Science of Aphasia 2026

September 22 to 25, 2026
at the University of Bordeaux
campus Carreire

Programme

Conference programme SoA Bordeaux 2026

Day 1 (22 September 2026; Tuesday)

09.00-09.30 Registration
09.30-09.45 Welcome

Theme*:* Discourse/Mild aphasia

09.45-10.45 Keynote: Lucy Dipper

10.45-11.15 Break

11.15-12.15 Keynote: Eva Soroli

12.15-12.45 Discussion (Chair: Sylvia Martínez Ferreiro)

12.45 - 13.45 Lunch

13.45-14.15 3-minute poster presentation pitches

1. Sophie Arheix-Parras- Anterior Temporal Lobe stimulation using rTMS for post-stroke aphasia recovery (ATLAS): a study protocol for a single-blinded randomized clinical trial
2. Noor-E-Zannat Nusha - The complementiser que in Spanish connected speech: a morphosyntactic signature of dementia
3. Elizabeth Galletta- Tradeoff Patterns in Connected Language Production in Post-Stroke Aphasia Following Speech-Language Therapy
4. Claudia Peñaloza - Learning and maintenance of L3 novel words in a multilingual person with aphasia: Is there a cognate facilitation effect?
5. Christiana Spiliou - Syntactic abilities and short-term memory in the behavioral variant of frontotemporal dementia
6. Jo-Anne van der Sluijs - Morphological Relations and Word Networks in People with Post-Stroke Aphasia
7. Manon Winkler - A German Test for Thematic and Taxonomic Relations in Awake Brain Surgery: Standardization and Validation of the GB-WOBA
8. Lena Rybka - Longitudinal subjective accounts of co-occurring language and memory complaints in glioma patients undergoing brain surgery and comparison to neuropsychological assessment
9. Dilan Joseph García Inclán - Embedding-Based Discourse Coherence in Spanish Dementia Speech

14.15 - 15.15 Poster session 1

15.15-16.45 Oral presentation 1
1. Alycia Laks - Psycholinguistic effects on word-level comprehension: results from a novel web-based word-picture matching task
2. Michael Scimeca - Overall Language Impairment and Cross-Language Effects Shape Reading in Bilinguals with Aphasia
3. Marina Ruiter - Dynamic assessment of everyday communication using Virtual Reality (DCOM-VR)

4. Imke Wets - Investigating the comprehension of discourse particles via a selection test in Dutch-speaking persons with aphasia
5. Silvia Marchesotti - Behavioral and neurophysiological effects of transcranial direct current stimulation to the left inferior frontal gyrus in patients with expressive aphasia
6. Rosemarije Weterings - Measuring confrontation naming under increased cognitive demands in young stroke: the Picture Naming Cards Test

Day 2 (23 September 2026; Wednesday)

Theme: Language impairments in understudied diseases

09.00-10.00 Keynote: Adrià Rofes

10.00-10.30 Coffee break

10.30-11.30 Keynote: Agnès Trébuchon

11.30-12.00 Discussion (Chair: Yuval Z. Katz)

12.00 - 13.00 Lunch

13.00-13.30 3-minute poster presentation pitches

1. Solène Hameau - Where is the locus of language mixing in bilingual aphasia? A preliminary comparison of two cases.
2. Cassiady Perard - Silent and Filled Pauses as Markers of Longitudinal Language Change in Mild Primary Progressive Aphasia: A Case Study
3. Marie Couvreu - Associative Strength and Semantic Proximity: A Cross-Modal Investigation of Lexical-Semantic Network Organisation
4. Silvia Hargrove - Dissociation of Retrieval and Inflection of Verbs and Adjectives in Bilingual Speakers with Aphasia
5. Ana Matić - Elaboration of events in a picture description task: a case of Croatian and Norwegian PWA
6. Dörte de Kok - The multilingual Token Test App
7. Georgia Roumpea - Investigating deverbal nouns in individuals with neurodegenerative syndromes: evidence from Slovene
8. Irina Chupina - How interrelated are semantic and episodic memory retrieval? An MEG study of healthy individuals and people after left temporal lobe epilepsy surgery
9. Ioannis Papakyritsis - The effect of lingustic complexity on temporal aspects of speech production in normal Aging: Data from Greek

13.30-14.30 Poster session 2

14.30-16.00 Oral presentations 2
1. Yuval Z. Katz - Hidden in the treetop: Moved and non-moved prepositions in agrammatism
2. Michal Láznička - Disfluencies in the production of Czech prepositional phrases
3. Mauro Viganò - Abstract morphosyntactic features versus overt morphophonological cues in the comprehension of French object relatives in Primary Progressive Aphasia
4. Abdullah Mahmud Alauddin - Frequency and Variation of Code-Switching Patterns in Spanish-Galician Bilinguals with Cognitive Decline

5. Sopho Tsagareishvili - First insights on nonfluent aphasia in Georgian: A quantitative analysis of spontaneous speech with a focus on morphosyntax
6. Jaya Ratane - Aphasia and Neurodegenerative Language Disorders in Creole Speakers: A Scoping Review

Day 3 (24 September 2026; Thursday)

Symposia

09.00-10.15 Symposium 1
Jean-Marie Annoni - Living well with aphasia: research and topics priorities identified by stroke survivors and aphasia associations

10.15-10.45 Coffee break

10.45-12.00 Symposium 2
Maria Varkanitsa - Rethinking Assessment in Aphasia: Discourse Ecosystems, Interpretable Metrics, and Clinical Relevance

12.00 - 13.00 Lunch

13.00-13.30 3-minute poster presentation pitches

1. Rola Andrawes - Lexical diversity and language mixing in two Egyptian Colloquial Arabic - English bilingual people with aphasia
2. Jean-Marie Annoni - Case report on the transition from phonetic error to correct production in a chronic PWA with spectral analysis
3. Gregoire Python - A group therapy approach using generative artificial intelligence and accessibility features of smart devices to support written communication and social participation in aphasia
4. Rinn Wang - A Usage-based Approach to Assessing Sentence Comprehension
5. Madeline Philipsen - Noun and Verb Association in German: Test Standardisation and Validation of the GB-WoBA
6. Tamás Káldi - Dynamic Scaling of Sentence Processing Effort in Hungarian Speakers with Aphasia: A Pupillometry Study
7. Io Salmons - Main Concept Analysis in Catalan and Spanish: performance of unimpaired bilingual speakers
8. Johanna Scholtz - Sentence completion for pre- and postoperative language testing in neurosurgical patients
9. Katerina Gkatzioni - Syntactic impairment in Huntington’s disease: a case series study

13.30-14.30 Poster session 3

Social event

Day 4 (25 September 2026; Friday)

Theme: Advances in understanding aphasia recovery

09.00-10.00 Keynote: Fatemeh Geranmayeh

10.00-10.30 Coffee break

10.30-11.30 Keynote: Gesa Hartwigsen

11.30-12.00 Discussion (Chair: Vitória Piai)

12.00 - 13.00 Lunch

13.00-13.30 3-minute poster presentation pitches

1. Anaëlle Filisetti - A study of the impact of speech rhythm and semantic context on comprehension and production in Parkinson's disease
2. Djaina Satoer - Mild aphasia in patients with brain tumors and after stroke: objective and subjective insights
3. Lotta Heidemann - Beyond words: Capturing Creativity in Speech-Replacing gestures in people with aphasia
4. Alberto Osa Garcia - Acute white matter integrity differentially predicts early connected speech recovery in post-stroke aphasia
5. Vivien Vásári - Hungarian adaptation of the Apraxia of Speech Rating Scale–3.5: A pilot study of differential diagnosis in post-stroke AOS and aphasia
6. Laure Salpéteur - Effects of psycholinguistic variables on repeated written naming attempts across aphasic deficits
7. Mareike Hartmann - Novelty meets Effectiveness: Creative Naming Strategies in Persons with Aphasia
8. Caroline Rogue - Language assessment using immersive virtual-reality relative to standard tests: results on neurotypical adults and preliminary validation on people with aphasia
9. Anna-Maria Strinzel - 'Striped shirt in blue and orange' – Verbose discourse production in people with focal epilepsy

13.30-14.30 Poster session 4

14.30-16.00 Oral presentations 3
1. Ioana Sederias - Understanding language change in cognitive aging from a usage-based perspective: Implications for Aphasia
2. Frank Tsiwah - Towards preserving errors: fine-tuning automatic speech recognition for literal transcription of aphasic speech
3. Ludovica Onofri - Grammatical Tense Impairment in Aphasia: a usage-based investigation of verb forms in spontaneous language production
4. Vitor Zimmerer - A usage-based approach to language in Alzheimer's disease: the effect of multiword-frequency variables on patterns of impairment
5. Josefina Castillo Cerutti - The Emoji Naming Test: Comparing naming performance in controls and people with aphasia
6. Ana María Estrada Pineda Temporal Indicators of Bilinguals' Early Cognitive Decline in Verbal Fluency Tasks

Abstracts

Abstracts are provided in alphabetical order (surname 1st author). Email addresses of the corresponding author are provided on page 168. The description and short abstracts of the symposia can be found at the end of this booklet.

