroni* correction) suggested a non-linear trajectory in recovery patterns. Accordingly, to explore potential non-linear effects of TSO on task accuracy, we divided the sample into four TSO categories: 0.2–0.4 years, 0.8–0.9 years, 1.9–3 years, and 6–8 years. A Kruskal-Wallis test revealed a significant difference in accuracy across groups ($\chi^2 = 14.36$, $p = 0.002$), prompting post-hoc comparisons using Dunn's test. Results indicated that the group with TSO 0.8–0.9 years performed significantly better than both the 1.9–3 years group ($Z = 3.17$, $p = 0.009$) and the 6–8 years group ($Z = 3.00$, $p = 0.013$). Overall, these findings suggest a non-monotonic relationship between TSO and task accuracy, with a potential peak in performance occurring within the first year post-stroke — particularly around 8–11 months — followed by a decrease or plateau over time.

The YOE pattern suggests that higher educational attainment may be associated with better performance. These results align with previous research indicating that individuals with a higher level of education may benefit from greater cognitive reserve, which can mitigate the impact of brain damage and age-related cognitive decline, thus supporting performance in cognitively demanding tasks (Stern, 2002; Tucker & Stern, 2014; Valenzuela & Sachdev, 2006). Regarding TSO, the observed pattern does not follow a simple monotonic improvement with time. Instead, it reflects the dynamic and non-linear nature of post-stroke neuroplasticity. Neuroimaging studies show, in fact, that the most robust phase of neuroplastic reorganisation occurs within the first 6–12 months after stroke, followed by a relative stabilisation or plateau (Breitenstein et al., 2017; Corbetta et al., 2015; Hamilton, Chrysikou, & Coslett, 2011; Saur et al., 2006). This evidence on longitudinal studies is compatible with the peak in performance observed around 0.8–0.9 years post-stroke in our sample. A subsequent decline may additionally be shaped by reduced rehabilitation intensity, environmental and psychosocial factors, or the cumulative impact of vascular and cognitive comorbidities, all of which are known to modulate long-term outcomes (Plowman, Hentz, & Ellis, 2012). In addition, the variability observed across patients further underscores the heterogeneity of individual recovery trajectories. This effect is likely amplified by our relatively small sample, where factors such as age, years of education, lesion site, and comorbidities may interact with time since onset and blur linear relationships (see Section 4.4).

3.2 Predictive modelling

Building on these preliminary findings, a logistic regression model was applied to assess the predictive contribution of years of education (YOE), time since lesion onset (TSO), and type of lesion (LESION) – cortical, subcortical, cortico-subcortical. All three variables were found to be significant predictors of response accuracy: YOE ($\beta = 0.26$, $p < 0.001$), TSO ($\beta = -0.18$, $p = 0.010$), and LESION (with cortico-subcortical lesions showing the strongest effect; $\beta = 1.58$, $p = 0.002$). The patient group (left vs. right hemisphere lesion) also contributed significantly ($\beta = -0.61$, $p = 0.025$), with lower accuracy observed in the right hemisphere group. Despite the statistical significance of individual predictors, the model showed modest overall explanatory power, with a modest pseudo-R-squared values (McFadden's $R^2 = 0.078$, corresponding to approximately 8% of the explained deviance), highlighting the complexity and multifactorial nature of accuracy outcomes in our clinical sample (see Table B2 in Appendix B).

To further explore the influence of recovery time, we fitted an additional logistic model (a generalised additive model, or *GAM*) including a binarized version of time since onset (with TSO < 1 vs. > 1 year) (Brady, 2021; Wilson et al., 2023). The model included lesion type, lesion group, years of education (YOE), and their interactions with the binary TSO variable (see Table B3 in Appendix B). This exploratory analysis yielded patterns broadly consistent with some previously observed effects. However, given the limited sample size, the interaction of terms should be interpreted cautiously. Specifically, YOE was positively associated with accuracy in patients assessed within the first year post-onset (TSO < 1 year, ($p = 0.0013$). This effect was attenuated in patients with a longer recovery time (TSO > 1 year), as reflected by a non-significant marginal negative interaction between YOE and TSO > 1 ($\beta = -0.05$, $p = 0.63$), suggesting that the influence of educational background may be more prominent in the early stages of recovery – possibly due to a greater reliance on cognitive reserve mechanisms in the acute and subacute phases (Bertoni, Bruni, Saviola, De Tanti, & Costantino, 2024). In addition, patients with cortico-subcortical lesions showed significantly better performance than those with cortical lesions ($p = 0.007$). A main effect of lesion group indicated lower accuracy in the right-hemisphere group ($p < 0.0001$), and a significant interaction between lesion type and lesion group further highlighted the importance of lesion location in modulating performance profiles ($p < 0.0001$). Overall, the model explained a modest proportion of variance (adjusted $R^2 = 3.8$, deviance explained = 14.2%), consistent with the expected complexity and inter-subject variability of post-stroke recovery trajectories.

In both models, years of education, time since onset, and lesion type emerged as significant predictors of response accuracy. However, the limited variance explained suggests that additional factors likely contribute to performance outcomes. To capture more fine-grained effects and better understand the cognitive mechanisms underlying performance across different lesion profiles, we next examined the influence of stimulus type and lexical category.

In what follows, we examined response accuracy in relation to specific task demands and lesion profiles. We focused, in particular, on dissociation between production, comprehension, and other linguistic tasks, considering the lesion laterality (left vs. right hemisphere) and type (cortical, subcortical, cortico-subcortical) as key grouping variables.

3.3 Task analyses

Following the preliminary analyses, we examined patients' performance across different linguistic tasks to better understand how lesion laterality and type influence accuracy.

We conducted post-hoc *Dunn* test within each group. Patients with right-hemisphere lesions exhibited significantly lower accuracy in comprehension than in production tasks ($Z = -7.22$, $p < 0.0001$), consistent with the hypothesised role of the right hemisphere in inferential and pragmatic language processing (Gajardo-Vidal et al., 2018).

In contrast, patients with left-hemisphere lesions showed a significant dissociation between interface and syntactic tasks. Specifically, performance in the syntactic tasks was significantly better than in the interface tasks ($Z = -5.02$, $p < 0.0001$). This pattern is consistent with the role of the left hemisphere in processing both morphosyntax and interface-level linguistic structures, but it suggests that the left hemisphere's involvement in semantically driven syntax might be more disrupted in this group (Romagno et al., 2012). These complementary patterns underscore a functional lateralisation in post-stroke linguistic performance, motivating a more fine-grained analysis of lesion type and its interaction with task performance.

To further explore these effects, we analysed response accuracy across lesion types (cortical, subcortical, cortico-subcortical) within each lesion group (left vs right hemisphere). Figure 1 illustrates the mean accuracy and bootstrapped confidence intervals as a function of lesion type and hemisphere (see Table B4 in Appendix B for the corresponding fitted GAM).