Frequency and Variation of Code-Switching Patterns in Spanish-Galician Bilinguals with Cognitive Decline

Abdullah Mahmud Alauddin | Silvia Martínez-Ferreiro

Introduction and aims

This study investigates code-switching (CS) patterns in older Spanish dominant Spanish-Galician bilingual adults with mild cognitive impairment (MCI) and dementia. The analysis focuses on the frequency and type of switches (word and clause level) with special attention to adverbs, prepositions, and pronouns. Building on the ProGram theory (Boye & Harder, 2012), which distinguishes lexical items (discursively primary and independent) from grammatical items (discursively secondary and dependent), Martínez-Ferreiro and Boye (2019) draws on evidence from bilingual aphasia showing that isolated insertions from a non-target language are predominantly lexical, while grammatical elements tend to appear only in combination with a lexical host from the same language. Further work on prepositions and pronouns demonstrates that these categories are internally heterogeneous and must be divided into lexical and grammatical subtypes, each showing distinct vulnerability profiles in aphasia (Martínez-Ferreiro et al., 2019). The current study characterizes CS patterns in older Spanish-Galician bilingual adults with MCI and dementia, focusing on (a) the frequency of switches across groups with differing degrees of cognitive decline, and (b) the types of switches, including instances of code-switching and code-mixing, and their lexical-grammatical status, to delve deep into the role of this phenomenon in cognitive decline. Special focus is given to whether isolated grammatical switches, items produced without a lexical host from the same language, increase as a function of neurodegeneration. Based on previous findings, the switching frequency is expected to follow a gradient of Dementia > MCI > Controls and word-level insertional switches to predominate over clause-level alternations in the clinical groups.

Methods

Spontaneous speech data was collected from 84 (68 females and 16 males) Spanish-dominant older Spanish-Galician bilingual adults (mean age = 78.8) divided into three groups — dementia (n = 28), MCI (n = 27), and neurotypical controls (n = 29) — through semi-structured interviews comprising picture description, narrative tasks, and conversational interaction. The participants were recruited through the LINDEGALES (ED431F 2024/018) and SECONDSGALES (PID2023-151673OB-I00) projects and they went through a cognitive and a linguistic evaluation. Samples were transcribed following CLAN conventions and were processed using CHAT. The transcription and subsequent review were completed by two trained linguists and disagreements were solved through consensus. Necessary variables were extracted from the .cha files using Python and the dataset was then analysed using R for results visualisation. For the frequency of code-switching, all switches were annotated. Participants were compared by group to evaluate group-wise frequency of code-switching. The switches were also marked at word, constituent, and clause levels before extracting a ratio of switch counts. All switches were combined and compared by groups to capture code-switching tendencies among participants who switched language during data collection. Statistical analyses are in progress at the time of submission.

Results

The proportion of participants who produced at least one switch was 24.1% in the control group (n = 7/29), 14.8% in MCI (n = 4/27), and 28.6% in dementia (n = 8/28). The dementia group showed the highest switching rate, followed by the control group. MCI had the lowest percentage of participants who switched language.

Among the code-switching participants, Spanish-Galician was the most common combination in all groups, with English intrusions only found in the speech output of one control and one participant with dementia.

Total switches were computed as the sum of clause and constituent switches, and word-level insertions per participant. This overall count was markedly higher in the dementia group (M = 42.1, Mdn = 13) compared to MCI (M = 3.2, Mdn = 3.5) and controls (M = 3, Mdn = 2.0), whose performance was similar. The dementia group also showed considerably greater within-group variability, driven by two participants with total switch counts of 212 and 56 respectively, against a background of more moderate switching in the remaining six participants with dementia. Both the control participants and the participants with MCI who switched languages clustered at low total counts, although both the mean and the median of switches in the MCI group was slightly higher.

Discussion

While the initial hypothesis expected the switching gradient to be Dementia > MCI > Controls, overall, participants with MCI code-switched less than both the control group and the participants with dementia. This may be explained by the fact that proactive monitoring remains relatively intact in MCI (Calabria et al. 2025), preventing them from switching languages as a compensatory strategy for incipient difficulties. On the other hand, participants with dementia showed the highest tendency to code-switch, complying with the hypothesis, which may suggest weakened cognitive control. Control participants, however, show comparatively high tendency of code-switching suggesting that the percentage of speakers choosing to switch language during a conversation alone cannot act as a reliable marker of cognitive decline without strict control of the sociolinguistic context.

If we focus only on those participants who produce examples of code-switching and mixing, the hypothesized pattern emerges: Dementia > MCI > Controls. This shows that although some MCI participants successfully monitor their speech and refrain from switching language through potential hypercorrection, those who do switch, do it more often than neurotypical participants. Ongoing fine-grained statistical analyses will help us to draw a clear picture of the role of code-switching in older Spanish-Galician bilingual adults and how this changes in the different stages of cognitive decline.

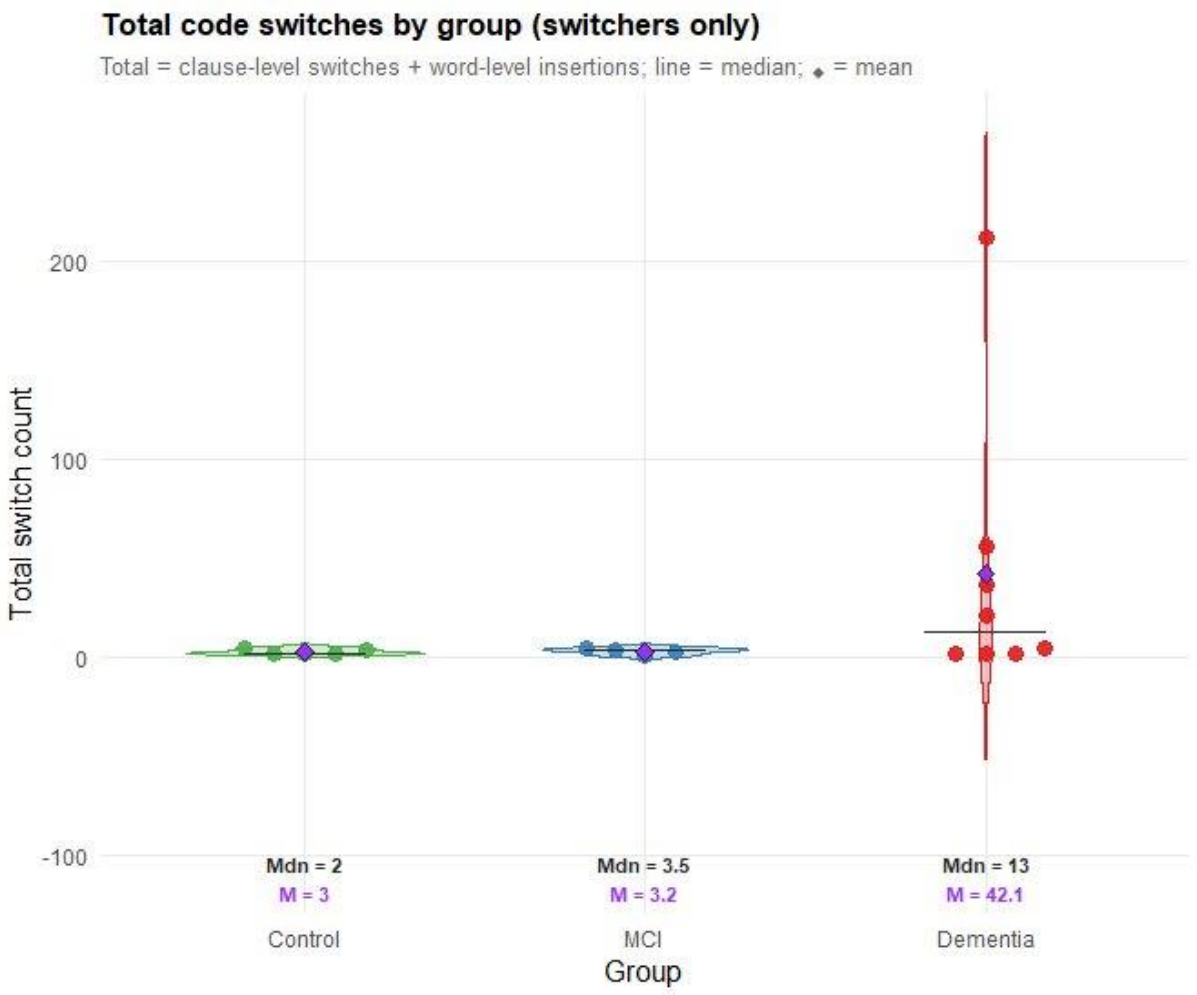
Acknowledgement

This work is a part of the SECONDSGALES research project led by Silvia Martinez-Ferreiro (PID2023-151673OB-I00 funded by MICIU/AEI/10.13039/501100011033 and by “ERDF/EU"). Data for this research has also been extracted from the LINDEGALES (ED431F 2024/018; Xunta de Galicia) project. SMF also acknowledges support of RYC2020-028927-1, MICIU/AEI/10.13039/501100011033 and ESF investing in your future. A special

thank you to the Gerontology and Geriatrics Research Group and its members for assistance with data collection, transcription, and providing institutional support.

References

Boye, K., & Harder, P. (2012). A usage-based theory of grammatical status and grammaticalization. *Language*, *88*(1), 1-44. https://dx.doi.org/10.1353/lan.2012.0020
Calabria, M., Suades, A., Juncadella, M., Ortiz-Gil, J., Ugas, L., Sala, I., & Lleo, A. (2025). Dissociating language switching from executive control in bilinguals with MCI. *Cortex*, *183*, 290-303. https://doi.org/10.1016/j.cortex.2024.11.020
Martínez-Ferreiro, S., & Boye, K. (2019). Bilingual strategies for aphasia: Evidence from typologically close languages. In S. Halupka-Rešetar & S. Martínez-Ferreiro (Éd.), *Studies in language and mind 3. Selected papers from the 6th & 7th Workshop in Psycholinguistics, Neurolinguistics and Clinical linguistic research* (p. 153-167). Novi Sad, Serbia : Filozofski fakultet u Novom. https://digitalna.ff.uns.ac.rs/sadrzaj/2019/978-86-6065-558-7
Martínez-Ferreiro, S., Ishkhanyan, B., Rosell-Clari, V., & Boye, K. (2019). Prepositions and pronouns in connected discourse of individuals with aphasia. *Clinical Linguistics & Phonetics*, *33*(6), 497-517. https://doi.org/10.1080/02699206.2018.1551935



Lexical diversity and language mixing in two Egyptian Colloquial Arabic - English bilingual people with aphasia

Rola Andrawes | Mauro Viganò | Seçkin Arslan

Introduction and aims

Language recovery in bilingual people with aphasia (PWA) can follow different patterns predicted by diverse factors such as age of acquisition, proficiency and frequency of use of languages in their daily life (Fabbro, 2001; Goral et al., 2002; Paradis, 2001; Muñoz & Marquardt, 2004). Typological distance between the language pairs and structural complexity also appear to influence these recovery patterns (Arslan & Felser, 2018; Kuzmina et al., 2019). A trend towards better-preserved lexical abilities and fairly intact verbal inflectional forms in the language with more complex morphology has often been reported (Abuom & Bastiaanse, 2012; Kambanaros & Grohmann, 2011). Mixing lexical elements between languages in spontaneous language production is also common in bilingual PWA in order to circumvent lexical retrieval difficulty (Goral et al., 2019). However, little is known about whether and how language mixing patterns are associated with lexical diversity in spontaneous speech production in bilingual aphasia. Egyptian Colloquial Arabic, a language with a rich and complex inflection paradigm, has fewer than a handful of studies (see Elkholay et al., 2025). The present study aims to examine spontaneous speech production in two individuals with Egyptian Colloquial Arabic—English bilingual aphasia, focusing on lexical diversity, syntactic structure, morphological accuracy, and language mixing.

Methods

Two female bilingual PWA who speak Egyptian Colloquial Arabic and English were recruited for this study. They both acquired English at approximately age four and use both languages in their daily lives. P1 was aged 47 at the time of testing, living in Canada and hence in a dominantly English-spoken environment. P2 was aged 48 and based in Cairo, where Arabic is the dominant language. Both PWA presented with a left-hemisphere ischemic stroke (60- and 14-months post-onset). The two PWA were screened for language impairments using selected subtests of the Bilingual Aphasia Test (BAT Arabic Screening version; Guilhem et al., 2013). Speech samples were collected using a semi-spontaneous interview and picture description tasks following Arslan et al. (2016).

Results

The results are presented in *Table 1*. P1's BAT performance for syntactic comprehension was slightly higher in English (0.70 vs. 0.60; h=+0.21), while word and sentence repetition were identical (h=0.00). In terms of lexical retrieval, P1's naming accuracy showed a strong advantage for Arabic over English (1.00 vs. 0.67; h=−1.22, *large effect size). P1's sentence translation and grammaticality judgment ability were overall preserved in Arabic (0.83 vs. 0.50; h=−0.72; 0.75 vs. 0.63; h=−0.26, respectively), while sentence correctness was around chance in both languages (0.50 vs. 0.43; h=+0.14). Word recognition and word translation tasks reported ceiling performance in both languages. Regarding spontaneous speech, P1's production in English showed a higher verb–noun ratio and noun TTR than in her Arabic (0.69 vs. 0.43; h=+0.55; 0.73 vs. 0.52; h=+0.46). P1 presented a larger number of language mixing in Arabic with insertion of English lexical elements than in English.

P2's performance on BAT for syntactic comprehension and sentence repetition was identical. Word repetition showed an English advantage (0.92 vs. 0.75; *h*=+0.49). In terms of lexical retrieval, P2's naming showed a strong English advantage (1.00 vs. 0.83; *h*=+0.83, *large effect). Her sentence translation ability was better preserved in English (0.83 vs. 0.50; *h*=+0.71). In contrast, grammaticality judgment and sentence correctness were better in Arabic (1.00 vs. 0.75; h=−0.83, 1.00 vs. 0.83; h=−0.55). P2's spontaneous speech outcomes showed better performance in English than in Arabic in producing sentences with embeddings (0.67 vs. 0.35; *h*=+0.67), with a higher noun TTR (0.78 vs. 0.62; *h*=+0.36), and higher verb–noun ratio (0.76 vs. 0.61; *h*=+0.34). Her language mixing patterns were present only in Arabic.

Discussion

This study examined using BAT and spontaneous speech production in two Egyptian Colloquial Arabic-English bilingual PWA, focusing on lexical diversity, syntactic structure, morphological patterns and language mixing. P1 displayed a relatively balanced performance across the two languages on the BAT, yet with a clear advantage for naming and sentence translation in Arabic. In contrast, P2 showed a better performance in English, especially in tasks involving naming and sentence translation while her grammaticality judgement responses were more accurate in Arabic. P2's production of complex embedded forms was also higher in English than in Arabic, a difference which is not found in P1. In both participants lexical diversity of nouns was slightly more intact in English over Arabic. Their lexical diversity of verbs appeared to be equally affected in both languages. While only minimal mixing was observed in their English production, both participants mixed English elements while speaking Arabic. An example of mixing pattern P1 produced is given in (1).

- el-walad waʔef ʿala ***chair*** w okht betsaʿdo beyākhod ***cookie*** men el-***cupboard*** mesh ʿarfa esmo eh bel-ʿarabi

the-boy standing on *chair* and sister she-help-him he-take *cookie* from the-*cupboard* not know.F name-its what in-the-Arabic

The data presented here largely converge with Goral et al. (2019), who observed that PWA often mix lexical elements from the non-target language to compensate for lexical retrieval difficulties. This explanation aligns well with the mixing patterns reported in the present study. For both P1 and P2, noun lexical diversity was lower in Arabic than in English, and both participants were observed to more frequently insert English lexical items when speaking Arabic, likely reflecting greater accessibility of English lexical representations. A disparity however is that P1 was found to have better naming ability in Arabic than in English, which might be related to task effects, that is, confrontational naming and lexical retrieval during connected speech may be differentially affected across languages in this bilingual PWA.