While this summary plot captures overall performance differences across lesion types and hemispheric groups, it does not account for task-specific variation. Based on the previously observed group-by-task interactions, we next exam-

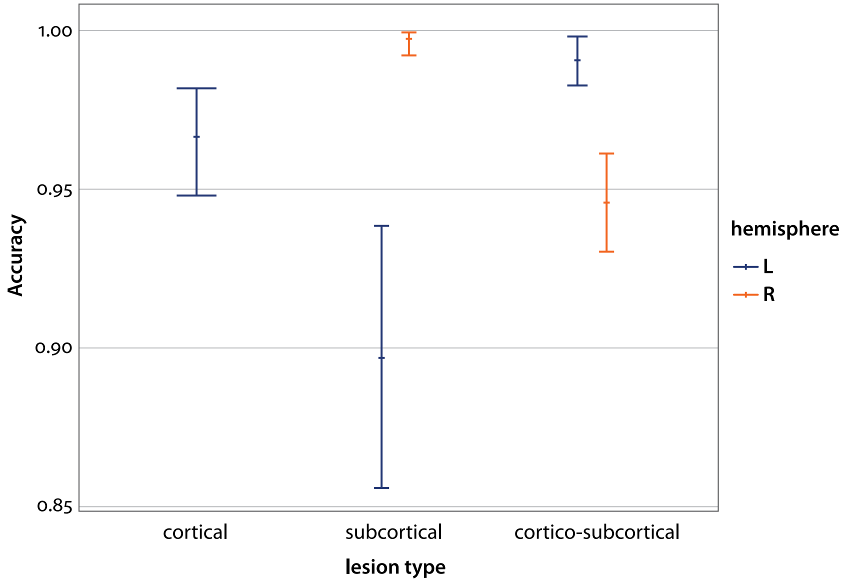

Figure 1. Response accuracy across lesion types, by hemisphere group. Dots represent mean accuracy scores for each lesion type (cortical, subcortical, cortico-subcortical) within the left (L, in blue) and right (R, in coral) hemisphere lesion groups. Error bars indicate 95% confidence intervals

ined task performance separately within each lesion group, focusing on the specific dissociation that had emerged as significant.

Within the right-hemisphere lesion group, a significant dissociation between production and comprehension tasks emerged across different lesion types. Patients with mixed cortical-subcortical lesions showed a robust difference ($Z = 4.17$, $p < 0.001$), while those with subcortical-only lesions also exhibited a significant, though more moderate, dissociation ($Z = 2.94$, $p = 0.03$). Notably, no patients in this group presented with exclusively cortical lesions. These findings suggest that disruption of subcortical structures alone is sufficient to impair task performance, while combined cortical and subcortical damage may exert a compounding effect.

A parallel analysis was conducted for the left-hemisphere lesion group. Initial results revealed a significant difference between interface and morphosyntactic tasks, pointing to a possible sensitivity of the left hemisphere to these aspects of language processing. However, when breaking down performance by lesion type (cortical, subcortical, cortico-subcortical), no significant differences emerged between the two tasks within any of the three lesion subgroups. This finding suggests that the overall dissociation between interface and morphosyntactic tasks observed at the group level is likely driven by the interaction between lesion type and task performance. Specifically, it appears that the type of lesion in combination with the lateralisation play a key role in determining whether these two types of tasks show a significant difference, and that each type of lesion may impact performance on these tasks differently.

To further test the specificity of the observed task effects within each lesion group, we compared performance between left- and right-hemisphere patients on the tasks that showed significant intra-group dissociations. Comprehension tasks revealed a significant difference between the two groups ($W = 94536$, $p < 0.001$), with right-hemisphere patients performing significantly worse ($M = 0.94$) than left-hemisphere patients ($M = 0.96$). Interface tasks showed a trend-level difference ($W = 45{,}936$, $p = 0.058$), with the pattern reversed — right-lesion patients performed slightly better ($M = 0.97$) than left-lesion patients ($M = 0.96$) — suggesting a possible lateralised effect for this task type. By contrast, group differences in production and morphosyntactic tasks were not statistically significant. Nonetheless, right-lesion patients showed numerically higher accuracy in production, whereas left-lesion patients performed slightly better in morphosyntactic tasks, as shown in Figure 2.

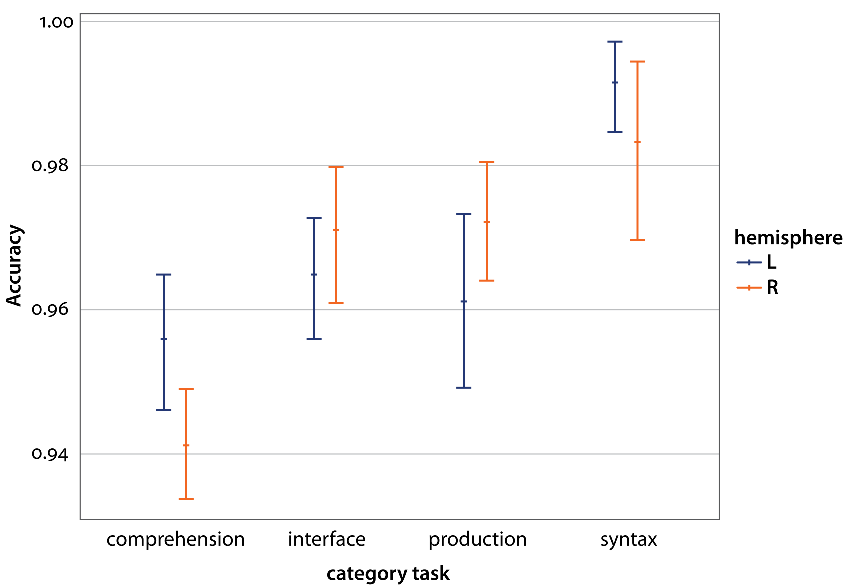

Figure 2. Response accuracy on task categories, by hemisphere group. Dots represent mean accuracy scores for comprehension, interface, production, (morpho)syntactic tasks (*syntax*) within the left (L, in blue) and right (R, in coral) hemisphere lesion groups. Error bars indicate 95% confidence intervals

Patients with left-hemisphere lesions showed significantly higher accuracy on comprehension tasks, and a non-significant trend toward better performance on morphosyntactic tasks. These findings are consistent with the well-established role of the left hemisphere in morphosyntactic parsing and sentence-level integration (Dronkers et al., 2004; Friederici, 2011). In contrast, patients with right-hemisphere lesions performed marginally better on interface tasks and descriptively higher on production tasks. Overall, these results support the view of a distributed yet functionally specialised contribution of each hemisphere to language processing, modulated by task type.

When breaking down comprehension accuracy by lexical category, a more nuanced hemispheric pattern emerges (see Figure 3). Patients with left-hemisphere lesions outperformed their right-hemisphere counterparts on adjectives and nouns, whereas they showed markedly lower accuracy on adverbs and slightly lower performance on verbs. Patients with left-hemisphere lesions showed significantly higher comprehension accuracy than those with right-hemisphere lesions for adjectives ($W = 4788$, $p < 0.001$) and nouns ($W = 17748$, $p < 0.001$). Conversely, right-hemisphere patients outperformed those with left-hemisphere lesions for adverbs ($W = 180$, $p = 0.014$). No significant group difference was found for verbs ($W = 8118$, $p = 0.71$).

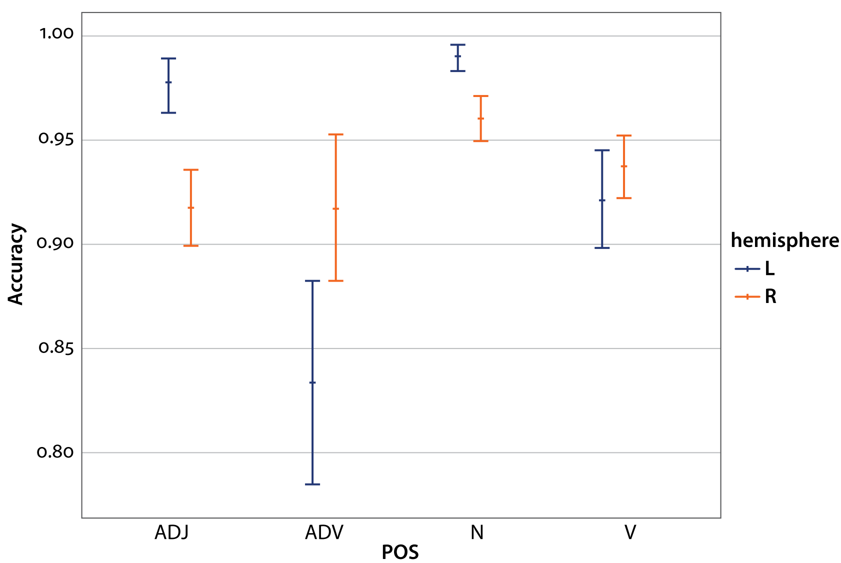

Figure 3. Response accuracy on comprehension tasks by lexical category (POS) and hemisphere group. Dots represent mean accuracy scores for adjectives, adverbs, nouns, and verbs in the comprehension task, plotted separately for patients with left- (L, in blue) and right-hemisphere (R, in coral) lesions. Error bars indicate 95% confidence intervals

This pattern may reflect differential hemispheric involvement in processing specific lexical-semantic domains. The left lesion group exhibited a poorer performance in processing verbs and adverbs, whereas nouns and adjectives were more challenging for the right lesion cohort. These results point to a functional lateralisation of the interface semantic dimensions involved in the word-class effect: the left lesion group exhibited a selective impairment in processing word classes that prototypically convey event-related information, while right hemisphere lesions produced an impairment in the processing of visuospatial information related to nouns and adjectives. Verbs, given their dual nature as both lexical and morphosyntactic carriers, are reported to engage a more bilateral neural network during language comprehension, which may explain the smaller performance differences observed between lesion groups for this category (Cappa & Perani, 2003; K.A. Shapiro et al., 2005). Overall, these results align with existing literature on word-class distinctions and reinforce the idea that the different semantic dimensions that underlay word class distinctions are supported by specialised neural networks across the cerebral hemispheres (Handjaras et al., 2016; Law et al., 2017; Papagno et al., 2009; Peelen et al., 2012; Shapiro & Caramazza, 2003; Shapiro, Shelton, & Caramazza, 2000; Williams et al., 2017; Xia & Peng, 2022; Zhang et al., 2018).

3.4 Correlation with MoCA performance

To explore the clinical relevance of performance on the proposed linguistic tasks, we examined correlations between accuracy scores and the total score on the *Montreal Cognitive Assessment* (MoCA). We also specifically considered the MoCA language subscore, designed to assess language abilities more directly.

A significant positive correlation emerged in the whole sample ($\rho = 0.71$, $p = 0.009$), indicating that higher overall MoCA scores were associated with better performance across all linguistic tasks in the BISMoS battery. When analysed separately by lesion group, this relationship remained strong and significant for patients with left-hemisphere lesions ($\rho = 0.93$, $p = 0.008$), but was weaker and not significant for the right-hemisphere group ($\rho = 0.49$, $p > 0.05$), suggesting a tighter coupling between clinical and experimental language measures in the left-lesioned population.

By contrast, the correlation between the MoCA language subscore and BISMoS accuracy was lower and not statistically significant ($\rho = 0.39$, $p > 0.05$), suggesting that the global MoCA score may better reflect the range of linguistic abilities tapped by the BISMoS battery than the dedicated language subtest alone. Interestingly, the correlation with BISMoS accuracy remained higher in patients with left-hemisphere lesions ($\rho = 0.56$) compared to those with right-hemisphere lesions ($\rho = 0.27$), although neither correlation reached significance. This pattern aligns with the expected dominant role of left-hemisphere structures in language processing.

To better understand which aspects of linguistic functioning contributed most to the observed association with MoCA scores, we conducted separate analyses for the four BISMoS task categories. The results revealed strong and significant correlations for comprehension ($\rho = 0.80$, $p = 0.002$), morphosyntax ($\rho = 0.76$, $p = 0.004$), and production tasks ($\rho = 0.68$, $p = 0.014$), indicating that performance in these linguistic domains closely tracked patients' general cognitive profile. By contrast, the correlation with interface tasks was weaker and not statistically significant ($\rho = 0.48$, $p > 0.05$), suggesting that this category may involve more complex or heterogeneous processes that are not directly captured by the MoCA. These correlations suggest that comprehension, non semantically determined morphosyntax, and production tasks may serve as sensitive markers of cognitive-linguistic integrity, while performance in interface tasks – potentially engaging broader integrative skills – might be less directly tied to general cognitive functioning as assessed by the MoCA.

When we examined the relationship between MoCA total scores and performance in specific linguistic domains, distinct patterns emerged across groups. In the left-lesioned group, comprehension was most strongly associated with overall cognitive functioning ($\rho = 0.97$, $p = 0.001$), with weaker but positive correlations observed for production, morphosyntax, and interface domains. In contrast, patients with right-hemisphere lesions showed significant correlations with both comprehension and morphosyntactic tasks ($\rho = 0.83$ for both, $p = 0.042$), while the correlation with production approached significance ($\rho = 0.78$, $p = 0.069$). No meaningful relationship was found for interface measures in either group. For a complete breakdown of domain-specific correlations, see Table 2. This pattern may suggest that, although left-hemisphere damage is classically associated with language deficits, right-hemisphere lesions may also impact linguistic performance in structured or morphosyntactically demanding tasks, possibly due to their broader role in executive control, inferencing, or prosodic processing. The strong comprehension correlations across both groups highlight the centrality of this domain in capturing cognitive-linguistic integrity.

Together, these findings support the convergent validity of the BISMoS tasks and highlight the added value of detailed linguistic assessments alongside global cognitive screening tools. In particular, BISMoS appears to capture individual variability in language functioning more finely than the MoCA, making it a valuable complement in the clinical profiling of language deficits. In fact, in our pilot study, BISMoS has been instrumental in revealing deficits at the interface-level of linguistic processing, highlighting its potential to identify dissociation patterns that may otherwise go undetected using traditional screening tools.

Table 2. Spearman correlations (ρ) between MoCA total scores and BISMoS domain accuracies, reported separately for left-hemisphere (L) and right-hemisphere (R) lesion groups

Domain	ρ (L)	p-value (L)	ρ (R)	p-value (R)
Comprehension	0.97	0.0013	0.83	0.0418
Production	0.69	0.1302	0.78	0.0687
Syntax	0.66	0.1502	0.83	0.0418
Interface	0.57	0.2404	0.17	0.7489

4. Discussion

4.1 Key findings

Our results highlight the need to consider both lateralisation and the interaction between cortical and subcortical structures when evaluating language outcomes following brain injury, as lateralisation alone did not produce significant differences in overall accuracy, whereas predictive models revealed effects related to lesion type.

The patient with a left subcortical lesion shows the most impaired and variable performance, underscoring the importance of left subcortical structures (such as the arcuate fasciculus and adjacent white matter pathways), in the processing of complex and multimodal linguistic information, including mapping meaning into morphosyntactic structures (Hickok & Poeppel, 2004) (Catani et al., 2005). The lack of a similar impairment for left hemisphere cortical lesions also highlights the crucial role of subcortical structures in information integration, indicating that a purely cortical lesion was not sufficient to significantly impair performance in the study population. Notably, the two patients with a contra-lateral subcortical lesion do not show the same impairment, suggesting hemispheric specialisation, with the left subcortical network playing a more dominant role for this aspect of linguistic processing (Kang, Herron, Ettlinger, & Woods, 2015).

In mixed lesions, the opposite pattern is observed, with accuracy dropping for the right lesion group. This trend could point in the direction of a compounding effect, where damage to both cortical and subcortical components of the right perisylvian network disrupts inferential or pragmatic aspects of comprehension, functions more commonly associated with the right hemisphere (Federmeier, Wlotko, & Meyer, 2008).

As mentioned in the population section (see 2.1), the site of stroke onset (M1-MCA) was consistent with damage to structures of the perisylvian language network (Catani et al., 2005).

The left temporal regions of the perisylvian language network (particularly the posterior superior temporal gyrus (pSTG), middle temporal gyrus (MTG), and superior temporal sulcus (STS), as well as underlying white matter structures) have been extensively described in literature as playing a crucial role in syntactic processing (Friederici, 2011) (Hickok & Poeppel, 2004). In fact, these regions show robust activation during the comprehension of syntactically complex sentences in healthy adults, and lesions to these areas are associated with selective deficits in parsing abilities in individuals with brain damage (Biondo, Ivanova, Pracar, Baldo, & Dronkers, 2024), (Sheppard et al., 2021) (Hickok, 2009). Despite these findings, two elements need to be taken into consideration.