References

Abuom, T.O., & Bastiaanse, R. (2012). Characteristics of Swahili–English bilingual agrammatic spontaneous speech and the consequences for understanding agrammatic aphasia. *Journal of neurolinguistics*, *25*(4), 276-293.

Arslan, S., & Felser, C. (2018). Comprehension of *wh*-questions in Turkish–German bilinguals with aphasia: a dual-case study. *Clinical linguistics & phonetics*, *32*(7), 640-660.

Arslan, S., Bamyacı, E., & Bastiaanse, R. (2016). A characterization of verb use in Turkish agrammatic narrative speech. *Clinical Linguistics & Phonetics*, *30*(6), 449-469.
Elkholay, A., Arslan, S., & Jonkers, R. (2025). Lexical Retrieval in Diglossic Aphasia: Naming Accuracy for Cognates and Non-Cognates in Modern Standard Arabic and Egyptian Colloquial Arabic. Paper presented at *the Science of Aphasia Conference: Stem-, Spraak- en Taalpathologie* (Vol. 30, pp. 141-143).
Fabbro, F. (2001). The bilingual brain: Bilingual aphasia. *Brain and language*, *79*(2), 201-210.
Goral, M., Levy, E. S., & Obler, L. K. (2002). Neurolinguistic aspects of bilingualism. *International Journal of Bilingualism*, *6*(4), 411-440.
Goral, M., Norvik, M., & Jensen, B.U. (2019). Variation in language mixing in multilingual aphasia. *Clinical linguistics & phonetics*, *33*(10-11), 915-929.
Guilhem, V., Gomes, S.,…& Köpke, B. (2013). The BAT screening. *Revue de Rééducation Orthophonique*, 253, 121–142.
Kambanaros, M., & Grohmann, K. K. (2011). Profiling performance in L1 and L2 observed in Greek–English bilingual aphasia using the Bilingual Aphasia Test: a case study from Cyprus. *Clinical linguistics & phonetics*, *25*(6-7), 513-529.
Kuzmina, E., Goral, M., Norvik, M., & Weekes, B. S. (2019). What influences language impairment in bilingual aphasia? A meta-analytic review. *Frontiers in psychology*, *10*, 445.
Muñoz, M. L., & Marquardt, T. P. (2004). The influence of language context on lexical retrieval in the discourse of bilingual speakers with aphasia. *Journal of Multilingual Communication Disorders*, *2*(1), 1-17.
Paradis, M. (2001). *Bilingual and polyglot aphasia*. Elsevier.

Table 1. Outcome measures for P1 and P2 across BAT and Spontaneous speech measures. Cohen's *h* indicates effect sizes, where bolded comparisons indicate large effect sizes. Positive *h* indicates English > Arabic, negative *h* indicates Arabic > English (we considered as 0.2 small, 0.5 medium, over 0.8 large effect sizes).

Measure	P1 Arabic	P1 English	Statistics Cohen's *h*	P2 Arabic	P2 English	Statistics Cohen's *h*
Bilingual Aphasia Test Outcomes						
Syntactic comprehension	0.60	0.70	*+0.21*	0.90	0.90	*0.00*
Word repetition	0.75	0.75	*0.00*	0.75	0.92	*+0.49*
Sentence repetition	0.67	0.67	*0.00*	0.67	0.67	*0.00*
Naming	**1.00**	**0.67**	***−1.22****	**0.83**	**1.00**	***+0.83****
Commands	0.50	0.67	*+0.35*	0.67	0.67	*0.00*
Word recognition	1.00	1.00	*0.00*	1.00	1.00	*0.00*
Word translation	1.00	1.00	*0.00*	0.90	0.90	*0.00*
Sentence translation	0.83	0.50	*−0.72*	0.50	0.83	*+0.71*
Grammaticality judgment	0.75	0.63	*−0.26*	**1.00**	**0.75**	***−0.83****
Sentence correctness	0.43	0.50	*+0.14*	1.00	0.83	*−0.55*
Spontaneous Speech Outcomes						
Total utterances	27	33		33	27	
Well-formed utterances	13	20		20	12	
Utterances with embeddings	5	7		7	8	
% well-formed utterances	0.48	0.60	*+0.24*	0.60	0.44	*−0.32*
% embeddings	0.38	0.35	*−0.06*	0.35	0.67	*+0.67*
Type–token ratio (TTR)	0.49	0.50	*+0.02*	0.51	0.49	*−0.04*
Number of verbs	37	41		43	39	
Verb TTR	0.62	0.61	*−0.02*	0.63	0.59	*−0.08*
Number of nouns	87	59		71	51	
Noun TTR	0.52	0.73	*+0.46*	0.62	0.78	*+0.36*
Verb–noun ratio	0.43	0.69	*+0.55*	0.61	0.76	*+0.34*
Nr. Code-mixing	17	2		8	0	
Nr. Code-switching	2	4		2	0	

Notes. For word recognition, word translation, sentence translation, grammaticality judgment, and sentence correction, Arabic scores correspond to performance from English to Arabic, and English scores correspond to performance from Arabic to English.

CASE REPORT ON THE TRANSITION FROM PHONETIC ERROR TO CORRECT PRODUCTION IN A CHRONIC PWA WITH SPECTRAL ANALYSIS.

Jean Marie Annoni, Giovanni Berti

INTRODUCTION

Overlapping effects and ambiguities in the identification of acoustic correlates of articulatory properties are widely regarded as among the principal challenges in the analysis of phonetic production recordings (1). In the presence of neurological disorders affecting language, additional factors must also be taken into account. These include difficulties and uncertainties in the temporal and spectral control of both motor planning (apraxia) and execution (dysarthria) of speech; in some cases, these impairments co-occur.

METHOD

A 73 years old male surveyor, affected by chronic post-stroke aphasia (left hemispheric stroke resulting in aphasia and hemiplegia for 10 years), exhibiting signs of probable apraxia of speech, more pronounced than a concomitant mild dysarthria, and presenting with subsequent partial hearing loss, was recorded while performing language tasks involving the identification of synonyms and/or antonyms. A frequent but non-systematic substitution of the speech sound /v/ with /b/ (phonemic paraphasia) was observed.

The subject was encouraged to adopt self-directed training strategies at home aimed at enhancing concentration, with the goal of focusing cognitive resources on the retrieval of the correct motor planning for the phoneme /v/ rather than /b/.

An internet loaded recording software, audacity 3,2,3 (https://www.audacityteam.org) was used for frequency analysis Oscillograms derived from phonetic recordings (expressed in dB/sec) constitute sequences of forms that envelope each individual oscillation. Their corresponding spectra provide a powerful analytical tool for examining the frequency components of these oscillations. Both representations inherently reflect the aforementioned effects and ambiguities.

RESULTS

This training allowed correction of certain specific errors. The figure (top right) presents oscillograms (decibel values in the time domain) corresponding to the incorrect and correct productions of the word "*vermiglio*" (*vermilion*, upper and lower panels, respectively). The corresponding spectral analyses (decibel values in the frequency domain) are shown in the top left. The lower panel displays a comparison of the spectral analyses of five distinct occurrences of the word "vermiglio," produced at different times but under consistent environmental and procedural conditions.

The graphs were rendered comparable through numerical processing designed to normalize decibel values relative to an appropriate reference point within the frequency spectrum. The highest frequency vale was arbitrarily selected. A significant difference emerges, expressible as $g = \Delta dB/\Delta Hz$, to which several boundary conditions may be attributed. This processing highlights distinct modulation patterns among the erroneous productions and the correct production (blue signal). The magnified inset further emphasizes this distinction: following normalization, the correctly produced "vermiglio" (blue) is clearly distinguishable from the

other instances. Although the present study does not quantitatively characterize these modulation differences, current technological capabilities would make such analysis feasible.

DISCUSSION

If words are considered as compositions of phonemes and speech sound, being as physical observables, then, in the presence of a neurological lesion, the interlocutor of an aphasic speaker perceives only the final output. Such transformations serve to align the perceived utterance with the meaning that the speaker most likely intended to convey.

This largely automatic interpretative process may be described in algorithmic terms, where the initial step consists of hypothesizing a phonetic–phonemic transformation, i.e., a transformation of forms described in both the time and frequency domains. Representations in the frequency domain may be directly related to the Strengthening of Connections (in term of Hebbian Theory) required to activate synaptic processes underlying motor planning and execution to activate the synaptic configurations necessary for the relearning of correct phonemic forms (2).