Firstly, many studies do not clearly disentangle syntactic complexity from semantic complexity, and the two dimensions are often interrelated in stimuli design. For example, a syntactically simple sentence such as "The boy kicks the ball" typically involves a high-frequency, prototypical agent–theme relationship. In contrast, a syntactically complex sentence such as "The ball is kicked by the boy who the woman saw" not only includes passive voice and embedded clauses, but also disrupts typical thematic role assignment and introduces elements that are rare in everyday speech. In such cases, it is difficult to conclude that observed deficits are driven purely by syntactic complexity without the confounding influence of semantic or pragmatic factors (Caplan, Michaud, & Hufford, 2013). Additionally, the role of interface semantics has often been overlooked, as syntax is frequently treated as a homogeneous domain, despite evidence suggesting it encompasses distinct levels.

In our study, only patients with left temporal lesions exhibited deficits in interface tasks, with a significant difference with tasks involving non semantically determined morphosyntax. This pattern suggests that left hemisphere damage can selectively impair the ability to map meaning onto morphosyntax, while sparing purely formal morphosyntactic operations. This finding aligns with previous research suggesting that the processing of specific interface phenomena, such as telicity, relies on regions within the left temporal lobe (Romagno et al., 2012).

The observed dissociation between comprehension and production tasks in patients with right-hemisphere lesions adds to a growing body of evidence that implicates the right hemisphere in executive and inferential aspects of language processing (Federmeier et al., 2008). The more pronounced deficits observed in patients with cortico–subcortical lesions and the moderate yet significant impairments in those with subcortical-only lesions underscore the contribution of subcortical structures in supporting language networks that extend beyond cortical territories (Sihvonen et al., 2023). Although the absence of patients with exclusively cortical lesions limits conclusions about purely cortical contributions in this group, the data nonetheless seem to indicate that subcortical disruption alone can impair comprehension, while combined lesions amplify the deficit. These findings support a network-based model of language, in which the right hemisphere contributes critically to higher-level integration processes, and subcortical pathways play a key role in coordinating this function (Hickok & Poeppel, 2004) (Catani et al., 2005).

A finer-grained analysis of lexical categories in comprehension revealed that left-hemisphere lesions selectively impaired processing of verbs and adverbs — word classes conveying event-related information — while right-hemisphere lesions disproportionately affected nouns and adjectives, linked to visuospatial semantic dimensions. Verbs showed minimal group differences, consistent with their engagement of bilateral neural networks.

Finally, our correlational analyses confirmed the convergent validity of adopting a more focused linguistic approach, with a strong association between BISMoS performance and MoCA scores in comprehension, and moderate correlation in morphosyntax and production domains. By targeting lexical class distinctions and interface-level semantic processing, the proposed battery proved sensitive to subtle cognitive-linguistic deficits that may go undetected by standard screening tools.

4.2 Implications for models of lexical processing

The pattern we observed suggests that the neural organisation of lexical processing cannot be fully captured by word class distinctions alone, but rather reflects the interaction between semantic content, morphosyntactic properties, and the processing demands associated with different types of lexical items. Specifically, verbs and adverbs typically encode information about events, actions, and temporal or argument structure relations, which may preferentially engage left-lateralised networks involved in integrating meaning with morphosyntactic structure. In contrast, nouns and adjectives often convey information about object features, spatial relations, and perceptual properties, which may recruit right-lateralised networks more attuned to visuospatial and context-dependent semantic processing. Taken together, these preliminary findings offer tentative insights into the organisation of lexical knowledge and its neural underpinnings. They align with models proposing that lexical entries are represented in a multidimensional, distributed manner across hemispheres, encoding not only categorical word-class information but also semantic, morphosyntactic, and event-related features. Such a representation allows for both selective and interactive access to lexical knowledge during comprehension, depending on the properties of the word and the demands of the task (Hagoort, 2013; Levelt, Roelofs, & Meyer, 1999; Pulvermüller, 1999). These results, albeit based on a small pilot sample, suggest that studying lesion-specific deficits can provide a window into the neurocognitive architecture of lexical knowledge and the principles governing the mental lexicon. Accordingly, retrieval is not solely driven by categorical distinctions but dynamically shaped by the interaction between semantic content and structural integration demands. This perspective supports interactive models of the mental lexicon, where lexical entries are accessed through distributed networks sensitive to both linguistic and contextual constraints (Caramazza, 1997; Hagoort, 2013; Levelt et al., 1999; Pulvermüller, 1999). Furthermore, the differential impact of left- versus right-hemisphere lesions on category-specific comprehension underscores the distributed and partially redundant nature of lexical representations, which may support compensatory mechanisms following brain damage. Although our conclusions are necessarily tentative given the small size of the pilot sample, they suggest that lesion-specific patterns of impairment can provide a valuable window into the neurocognitive architecture of lexical processing and the principles governing the organisation of the mental lexicon. These theoretical considerations complement our methodological findings, reinforcing the relevance of BISMoS as a tool to probe lexical architecture and interface-level processing in clinical populations.

4.3 Methodological and clinical implications

Results of our pilot study highlight the critical role of stimulus design in experimental batteries seeking to target linguistic abilities. Specifically, the inclusion of tasks capable of isolating semantically driven features from purely morphosyntactic ones has proven essential for uncovering dissociations between the two levels of processing – a distinction that may have remained undetected with a less refined stimulus structure.

The observed divergence between interface and non-interface processing – and the fact that this divergence reached significance only in the left-lesion group – has important theoretical implications. Our findings not only support the view that interface and non-interface levels of morphosyntactic processing constitute distinct dimensions that can be selectively impaired, but they also contribute to provide evidence for the lateralisation of interface semantics processing in the left hemisphere.

Furthermore, the observed word-class effect in the two groups adds a valuable contribution to the ongoing debate on the subject, showing how lateralisation may differentially influence the processing of verbs/adverbs and nouns/adjectives in isolation, thus lending support to theoretical frameworks that view linguistic category organisation as grounded in broader, domain-general cognitive structures. These theoretical insights have direct clinical relevance, as they guide the design of assessment tools capable of capturing selective deficits.

From a clinical perspective, the study makes a compelling case for the integration of specifically designed linguistic stimuli to support improvement of linguistic abilities in post-stroke aphasia. In particular, incorporating tasks that target distinct linguistic dimensions in clinical evaluations could enhance the assessment of impaired domains which, in turn, could enable clinicians to implement more targeted therapeutic interventions and to monitor progress across specific linguistic domains over time.

The adoption of experimental protocols tailored to the structural and functional properties of a target language allows for a more accurate investigation of the mechanisms underlying linguistic processing. Such approaches enhance the sensitivity of experimental tasks to capture subtle variations in performance, especially in clinical populations. In our study, these methodological principles were operationalised through a battery of stimuli developed for Italian, a language whose morphosyntactic and lexical properties enabled fine-grained control of the linguistic variables under investigation. While the current design is language-specific, the underlying approach can be adapted to other languages, thus offering a flexible framework for studying the architecture of language processing in diverse linguistic systems.

Overall, the implementation of carefully designed batteries of linguistic stimuli – grounded in theoretical linguistics and aimed at disentangling the multiple dimensions involved in language processing – represents a promising direction for future research, with potential applications across a range of neurocognitive and clinical contexts.

4.4 Limitations

This study has several limitations that should be acknowledged and, possibly, overcome in future studies. Firstly, the limited size of the population hinders the statistical power of our findings. Additionally, the distribution of the cognitive reserve variable (YOE) was uneven, particularly in the left hemisphere lesion group, where only two levels of education (8 and 18 years) were represented, with just one participant at the higher level. Lesion type distribution was also unbalanced: the right hemisphere group included no participants with purely cortical lesions, and the left hemisphere group included only one subject with subcortical damage. Further, the time elapsed since symptoms onset (TSO) varied widely among participants, ranging from 2 months to 8 years, introducing further variability in recovery stage and neuroplasticity-related adaptation. For future studies, it would be desirable to target recruitment to ensure a broader spread of participants across lesion types, hemispheres, and years of education, given the known impact of cognitive reserve on recovery and linguistic outcomes.

Nonetheless, such limitations underscore a broader methodological challenge in clinical research: the inherent difficulty of assembling well-balanced experimental groups under real-world conditions.

A further limitation concerns lesion characterisation: although all patients presented with MCA-M1 lesions, the available clinical imaging did not allow for reliable estimation of lesion size at onset nor for standardised graphical mapping. As a result, our distinction between cortical, subcortical, and cortico-subcortical lesions is qualitative. Future studies with harmonised neuroimaging protocols will be essential to assess the contribution of lesion size and its interaction with lesion location, which have been shown to influence language outcomes in post-stroke aphasia (Thye & Mirman, 2018) and to account for aphasia severity (Meier, Johnson, Pan, & Kiran, 2019).

Finally, other cognitive factors known to influence language processing, such as working memory or executive functions, were not systematically assessed in the present study. Although MoCA measures indicated substantial inter-individual variability in attention and working memory – with the lowest scores observed in the participant with the lowest years of education – these measures were not included as covariates in the statistical analyses due to the limited sample size. In a small and heterogeneous cohort, such differences in cognitive reserve and domain-general processing may have contributed to variability in accuracy and to the non-linear relationship observed with time since stroke onset. This limitation is particularly relevant given that some interface-level tasks involve complex syntactic constructions, which may impose additional working memory demands and interact with domain-general deficits. Future research with a larger cohort should include standardised cognitive batteries as covariates to better disentangle the contributions of language-specific and domain-general cognitive processes.

Funding

This article was made Open Access under a CC BY-NC 4.0 license through payment of an APC by or on behalf of the authors.

Author roles

The authors' contribution to the present work, according to the CRediT taxonomy, is as follows: Conceptualisation, A.T., D.R., M.M., C.M.; methodology, A.T., D.R., M.M., C.M.; formal analysis, C.M.; investigation, A.T.; stimuli classification: D.R.; writing–original draft preparation, A.T., C.M.; writing–review & editing, A.T., D.R., C.M.; visualisation, C.M.; supervision, D.R., M.M., C.M. All authors have read and agreed to the published version of the manuscript.

Ethics statement

All participants were evaluated as part of their standard clinical care. This study, based exclusively on anonymised data, was conducted in accordance with the ethical standards of the institutional and national research committees and with the Declaration of Helsinki and its later amendments. According to institutional guidelines, the analysis did not require formal ethics committee approval.

References

doi Andersen, K. K., Olsen, T. S., Dehlendorff, C., & Kammersgaard, L. P. (2009). Hemorrhagic and ischemic strokes compared: stroke severity, mortality, and risk factors. *Stroke*, *40* (6), 2068–2072.

doi Arcara, G., & Bambini, V. (2016). A test for the assessment of pragmatic abilities and cognitive substrates (apacs): Normative data and psychometric properties. *Frontiers in psychology*, *7*, 70.

doi Baker, M. C. (2003). *Lexical categories: Verbs, nouns and adjectives* (Vol. *102*). Cambridge University Press.

doi Barner, D., & Bale, A. (2002). No nouns, no verbs: psycholinguistic arguments in favor of lexical underspecification. *Lingua*, *112* (10), 771–791.

doi Barner, D., … Bale, A. (2005). No nouns, no verbs? a rejoinder to panagiotidis. *Lingua*, *115* (9), 1169–1179.

doi Battista, P., Miozzo, A., Piccininni, M., Catricalà, E., Capozzo, R., Tortelli, R., … Logroscino, G. (2017, 12). Primary progressive aphasia: a review of neuropsychological tests for the assessment of speech and language disorders. *Aphasiology*, *31*, 1359–1378.

Bertinetto, P. M. (1986). *Tempo, aspetto e azione nel verbo italiano. il sistema dell'indicativo*. Accademia della Crusca.

doi Bertoni, D., Bruni, S., Saviola, D., De Tanti, A., & Costantino, C. (2024). The role of cognitive reserve in post-stroke rehabilitation outcomes: A systematic review. *Brain Sciences*, *14* (11), 1144.

doi Biondo, N., Ivanova, M. V., Pracar, A. L., Baldo, J., & Dronkers, N. F. (2024). Mapping sentence comprehension and syntactic complexity: evidence from 131 stroke survivors. *Brain Communications*, *6* (6), fcae379.

Bishop, D. (2009). *Trog 2: test for reception of grammar-version 2*. Giunti.

doi Brady, M. C. (2021). Predictors of poststroke aphasia recovery: A systematic review-informed individual participant data meta-analysis. *Stroke*, *52* (5), 1778.

doi Breitenstein, C., Grewe, T., Flöel, A., Ziegler, W., Springer, L., Martus, P., … others (2017). Intensive speech and language therapy in patients with chronic aphasia after stroke: a randomised, open-label, blinded-endpoint, controlled trial in a health-care setting. *The Lancet*, *389* (10078), 1528–1538.

doi Briggs, R. G., Pryor, D. P., Conner, A. K., Nix, C. E., Milton, C. K., Kuiper, J. K., … Sughrue, M. E. (2019, 6). The artery of aphasia, a uniquely sensitive posterior temporal middle cerebral artery branch that supplies language areas in the brain: Anatomy and report of four cases. *World Neurosurgery*, *126*, e65–e76.

doi Brownsett, S. L., Warren, J. E., Geranmayeh, F., Woodhead, Z., Leech, R., & Wise, R. J. (2014). Cognitive control and its impact on recovery from aphasic stroke. *Brain*, *137* (1), 242–254.

Capasso, R., & Miceli, G. (2001). *Esame neuropsicologico per l'afasia: Enpa*. Springer.

doi Caplan, D., Michaud, J., & Hufford, R. (2013). Dissociations and associations of performance in syntactic comprehension in aphasia and their implications for the nature of aphasic deficits. *Brain and language*, *127* (1), 21–33.

doi Cappa, S. F., & Perani, D. (2003). The neural correlates of noun and verb processing. *Journal of Neurolinguistics*, *16* (2–3), 183–189.

doi Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive neuropsychology*, *14* (1), 177–208.

doi Catani, M., Jones, D. K., & Ffytche, D. H. (2005). Perisylvian language networks of the human brain. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, *57* (1), 8–16.

doi Catricalà, E., Gobbi, E., Battista, P., Miozzo, A., Polito, C., Boschi, V., … others (2017). Sand: a screening for aphasia in neurodegeneration. development and normative data. *Neurological Sciences*, *38*, 1469–1483.

doi Catricalà, E., Rosa, P. A. D., Ginex, V., Mussetti, Z., Plebani, V., & Cappa, S. F. (2013, 6). An italian battery for the assessment of semantic memory disorders. *Neurological Sciences*, *34*, 985–993.