The present study is limited to reporting the observed phenomena. Nevertheless, it highlights the need for further investigation through additional recordings, focusing on:

- the validity of such methodology (analysis of form sequences in whole word spectral analysis)
- the systematic nature of the observed differences;
- correlations between distinct modulation patterns and variations in the configuration of epilaryngeal cavities;
- the systematicity of recall capacity, through self-training approaches aimed at reactivating pre-existing neural configurations that may have become functionally "frozen" following neurological injury.
- The potential impact on the speech educational approach or self education.

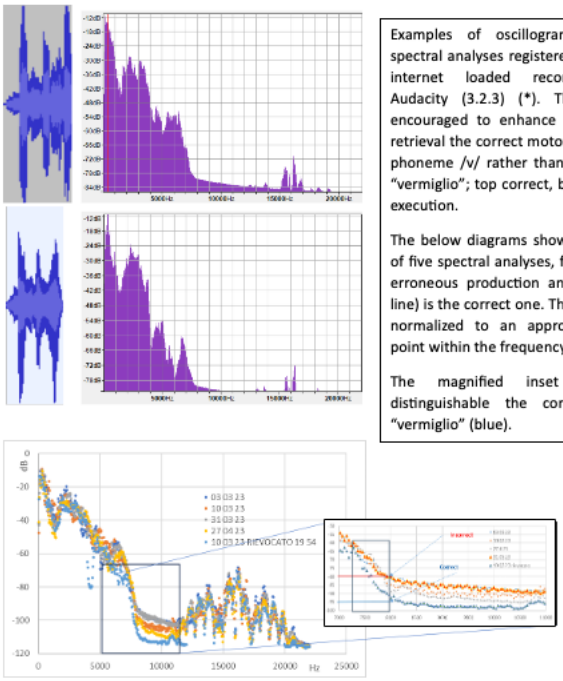
BIBLIOGRAPHY

(1) A. Romano, *L'analisi fonetica strumentale*, in *Manuale di Fonetica*, edited by N. Minissi, M. Rivoira, and A. Romano, Edizioni Dell'Orso, Alessandria (forthcoming).

(2) G. Berti, *Rievocazione Congelata, il fattore di forma e l'energia latente per il riapprendimento intenzionale,* Campano Editore, Pisa, 2022. ISBN 978-8.86528-585-5 (out of print). info.afasicitoscana.pisa@gmail.com.

Acknowledgements

We would like to acknowledge L. I. living with aphasia, who participated to this observation, his family, and Annalisa Rosadoni (student for SLT) who participated to this observation in the frame of her bachelor thesis.



Examples of oscillograms and related spectral analyses registered from using the internet loaded recording software, Audacity (3.2.3) (*). The subject was encouraged to enhance concentration to retrieval the correct motor planning for the phoneme /v/ rather than /b/ in the word "vermiglio"; top correct, bottom erroneous execution.

The below diagrams show the comparison of five spectral analyses, four of thema are erroneous production and one (the blue line) is the correct one. The data have been normalized to an appropriate reference point within the frequency spectrum.

The magnified inset makes easly distinguishable the correctly produced "vermiglio" (blue).

(*) The recordings were collected using Audacity 3.2.3 by A. Rosadoni (student for speech and language therapist, SLT) and Giovanni Berti (lay tutor) as part of the degree thesis entitled *Esperienza in Associazione Afasici Toscana: uso di Smartbrain® Home WSN4Life™ in persone con afasia* (Bachelor's Degree in Speech and Language Therapy, School of Medicine, Department of Surgical, Medical, Molecular and Critical Area Pathology, Academic Year 2022/2023).

The chapter concerning numerical processing and its related discussion was excluded from the present manuscript, as it was considered beyond the scope of the thesis for speech and language therapy training.

Anterior Temporal Lobe stimulation using rTMS for post-stroke aphasia recovery (ATLAS): a study protocol for a single-blinded randomized clinical trial

Sophie Arheix-Parras | Sophia R. Moore | Chris Rorden | Leonardo Bonilha | Rutvik H. Desai

Introduction and aims

Post-stroke aphasia is a major public health concern, as it affects quality of life and community integration. Repetitive Transcranial Magnetic Stimulation (rTMS) has shown promise for improving post-stroke aphasia recovery, with most studies targeting the right Inferior Frontal Gyrus (IFG) using inhibitory rTMS or the left IFG using facilitatory rTMS, putatively to restore interhemispheric balance (1). However, the language system is widespread and complex, involving a large, interconnected network, and focusing on a single cortical target may be insufficient for all aphasia profiles (2).

Lexicosemantic impairments, manifesting as anomia and/or semantic errors, can result from deficits in lexical access or within the semantic system itself. In these processes, the Anterior Temporal Lobe (ATL) is considered as a semantic hub in the hub-and-spoke model (3), while other theories highlight its lexical sensitivity (4). Lesion-symptom mapping implicates the left ATL in Picture Naming (5), and a preliminary single-case study applying inhibitory rTMS targeting the right ATL showed improvement in picture naming for all post-stroke participants with lexicosemantic impairments (6). In this single case study, the stimulation was not associated with speech language therapy. Combining ATL stimulation with Semantic Feature Analysis (SFA), a therapy aiming to strengthen semantic links around a word (7), may further increase language recovery. Additionally, whether inhibition of the right (undamaged) hemisphere or facilitation of the left (damaged) hemisphere produces greater benefit remains unclear.

Furthermore, the mechanisms by which rTMS modulates language recovery remain poorly understood and have not been studied in the context of ATL targeting. One fMRI study (8) targeting the right IFG showed changes in bilateral recruitment during picture naming following inhibitory rTMS, suggesting that stimulation effects extend beyond the targeted region. We can hypothesize that rTMS will modulate not only the targeted ATL but also the ventral semantic stream (2), and fMRI offers a means to examine how each type of rTMS reorganizes intra- and interhemispheric connectivity in stroke recovery.

The ATLAS trial aims to (a) evaluate the effect of rTMS targeting the ATL combined with SFA on anomia in post-stroke aphasia with lexicosemantic impairments, (b) compare the efficacy of right ATL inhibition versus left ATL facilitation, and (c) investigate the neural mechanisms of each stimulation using fMRI.

Methods

This single-blinded sham-controlled randomized clinical trial will recruit 45 participants with chronic (>6 months post-onset) mild-to-moderate aphasia following a single left hemispheric (LH) stroke, from May 2026 to May 2027. Participants are randomized into three parallel groups: (1) inhibitory continuous Theta Burst Stimulation (cTBS) targeting the right ATL; (2) facilitatory intermittent Theta Burst Stimulation (iTBS) targeting the left

ATL; (3) sham stimulation (Figure 1). All groups receive rTMS five times per week for two consecutive weeks (10 sessions in total), each immediately followed by 30 minutes of SFA therapy (7). The stimulation intensity will be set at 80-90% of the resting motor threshold. The sham stimulation will use the same coil but flipped over and will replicate the auditory sensations produced by active stimulation. In addition, we will use electrical stimulation of the scalp in the Sham group to mimic the sensory effects of active stimulation. Target localization uses Brainsight® neuronavigation based on each participant's T1 MRI. The ATL target is defined as 10mm posterior to the tip of the middle anterior temporal lobe.

Assessments will be performed at baseline (S1), immediately post-intervention (S2), and at one-month follow-up (S3). The primary outcome is naming accuracy, assessed using Auditory Naming and Visual Naming Test (9). Secondary outcomes include verb naming (T-DAV 20 translated and adapted for United States participants (10)), the Comprehensive Aphasia Test (CAT) that includes the Aphasia Impact Questionnaire, the Trail Making Test, and a discourse measure (Peer Conflict Resolution task). All naming assessments are administered twice per time point to account for intra-individual variability.

FMRI data will be collected at S1 and S2. Each session includes structural T1, resting-state fMRI, a Picture Naming Task and a Semantic Matching Task (semantically related versus unrelated word pairs contrasted with false-font strings (11)).

Power analysis based on prior rTMS studies in post-stroke aphasia indicates that 15 participants per group (n = 45, accounting for approximately 10% attrition) provides 80% power to detect a significant group effect on Picture Naming accuracy ($\alpha = 0.05$).

Results

Data collection begins May 2026. Results are not yet available. We will present rationale, design, and preliminary results at the time of the conference.