Chomsky, N. (1970). Remarks on nominalization. In R. Jacobs & P. Rosenbaum (Eds.), *Readings in english transformational grammar* (pp. 184–221). Ginn and Company.

doi Cichon, N., Wlodarczyk, L., Saluk-Bijak, J., Bijak, M., Redlicka, J., Gorniak, L., & Miller, E. (2021). Novel advances to post-stroke aphasia pharmacology and rehabilitation. *Journal of Clinical Medicine*, *10* (17).

doi Corbetta, M., Ramsey, L. E., Callejas, A., Baldassarre, A., Hacker, C. D., Siegel, J. S., … others (2015). Common behavioral clusters and subcortical anatomy in stroke. *Neuron*, *85* (5), 927–941.

doi Croft, W. (2001). *Radical construction grammar: Syntactic theory in typological perspective*. Oxford University Press.

doi Croft, W. (2010, 5). Pragmatic functions, semantic classes, and lexical categories. *Linguistics*, *48*, 787–796.

doi Dronkers, N. F., Wilkins, D. P., Van Valin Jr, R. D., Redfern, B. B., & Jaeger, J. J. (2004). Lesion analysis of the brain areas involved in language comprehension. *Cognition*, *92* (1–2), 145–177.

doi Engelter, S. T., Gostynski, M., Papa, S., Frei, M., Born, C., Ajdacic-Gross, V., … Lyrer, P.A. (2006). Epidemiology of aphasia attributable to first ischemic stroke: incidence, severity, fluency, etiology, and thrombolysis. *Stroke*, *37* (6), 1379–1384.

doi Federmeier, K. D., Wlotko, E. W., & Meyer, A. M. (2008). What's 'right'in language comprehension: Event-related potentials reveal right hemisphere language capabilities. *Language and linguistics compass*, *2* (1), 1–17.

doi Feigin, V. L., Brainin, M., Norrving, B., Martins, S., Sacco, R. L., Hacke, W., … Lindsay, P. (2022). World stroke organization (wso): global stroke fact sheet 2022. *International Journal of Stroke*, *17* (1), 18–29.

doi 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. *Archives of physical medicine and rehabilitation*, *97* (12), 2188–2201.

doi Fridriksson, J., den Ouden, D.-B., Hillis, A. E., Hickok, G., Rorden, C., Basilakos, A., … Bonilha, L. (2018). Anatomy of aphasia revisited. *Brain*, *141* (3), 848–862.

doi Friederici, A. D. (2011). The brain basis of language processing: from structure to function. *Physiological reviews*, *91* (4), 1357–1392.

doi Fujii, M., Miyachi, S., Matsubara, N., Kinkori, T., Takebayashi, S., Izumi, T., … others (2011). Selective propofol injection into the m1 segment of the middle cerebral artery (mca wada test) reduces adverse effects and enhances the reliability of the wada test for determining speech dominance. *World neurosurgery*, *75* (3–4), 503–508.

doi Gajardo-Vidal, A., Lorca-Puls, D. L., Hope, T. M., Parker Jones, O., Seghier, M. L., Prejawa, S., … Price, C. J. (2018). How right hemisphere damage after stroke can impair speech comprehension. *Brain*, *141* (12), 3389–3404.

Garrard, P., & Ahmed, S. (2012). An abbreviated examination for the assessment of linguistic impairment in primary progressive aphasia. *Eur J Neurol*, *19*, 561.

doi Geschwind, N. (1965). Disconnexion syndromes in animals and man. *Brain*, *88* (3), 585–585.

doi Gorno-Tempini, M. L., Hillis, A. E., Weintraub, S., Kertesz, A., Mendez, M., Cappa, S. F., … Grossman, M. (2011, 3). Classification of primary progressive aphasia and its variants. *Neurology*, *76*, 1006–1014.

doi Hagoort, P. (2013). Muc (memory, unification, control) and beyond. *Frontiers in Psychology*, *4*, 416.

doi Halai, A. D., Woollams, A. M., & Ralph, M. A. L. (2017, 1). Using principal component analysis to capture individual differences within a unified neuropsychological model of chronic post-stroke aphasia: Revealing the unique neural correlates of speech fluency, phonology and semantics. *Cortex*, *86*, 275–289.

doi Hamilton, R. H., Chrysikou, E. G., & Coslett, B. (2011). Mechanisms of recovery in aphasia: evidence from functional neuroimaging studies. *Brain and Language*, *118* (1–2), 40–50.

doi Handjaras, G., Ricciardi, E., Leo, A., Lenci, A., Cecchetti, L., Cosottini, M., … Pietrini, P. (2016, 7). How concepts are encoded in the human brain: A modality independent, category-based cortical organization of semantic knowledge. *NeuroImage*, *135*, 232–242.

Haspelmath, M. (2012). *How to compare major word-classes across the world's languages* (Vol. *17*) (No. 1). UCLA Working Papers in Linguistics Los Angeles: UCLA.

doi Hickok, G. (2009). The functional neuroanatomy of language. *Physics of life reviews*, *6* (3), 121–143.

doi Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition*, *92* (1–2), 67–99.

doi Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature reviews neuroscience*, *8* (5), 393–402.

Huber, W., Poeck, K., Weniger, D., & Willmes, K. (1983). *Aachener aphasie test (aat): Handanweisung*. Verlag für Psychologie Hogrefe.

Jakubíček, M., Kilgarriff, A., Kovář, V., Rychlỳ, P., & Suchomel, V. (2013). *The tenten corpus family*. In *7th international corpus linguistics conference cl* (p. 125–127).

doi Kakinuma, K., Osawa, S.-i., Hosokawa, H., Oyafuso, M., Ota, S., Kobayashi, E., … others (2022). Determination of language areas in patients with epilepsy using the super-selective wada test. *IBRO Neuroscience Reports*, *13*, 156–163.

Kandel, E. (2021). *Principles of neural science (6th edition)*. McGraw Hill.

doi Kang, X., Herron, T. J., Ettlinger, M., & Woods, D. L. (2015). Hemispheric asymmetries in cortical and subcortical anatomy. *Laterality: Asymmetries of Body, Brain and Cognition*, *20* (6), 658–684.

Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *The boston naming test. lea & febiger*. Philadelphia, PA.

Kiran, S. (2012). What is the nature of poststroke language recovery and reorganization? *International Scholarly Research Notices*, *2012* (1), 786872.

doi Kümmerer, D., Hartwigsen, G., Kellmeyer, P., Glauche, V., Mader, I., Klöppel, S., … Saur, D. (2013). Damage to ventral and dorsal language pathways in acute aphasia. *Brain*, *136* (2), 619–629.

Lakoff, G. (2008). *Women, fire, and dangerous things: What categories reveal about the mind*. University of Chicago Press.

doi Law, S. P., Yum, Y. N., & Cheung, G. W.-L. (2017, 11). Neural correlates of concreteness effect in semantic processing of single chinese characters using mixed-effects modeling. *Journal of Neurolinguistics*, *44*, 223–238.

doi 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.

Levin, B., & Hovav, M. R. (1994). *Unaccusativity: At the syntax-lexical semantics interface* (Vol. *26*). MIT press.

doi Luzzatti, C., De Bleser, R., Scola, I., Frustaci, M., & Willmes, K. (2023). Update on the psychometric properties for the italian version of the aachen aphasia test (it-aat). *Aphasiology*, *37* (4), 658–695.

Luzzatti, C., Willmes, K., De Bleser, R., Bianchi, A., Chiesa, G., De Tanti, A., … Pozzoli, C. (1994). Nuovi dati normativi per la versione italiana dell'aachener aphasia test. *Archivio di Psicologia, Neurologia e Psichiatria*, *55*, 1086–1131.

doi Marantz, A. (2013). No escape from morphemes in morphological processing. *Language and cognitive processes*, *28* (7), 905–916.

doi Matchin, W., & Hickok, G. (2020). The cortical organization of syntax. *Cerebral Cortex*, *30* (3), 1481–1498.

doi 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*, *13* (6), 1510–1525.

Mendelow, A. D. (2015). *Stroke: Pathophysiology, diagnosis, and management*. Elsevier Health Sciences.

Miceli, G., Laudanna, A., & Capasso, R. (2006). *Batteria per l'analisi dei deficit afasici*. EMS.