Discussion

The ATLAS trial addresses gaps in the literature of rTMS in post-stroke aphasia. First, it proposes to move beyond IFG stimulation by choosing the brain target based on participants' specific lexicosemantic impairments. This approach was previously tested in Single Case Experimental Designs, but never in a controlled trial (6, 12). Second, by directly comparing inhibition of the right ATL to facilitation of the left ATL, it will allow comparison of the effect of both approaches in regard to the interhemispheric imbalance theory. Third, combining rTMS with SFA is based on the idea that both methods aim to strengthen lexicosemantic networks, and their combination may produce larger effect sizes than either alone. Finally, fMRI analysis of changes in connectivity will give insight into understanding of underlying changes induced by the stimulation.

If results confirm improvement in both active stimulation groups, these findings could support the choice of rTMS target based on the language profile of the participant, easily transferable into clinical practice. RTMS is already recommended for clinical care in France for post-stroke aphasia rehabilitation, and this study could be a first step toward the establishment of clinical guidelines.

References

1. Arheix-Parras et al. A systematic review of rTMS in aphasia rehabilitation.2021.
2. Duffau et al. A re-examination of neural basis of language processing: proposal of a dynamic hodotopical model from data provided by brain stimulation mapping during picture naming.2014.
3. Lambon Ralph et al. The neural and computational bases of semantic cognition.2017.
4. Binder, Desai. The neurobiology of semantic memory.2011.
5. Piai, Eikelboom. Brain Areas Critical for Picture Naming: A Systematic Review and Meta-Analysis of Lesion-Symptom Mapping Studies.2023.
6. Arheix-Parras et al. Improving Lexicosemantic Impairments in Post-Stroke Aphasia Using rTMS Targeting the Right ATL.2026.
7. Quique et al. Acquisition and Generalization Responses in Aphasia Naming Treatment: A Meta-Analysis of Semantic Feature Analysis Outcomes.2019.
8. Harvey et al. Functional Reorganization of Right Prefrontal Cortex Underlies Sustained Naming Improvements in Chronic Aphasia via rTMS.2017.
9. Hamberger et al. Naming in Older Adults: Complementary Auditory and Visual Assessment.2022.
10. Spigarelli et al. TQ-DAV: Action Naming Test with Videos for Quebec French.2026.
11. Wilson et al. An adaptive semantic matching paradigm for reliable and valid language mapping in individuals with aphasia.2018.
12. Arheix-Parras et al. Neuromodulation of the Right Motor Cortex of the Lips With rTMS to Reduce Phonological Impairment and Improve Naming,2024.

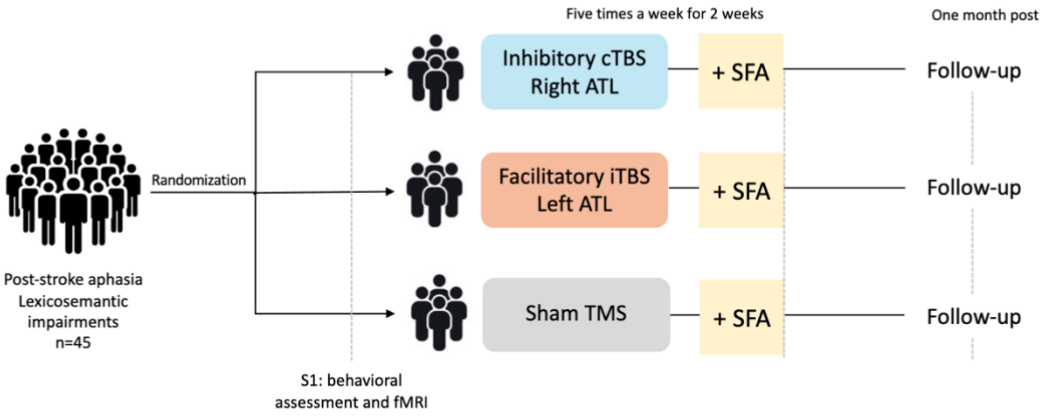


Figure 1: Flow Diagram of the randomized controlled trial with three groups including (1) inhibitory cTBS targeting the right ATL, (2) facilitatory iTBS targeting the left ATL, and (3) sham stimulation. SFA: Semantic Feature Analysis, fMRI: functional Magnetic Resonance Imaging

The Emoji Naming Test: Comparing naming performance in controls and people with aphasia

Josefina Castillo Cerutti | Macarena Martínez-Cuitiño |Martina Mercado | Nicolás Nahuel Romero | Lucía Trevisán | María Lourdes Mátar | Mora Torrente | Dolores Zamora

Introduction and aims

Picture naming is a key task in language assessment which has traditionally relied on black-and-white (B&W) line drawings. The most widely used tool is the Boston Naming Test (BNT; Kaplan et al., 1983), which consists of 60 B&W drawings arranged by increasing lexical difficulty. Despite its popularity, many items present cultural and pictorial limitations. In addition, studies have highlighted the importance of color in visual recognition (Bramão et al., 2011).

Emojis are a contemporary form of visual language created for instant messaging. They represent multiple semantic categories, ranging from emotions to objects. Studies have assessed differences in emoji use based on age, sex or culture (Bai et al., 2019), while others have examined their application in psychological evaluation (Davies et al., 2019) and cognitive tasks (Martínez-Cuitiño et al., 2023).

Given their widespread use and semantic attributes, emojis may provide cues that classical neuropsychological tools lack. Therefore, this study aimed to compare performance using traditional B&W drawings versus emojis in healthy controls and people with aphasia.

Methods

Participants

The sample included 74 native Spanish-speaking adults with no neurological or neuropsychiatric conditions (*M* age=53.9, *SD*=17.75; *M* education=15.7, *SD*=2.82) and nine people with post-stroke anomic aphasia (PWA; *M age*=59.9, *SD*=8.21; *M* education=14.6, *SD*=3.78). Controls and PWA did not differ in education ($F(1,9.12)=0.72$, $p=.417$) or age ($F(1,19.01)=3.09$, $p=.095$).

Materials

Two picture naming tasks were administered: One in B&W and one with emojis. DMDX software (Forster & Forster, 2003) was used to record accuracy rates (ARs) and response times (RTs). Each test consisted of the same 60 stimuli, differing only in picture format. The B&W drawings were extracted from the BNT (BNT-B&W), while emojis (BNT-e) from Emojipedia (https://emojipedia.org/) or designed with DALL-E 3 (OpenAI).

Procedure

Controls were divided into two groups: one was assessed using the BNT-B&W, whereas the other one with the BNT-e. PWA were assessed with both tasks. Each participant was asked to name aloud the stimuli with a single word, as quickly and accurately as possible.

Each trial began with a 400-millisecond (ms) display of a fixation point followed by an 800-ms stimulus presentation. This was followed by a black screen of 4000-ms for controls, and 20000-ms for patients, during which participants could name the stimulus.

ARs and RTs were collected, and all responses were examined using CheckVocal software (Protopapas, 2007).

Analyses

A generalized linear mixed-effects model (GLMM) with a binomial distribution was used to analyze the ARs. RTs were analyzed using a linear mixed-effects model (LMM). In both cases, visual format (BNT-B&W vs. BNT-e) and lexical availability (high/medium vs. low) were assigned as fixed factors, while subjects (ID) and items were included as random factors. The first and second models analyzed controls and PWA, respectively. The third model included controls and PWA.

Results

Considering controls´ performance, the ARs analysis revealed effects of format ($\chi^2(1)=5.05$, $p=.02$), lexical availability ($\chi^2(1)=37.66$, $p<.001$) and their interaction ($\chi^2(1)=9.24$, $p=.002$). Post hoc comparisons indicated that for high/medium-availability items, BNT-B&W was associated with lower odds of correct responses than emojis ($OR=0.47$, $p=.007$). No format differences were observed for low availability items ($OR=0.84$, $p=1.000$). The RTs analysis showed only an effect of lexical availability ($F(1,57.4)=29.57$, $p<.001$), with low-availability stimuli being 375-ms slower than high/medium-availability stimuli.