Mtui, E., Gruener, G., & Dockery, P. (2020). *Fitzgerald's clinical neuroanatomy and neuroscience e-book: Fitzgerald's clinical neuroanatomy and neuroscience e-book*. Elsevier Health Sciences.

doi Panagiotidis, P. (2005). Against category-less roots in syntax and word learning: objections to barner and bale (2002). *Lingua*, *115* (9), 1181–1194.

Panagiotidis, P. (2015). *Categorial features* (Vol. *145*). Cambridge University Press.

doi Papagno, C., Capasso, R., & Miceli, G. (2009, 3). Reversed concreteness effect for nouns in a subject with semantic dementia. *Neuropsychologia*, *47*, 1138–1148.

doi Peelen, M. V., Romagno, D., & Caramazza, A. (2012). Independent representations of verbs and actions in left lateral temporal cortex. *Journal of cognitive neuroscience*, *24* (10), 2096–2107.

doi Plowman, E., Hentz, B., & Ellis, C. (2012). Post-stroke aphasia prognosis: A review of patient-related and stroke-related factors. *Journal of evaluation in clinical practice*, *18* (3), 689–694.

doi Pulvermüller, F. (1999). Words in the brain's language. *Behavioral and brain sciences*, *22* (2), 253–279.

R Core-Team. (2024). *R: A language and environment for statistical computing*. Foundation for Statistical Computing. Vienna: Austria.

Romagno, D. (2012). Beyond the mental lexicon: evidence from neuropsychology and implications for the morphosyntax/semantics interface. *Studi e Saggi Linguistici*, *50* (2), 97–118.

Romagno, D. (2016). Principles of categorization: the case of word classes. *Archivio Glottologico Italiano*.

Romagno, D. (2019). Exploring the neurocognitive basis of language processing: the case of word classes and the key role of interface semantics. *Atti del Sodalizio Glottologico Milanese*, *2018* (13 ns).

doi Romagno, D. (2024). Word classes and neurolinguistics. In A. De Cesare Greenwald & G. Salvi (Eds.), *Manuals of romance linguistics* (p. 797–818). Boston: De Gruyter.

Romagno, D., Busso, L., & Elce, V. (2020). *L'architettura neurocognitiva del linguaggio fra concetti e grammatica. BISMoS: una batteria di test per lo studio dell'interfaccia fra semantica e morfosintassi in italiano*. Edizioni ETS.

doi Romagno, D., Rota, G., Ricciardi, E., & Pietrini, P. (2012). Where the brain appreciates the final state of an event: The neural correlates of telicity. *Brain and language*, *123* (1), 68–74.

doi Rosch, E. H. (1973). Natural categories. *Cognitive psychology*, *4* (3), 328–350.

doi Saur, D., Lange, R., Baumgaertner, A., Schraknepper, V., Willmes, K., Rijntjes, M., & Weiller, C. (2006). Dynamics of language reorganization after stroke. *Brain*, *129* (6), 1371–1384.

doi Shapiro, K., & Caramazza, A. (2003). Looming a loom: Evidence for independent access to grammatical and phonological properties in verb retrieval. *Journal of Neurolinguistics*, *16* (2–3), 85–111.

doi Shapiro, K., Shelton, J., & Caramazza, A. (2000). Grammatical class in lexical production and morphological processing: Evidence from a case of fluent aphasia. *Cognitive Neuropsychology*, *17*, 665–682.

doi Shapiro, K. A., Mottaghy, F. M., Schiller, N. O., Poeppel, T. D., Flüß, M. O., Müller, H.-W., ... Krause, B. J. (2005). Dissociating neural correlates for nouns and verbs. *Neuroimage*, *24* (4), 1058–1067.

doi Sheppard, S. M., Meier, E. L., Kim, K. T., Breining, B. L., Keator, L. M., Tang, B., ... Hillis, A. E. (2021). Neural correlates of syntactic comprehension: A longitudinal study. *Brain and language*, *225*, 105068.

doi Shetreet, E., Friedmann, N., & Hadar, U. (2010). The neural correlates of linguistic distinctions: Unaccusative and unergative verbs. *Journal of Cognitive Neuroscience*, *22* (10), 2306–2315.

doi Sihvonen, A. J., Vadinova, V., Garden, K. L., Meinzer, M., Roxbury, T., O'Brien, K., ... Brownsett, S. L. (2023). Right hemispheric structural connectivity and poststroke language recovery. *Human Brain Mapping*, *44* (7), 2897–2904.

doi Stern, Y. (2002). What is cognitive reserve? theory and research application of the reserve concept. *Journal of the international neuropsychological society*, *8* (3), 448–460.

doi Stern, Y. (2012). Cognitive reserve in ageing and alzheimer's disease. *The Lancet Neurology*, *11* (11), 1006–1012.

doi Tenny, C. L. (1994). *Aspectual roles and the syntax-semantics interface* (Vol. 52). Springer Netherlands. Retrieved from http://link.springer.com/10.1007/978-94-011-1150-8.

doi Thye, M., & Mirman, D. (2018). Relative contributions of lesion location and lesion size to predictions of varied language deficits in post-stroke aphasia. *NeuroImage: Clinical*, *20*, 1129–1138.

doi Tucker, A. M., & Stern, Y. (2014). Cognitive reserve and the aging brain. *Geriatric neurology*, p. 118–125.

doi Valenzuela, M. J., & Sachdev, P. (2006). Brain reserve and dementia: a systematic review. *Psychological medicine*, *36* (4), 441–454.

doi Van Valin, R. D., & LaPolla, R. J. (1997). *Syntax: Structure, meaning, and function*. Cambridge University Press.

doi Van Valin Jr, R. D. (1990). Semantic parameters of split intransitivity. *Language*, 221–260.

doi Van Valin Jr, R. D. (2005). *Exploring the syntax-semantics interface*. Cambridge University Press.

Vendler, Z. (1967). Verbs and times. *The Philosophical Review*, *2*, 143–160.

doi Williams, A., Reddigari, S., & Pylkkänen, L. (2017, 6). Early sensitivity of left perisylvian cortex to relationality in nouns and verbs. *Neuropsychologia*, *100*, 131–143.

doi Wilson, S. M., Entrup, J. L., Schneck, S. M., Onuscheck, C. F., Levy, D. F., Rahman, M., ... others (2023). Recovery from aphasia in the first year after stroke. *Brain*, *146* (3), 1021–1039.

doi Xia, Q., & Peng, G. (2022, 2). The roles of object and action, and concreteness and imageability, in the distinction between nouns and verbs: An erp study on monosyllabic words in chinese. *Journal of Neurolinguistics*, *61*.

doi Zhang, Y., Wang, K., Yue, C., Mo, N., Wu, D., Wen, X., & Qiu, J. (2018, 4). The motor features of action verbs: fmri evidence using picture naming. *Brain and Language*, *179*, 22–32.

Appendix A

In this section, we provide a selection of items from the experimental battery, organised according to the main linguistic domains assessed: interface-level processing (Table 3), morphosyntactic features (Table 4), lexical production (Table 5), and sentence comprehension (Table 6). An asterisk (*) marks either the incorrect alternative or the unacceptable phrase. English translations follow each stimulus.