The analysis considering PWA in ARs indicated effects of format ($\chi^2(1)=13.24$, $p<.001$) and lexical availability ($\chi^2(1)=50.25$, $p<.001$). BNT-B&W was associated with lower odds of correct responses than BNT-e ($OR=0.49$, $p<.001$). High/medium-availability items were associated with higher odds of correct responses than low availability items ($OR=20.2$, $p<.001$). The RTs analysis denoted a significant effect of format ($F(1,647.0)=17.72$, $p<.001$) and of lexical availability ($F(1,64.7)=35.95$, $p<.001$). Responses to BNT-B&W drawings were, on average, 965-ms slower, and responses to low-availability items were slower than responses to high- and medium-availability items.

When controls and PWA were included, only an effect of format was observed for ARs ($\chi^2(1)=23.17$, $p<.001$). BNT-e was associated with higher odds of correct responses ($OR=1.88$, $p<.001$). For RTs, the model showed effects of format ($F(1,1585.6)=57.8$, $p<.001$), group ($F(1,66.4)=659.4$, $p<.001$) and their interaction ($F(1,1501.4)=60.4$, $p<.001$). Post hoc showed no format differences in controls, whereas PWA were slower with BNT-B&W, with a mean difference of 726.53-ms.

Discussion

Emojis demonstrated an equal or better performance compared with traditional B&W drawings in visual naming. They were named more correctly by both groups. In controls, this effect was restricted to high/medium-availability items. In PWA, however, the advantage was more pronounced, being named not only more accurately but also eliciting faster RTs. These results suggest that emojis may offer a more ecological and semantically enriched alternative

to classic naming tools. Future studies should expand clinical sample sizes, include diverse aphasia profiles and examine performance across different neurological conditions.

References

Bai, Q., Dan, Q., Mu, Z., & Yang, M. (2019). A systematic review of emoji: Current research and future perspectives. *Frontiers in psychology, 10*, 476737. https://doi.org/10.3389/fpsyg.2019.02221

Bramão, I., Reis, A., Petersson, K. M., & Faísca, L. (2011). The role of color information on object recognition: A review and meta-analysis. *Acta psychologica, 138*(1), 244-253. https://doi.org/10.1016/j.actpsy.2011.06.010

Davies, J., McKenna, M., Denner, K., Bayley, J., & Morgan, M. (2024). The emoji current mood and experience scale: the development and initial validation of an ultra-brief, literacy independent measure of psychological health. *Journal of mental health, 33*(2), 218–226. https://doi.org/10.1080/09638237.2022.2069694

Kaplan, E., Goodglass, H., & Weintraub, S. (1983). Boston Naming Test. Lea & Febiger.

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers, 35*, 116–124. https://doi.org/10.3758/BF03195503

Protopapas, A. (2007). Check Vocal: A program to facilitate checking the accuracy and response time of vocal responses from DMDX. *Behavior Research Methods, 39*(4), 859–862. https://doi.org/10.3758/BF03192979

How interrelated are semantic and episodic memory retrieval? An MEG study of healthy individuals and people after left temporal lobe epilepsy surgery

Irina Chupina | Britta U. Westner | Marc Petrus Hendriks | Roy P.C. Kessels | Joanna Sierpowska | Vitória Piai

Introduction and aims

Although the separation of episodic (verbal) and semantic memory systems has proven useful in clinical practice for categorising differently presenting memory impairments, empirical evidence suggests that the two memory types are closely intertwined (Greenberg and Verfaellie, 2010). This observation might be particularly relevant for individuals with left temporal lobe epilepsy (TLE), as their performance on semantic and episodic memory tests has been shown to correlate at the individual level (Barrett Jones et al., 2022). Furthermore, following neural surgery conducted to control the seizures, TLE patients frequently experience verbal memory as well as language problems (e.g., Helmstaedter, 2013), including word-finding difficulties (Ives-Delipery & Butler, 2012). Despite these findings, the extent of the interrelation between episodic and semantic memory remains poorly understood. To gain a more nuanced understanding of this interrelation, we aimed to compare retrieval from episodic and semantic memory at the behavioural and neural levels, as well as characterise the changes these cognitive processes might exhibit after left TLE surgery. We asked: (a) whether the behavioural performances on episodic and semantic memory tasks are correlated (accuracy, response times); (b) to what extent electrophysiological retrieval signatures are similar across the tasks; and (c) whether history of left TLE and epilepsy surgery leads to co-dependent changes reflected in the electrophysiological retrieval signatures and, if so, how these changes relate to behavioural retrieval efficiency.

Methods

We recruited cognitively healthy participants (N = 22) and epilepsy patients at least 1 year after a standard left temporal lobe resection in which the basal temporal lobe area (crucial for semantic memory and language) and (parts of) the hippocampus (crucial for episodic memory) were removed (N = 10). Participants underwent structural MRI and performed a language (semantic memory) and an episodic verbal memory task while magnetoencephalography (MEG) was recorded. In the language task, participants named pictures following semantically constraining and unconstraining sentences, the former eliciting word planning in the pre-picture interval (Piai et al., 2014). This task also served as the encoding phase of the memory task, in which participants read a word and pressed a button to indicate whether they had encountered the word in the naming task (old) or not (new) (Perrone-Bertolotti et al., 2015). We obtained electrophysiological signatures of retrieval for language and memory by contrasting the conditions within tasks, where retrieval of information occurs in one (constrained, old) but not the other (unconstrained, new) condition. For data analysis, we first correlated error rates in all participants to assess the interrelation of retrieval accuracy across the tasks. Then, we conducted within-task time-frequency analysis within each participant group to characterise the retrieval signatures of the semantic and episodic memory tasks. Next, we plan to correlate mean reaction times (RTs; speech onset vs button-press times) within-participant to further assess the interrelation of time-based retrieval efficiency across the tasks. Moreover, we will reconstruct the neural sources for the retrieval activity in the time-frequency representations within-task.

Additionally, we will use mutual information in the MEG signal to compare the brain-wide signatures across tasks in healthy and patient groups (Westner et al., 2026). Finally, we will descriptively report potential changes in the retrieval signatures demonstrated by the patient group. Any co-dependent changes will be interpreted as evidence of the interrelatedness of retrieval from episodic and semantic memory.

Results

Preliminary behavioural results demonstrate a moderate to strong correlation (rho = 0.7, $p < 0.001$) between error rates in language and memory tasks across all participants (controls: N = 15; patients: N = 7). In the TLE group, one individual experienced a more pronounced difficulty in the language task, while another experienced a more pronounced difficulty in the memory task. This highlights the potential dissociation in patient profiles, with either episodic or semantic memory being more impaired. In line with previous studies (Piai et al., 2014; Westner et al., 2026), the control group (N = 14) exhibited power decreases in the alpha-beta frequency bands (8 – 30 Hz) in both tasks. These electrophysiological signatures are associated with information retrieval. While sensor-level topographies indicated power decreases in the left hemisphere in both tasks, these appeared to be more temporal in the language task and more parietal-occipital in the memory task. We expect the source reconstruction to reveal partially non-overlapping sources, with some overlap in left temporo-parietal brain areas (Westner et al., 2026). In the patient group, we anticipate intra- or inter-hemispheric changes in retrieval from semantic and episodic memory. We hypothesise that functional language and memory signatures will be similar across tasks, but different from our control population.

Discussion

Our ongoing study investigates the behavioural performance as well as electrophysiological characteristics of information retrieval from semantic and episodic memory in healthy individuals and individuals after left TLE surgery. The preliminary results support the notion that language and memory are closely interrelated, yet are also distinct. First of all, although the behavioural performance between language and memory retrieval correlated in all participants, some patients may be relatively more impaired in one domain than the other. In daily life this may lead to either memory or language complaints. Furthermore, while the electrophysiological retrieval signatures in the control group showed similar spectral profiles (decreases in the alpha-beta frequency range), they might be supported by (partly) non-overlapping neural circuits. Ongoing analyses are expected to further clarify these patterns.

References

Barrett Jones, S., A. Miller, L., Kleitman, S., Nikpour, A., & Lah, S. (2022). Semantic and episodic memory in adults with temporal lobe epilepsy. *Applied Neuropsychology: Adult*, *29*(6), 1352–1361.

Greenberg, D. L., & Verfaellie, M. (2010). Interdependence of episodic and semantic memory: Evidence from neuropsychology. *Journal of the International Neuropsychological Society*, *16*(5), 748–753.

Helmstaedter, C. (2013). Cognitive outcomes of different surgical approaches in temporal lobe epilepsy. *Epileptic Disorders*, *15*(3), 221–239.