Table 3. Examples of BISMoS stimuli from the *Morphosyntax/semantics interface* category

Sub-category	Task Typology	Stimulus Example
Telicity	Guided phrase completion	*Pino ha trovato le chiavi _ (*PER/IN) un minuto* *Tom found his keys _ (*FOR/IN) one minute*
Agentivity I	Acceptability judgment	*È un governatore determinato* *He is a determined governor* **È un ricordatore serio* **He is a serious rememberer*
Agentivity II	Acceptability judgment	*Guarda dove vai, Stefania!* *Look where you're going, Lucy!* **Manca di intelligenza, subito!* **Lack intelligence, now!*
Dinamicity	Acceptability judgment	*La nostra squadra sta vincendo* *Our team is winning* **Un mio amico sta discendendo da Leopardi* **A friend of mine is descending from Leopardi*
Split Intransitivity	Guided phrase completion	*I soldati _ (*SONO/HANNO) marciat_ (O/*I) fino all'ingresso della città* *The soldiers _ (*ARE/HAVE) marched _ (M.SG/*M.PL) to the city's entrance*
Individuation	Guided phrase completion	*C'è ancora _ (DELLA/*UNA) neve sulla strada* *There still is _ (SOME/*ONE) snow on the street*

Table 4. Examples of BISMoS stimuli from the *Non-semantically determined morphology and syntax* category

Sub-category	Task Typology	Stimulus Example
Verbal Morphology	Acceptability judgment	*Pio e Gino hanno lavorato tutta la notte* *Jim and Henry worked all night long* **Ieri Mariagiulia è partata per le vacanze* **Yesterday Mary laft for her holidays*
Syntax	Acceptability judgment	*Le ragazze hanno chiacchierato per un'ora* *The girls have chatted for an hour* **Teo è andato comprare a un vocabolario di inglese* **Teo went buy to an English dictionary*

Table 5. Examples of BISMoS stimuli from the *Production* category

Sub-category	Task Typology	Stimulus Example
Coradical Verbs and Nouns	Guided phrase completion	*La corsa delle due amiche ogni mattina → Le due amiche _ ogni mattina* *The run of the two friends every morning → The two friends _ every morning*
Adjectives	Phrase completion	*David ama i dolci: è molto _* *David loves desserts: he is very _*

Table 6. Examples of BISMoS stimuli from the *Comprehension* category

Sub-category	Task Typology	Stimulus Example
Nouns	Semantic similarity judgment	*bambino — *strada — ragazzo* *child — *street — boy*
Verbs	Semantic similarity judgment	*lavare — *pensare — pulire* *to wash — *to think — to clean*
Adjectives	Semantic similarity judgment	*ultimo — primo — *disponibile* *last — first — *available*
Adverbs	Semantic similarity judgment	*dietro — *contro — avanti* *behind — *against — forward*

Table 6. *(continued)*

Sub-category	Task Typology	Stimulus Example
Nouns	Guided phrase completion	*Hai spento le _ (LUCI/*PENNE) prima di uscire?*
		*Did you switch off the _ (LIGHTS/*PENS) before going out?*
Verbs	Guided phrase completion	*Le ragazze _ (*NUOTANO/STUDIANO) in biblioteca*
		*The girls _ (*SWIM/STUDY) in the library*
Adjectives	Guided phrase completion	*Sono molto _ (STANCO/*BIONDO), vado a dormire*
		*I am very _ (TIRED/*BLOND), I'm going to sleep*

Appendix B

The following tables summarise the results of the regression models fitted to explore the effects of lesion lateralisation, and task type on participants' performance. For each model, relevant predictors, coefficients, and statistical significance levels are reported. These analyses were conducted to investigate the potential contribution of linguistic and clinical variables to performance variability across tasks.

Firstly, we fit a generalised linear mixed-effects model (GLMM) with a binary predictor for response accuracy based on lesion laterality group (L vs. R). A random intercept was included for patients (Table B1).

Table B1. Summary of the GLMM assessing the effect of lesion side on response accuracy

Fixed Effect	Estimate	Std. Error	z value	Pr(> |z|)
(Intercept)	4.181	0.681	6.144	¡ 0.001
Group (R)	0.088	0.956	0.092	0.926
AIC / BIC	626.3 / 643.6			
Log Likelihood	−310.2			
Number of observations	2328			
Random effect (Patients)	Variance = 2.27, SD = 1.507			
Conditional R^2 (total)	0.4085			
Deviance explained	40.8%			

To further explore the factors contributing to accuracy performance, a logistic regression model (GLM) was implemented. This model evaluated the predictive impact of lesion type (0=cortical, 1=subcortical, 2=cortico-subcortical), lesion lateralisation group (L vs. R), years of education (YOE), and time since lesion onset (TSO) (Table B2).

Table B2. Summary of the GLM assessing predictors of response accuracy

Predictor	**Estimate**	**Std. Error**	**z value**	**Pr(> $\|z\|$)**
(Intercept)	0.701	0.521	1.347	0.178
Lesion type 1	0.665	0.408	1.629	0.103
Lesion type 2	1.578	0.515	3.067	0.002
Group (R)	−0.612	0.273	−2.242	0.025
YOE	0.260	0.050	5.184	< 0.001
TSO	−0.177	0.068	−2.589	0.010
Model fit statistics:				
AIC	660.2			
residual deviance	648.2 on 2322 df			
Pseudo-R^2	(McFadden) = 0.078			

Additionally, we fitted a generalised additive model (*GAM*) with a binarized version of time since onset (with TSO <1 vs. >1 year) for years of education (YOE), and lesion type as a function of lesion group. The interaction term *Lesion type 2:Group (R)* returned *NaN* due to lack of variance in the corresponding predictor combination. While data were available for both lesion types across groups, the model could not estimate the interaction coefficient (Table B3).

Table B3. Summary of the logistic GAM assessing interaction of predictors of response accuracy

Predictor	**Estimate**	**Std. Error**	**z value**	**Pr(> $\|z\|$)**
(Intercept)	0.784	0.731	1.072	0.284
Lesion type 1	−1.014	0.774	−1.311	0.190
Lesion type 2	1.743	0.647	2.696	0.007
Group (R)	−2.283	0.568	−4.017	< 0.001
YOE	0.243	0.076	3.214	0.001
TSO > 1	0.831	1.295	0.641	0.521
Lesion type 1:Group (R)	5.918	1.491	3.970	< 0.001
Lesion type 2:Group (R)	0.000	0.000	NaN	NaN
YOE:TSO > 1	−0.048	0.099	−0.482	0.630

Table B3. *(continued)*

Predictor	Estimate	Std. Error	z value	Pr(> |z|)
Model fit statistics:				
REML score	304.6	N	2328	
Adjusted R^2	0.038			
Deviance explained	14.2%			

Finally, we fit a generalised additive model (GAM) with a binomial link to predict accuracy based on lesion laterality group (L vs. R), lesion type (1=subcortical, 2=cortico-subcortical), and their interaction. A random smooth term for patients was included to account for subject-level variability. Only patients with subcortical (1) or cortico-subcortical (2) lesions were included (Table B4).

Table B4. Summary of the logistic GAM model predicting trial-level accuracy from lesion type and group

Term	Estimate	Std. Error	z value	Pr(> |z|)
(Intercept)	2.163	1.456	1.486	0.1372
Group (R)	3.929	2.071	1.897	0.0578
Lesion type (2)	2.803	1.756	1.596	0.1105
Group (R):Lesion type (2)	−5.492	2.416	−2.274	0.0230
smooth terms:				
s(Patients)	4.486	6	27.07	6.21e-06
Model fit statistics:				
Adjusted R^2	0.0557			
Deviance explained	18.2%			

Address for correspondence

Claudia Marzi
Institute for Computational Linguistics ”A. Zampolli”
Italian National Research Council
Via G. Moruzzi
1-56124 Pisa
Italy

claudia.marzi@ilc.cnr.it
https://orcid.org/0000-0002-3427-2827

Co-author information

Alice Todesco
Italian National Research Council
Institute for Computational Linguistics "A. Zampolli"

a.todesco2@gmail.com
https://orcid.org/0009-0007-2763-5814

Domenica Romagno
Department of Philology, Literature and Linguistics
University of Pisa

domenica.romagno@unipi.it

Michelangelo Mancuso
Neurological Institute, Department of Clinical and Experimental Medicine
University of Pisa

michelangelo.mancuso@unipi.it
https://orcid.org/0000-0003-2738-8562

Publication history

Date received: 16 June 2025
Date accepted: 26 February 2026
Published online: 7 May 2026