Ives-Deliperi, V. L., & Butler, J. T. (2012). Naming outcomes of anterior temporal lobectomy in epilepsy patients: A systematic review of the literature. *Epilepsy & Behavior*, *24*(2), 194–198.
Perrone-Bertolotti, M., Girard, C., Cousin, E., Vidal, J. R., Pichat, C., Kahane, P., & Baciu, M. (2015). NEREC, an effective brain mapping protocol for combined language and long-term memory functions. *Epilepsy & Behavior*, *53*, 140–148.
Piai, V., Roelofs, A., & Maris, E. (2014). Oscillatory brain responses in spoken word production reflect lexical frequency and sentential constraint. *Neuropsychologia*, *53*, 146–156.
Westner, B. U., Luo, Y., & Piai, V. (2026). Shared and distinct oscillatory fingerprints underlying episodic memory and word retrieval. *bioRxiv*, 2026-04.

Associative Strength and Semantic Proximity: A Cross-Modal Investigation of Lexical-Semantic Network Organisation

Marie Couvreu | Marina Laganaro

Introduction

When a speaker attempts to produce a specific word, activation spreads within the lexical-semantic network, activating related words and features until the target word receives enough activation to be retrieved (Collins & Loftus, 1975). The organisation of this network may influence word access/retrieval, leading to a temporary access difficulty known as the Tip of the Tongue (ToT) phenomenon, observed across the lifespan. More persistent word-finding difficulties are also observed in clinical populations (i.e. primary progressive aphasia or post-stroke aphasia) and may be associated with a deregulation of the organisation of the lexical-semantic network (Lampe et al., 2021; Nickels et al., 2022).

The continuous free word association task is considered one of the most effective approaches to investigate the structure of the lexical-semantic network, providing measures of associative strength and semantic richness (De Deyne et al., 2013; Nelson et al., 2004). Previous works using this task with younger (Couvreu & Laganaro, 2026) and older neurotypical adults (Couvreu & Laganaro, in preparation) revealed that this organisation differs between ToT and "accessible" states, in terms of semantic richness and association strength, with more heterogeneous responses observed in ToT states. Building on these findings, the aim of the present study is to validate and complete the data collected with the word association task by evaluating its convergence with the judgment of semantic proximity in the written modality between a target word and three associates with different strength taken from the previous study (Couvreu & Laganaro, 2026).

More specifically, we compare the organisation of the lexical-semantic network across oral associative production and semantic judgement on written words, examining whether associates produced in an oral task are classified similarly in a semantic proximity judgment task we developed using the same stimuli. This work forms part of a broader project to develop a French lexical-semantic network database and tasks that can assess its changes in typical and atypical ageing.

Methods

80 French-speaking neurotypical adults participated in the study, divided into two groups (mean age =23y, 15 men), with no history of language impairment or neurological disease. Participants performed a task evaluating semantic proximity between a target (presented in two formats: word or picture) and 3 associates (see Fig 1).

The 70 low-frequency target words and their associates were selected from a previous study investigating the lexical-semantic network in the oral production (Couvreu & Laganaro, 2026), using a continuous free word association task. Associates were selected based on their frequency of occurrence and order of production across participants. A principal component analysis was then used to derive proximity scores, resulting in three levels: C90-100 (Associate1, closest), C70-80 (Associate 2), and C50-60 (Associate 3, more distant).

Results

Preliminary analyses on 20 participants revealed that Associate1 is strongly associated with "short" responses ($\beta = -2.68$, SE = 0.12, $z = -21.97$, $p < .001$), while Associate2 and Associate3 were both associated with further categories ($\beta = 2.72$, SE = 0.12, $z = 22.09$, $p<.001$) and Associate3 ($\beta = 2.68$, SE = 0.12, $z = 21.97$, $p<.001$). However, no significant difference was found between Associate2 and Associate3, suggesting an overlap between these two distance levels. This pattern was consistent whether the target was presented as a written word or a picture, with no significant interaction.

Conclusion

The aim of this study was to compare the organisation of the lexical-semantic network across oral association and written judgment modalities by testing whether associates produced in a continuous free word association task are classified similarly in a semantic proximity judgment task. Results showed that the strongest associate link to the target word/picture tends to be consistently classified in the same position across modalities. These preliminary findings highlight the relative stability of lexical-semantic network organisation of the target word for the first associate across participants, alongside greater heterogeneity for the second and third position (De Deyne et al., 2012; Nelson et al., 2000) across modalities. It also highlights the value of combining tasks across modalities to better capture the lexical-semantic network organisation and variability in associative strength.

Future studies will examine associative strength in older adults and extend this work to clinical populations. A better understanding of lexical-semantic organisation across modalities may provide valuable insight into early lexical-semantic difficulties in some neurodegenerative diseases or post-stroke aphasia in clinical settings.

References :

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological review, 82*(6), 407. doi.org/10.1037/0033-295X.82.6.407

Couvreu, M & Laganaro, M. (2026). *The speaker's lexical-semantic network in the tip of the tongue state.* ResearchSquare. doi.org/10.21203/rs.3.rs-8970418/v1

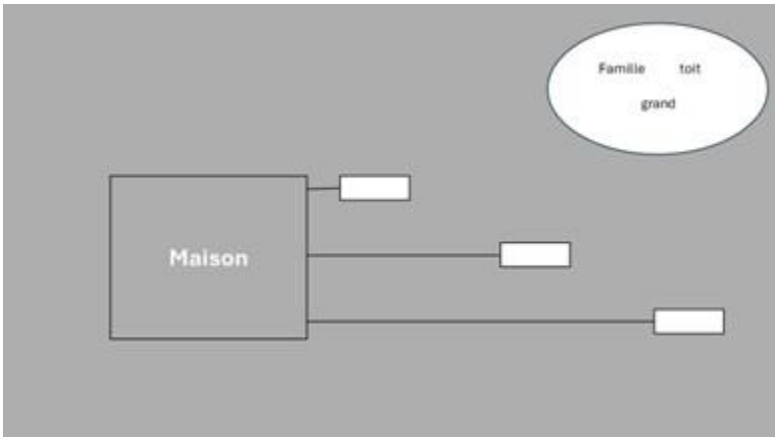
Couvreu, M., & Laganaro, M. (in preparation). Lexical-semantic organization and Tip-of-the-Tongue states in typical ageing.

De Deyne, S., Navarro, D. J., & Storms, G. (2013). Better explanations of lexical and semantic cognition using networks derived from continued rather than single-word associations. *Behavior research methods*, *45*(2), 480-498. doi.org/10.3758/s13428-012-0260-7

Lampe, L. F., Hameau, S., Fieder, N., & Nickels, L. (2021). Effects of semantic variables on word production in aphasia. *Cortex*, 141, 363–402. https://doi.org/10.1016/j.cortex.2021.02.020

Nelson, D. L., McEvoy, C. L., & Dennis, S. (2000). What is free association and what does it measure? *Memory & cognition*, *28*(6), 887-899. doi.org/10.3758/BF03209337

Nickels, L., Lampe, L. F., Mason, C., & Hameau, S. (2022). Investigating the influence of semantic factors on word retrieval: Reservations, results and recommendations. *Cognitive Neuropsychology, 39*(3-4), 113-154.


Famille
toit
grand
Maison

Temporal Indicators of Bilinguals' Early Cognitive Decline in Verbal Fluency Tasks

Ana María Estrada Pineda | Silvia Martínez Ferreiro

Introduction and aims

MCI currently affects 23.7% of older adults worldwide (Salari et al., 2025). However, despite its prevalence, it remains largely undiagnosed, with over half of cases in Spain and up to 90% in the U.S. going undetected (Mattke et al., 2023; Sociedad Española de Neurología, 2024). This presents a significant gap in global cognitive care today.

Performace in Verbal Fluency Tasks (VFTs) — listing words under a specific criterion within a time limit — could indicate cognitive deterioration in monolingual populations with MCI (Cintoli et al., 2024; Demetriou & Holtzer, 2017; McDonnell et al., 2020; Quaranta et al., 2019) Individuals with MCI exhibit delayed initial retrieval (Demetriou & Holtzer, 2017) and shorter mean latencies compared to healthy controls (Tröger et al., 2022).

However, whether the results of monolingual studies apply to bilinguals remains unclear. To the best of our knowledge, there are no studies exploring temporal parameters of VFT performance in bilinguals with cognitive decline. Nonetheless, evidence from healthy bilinguals suggests that their performance is fundamentally different to that of monolinguals, likely due to differences in executive functioning and cognitive reserve (Friesen et al., 2015; Luo et al., 2010; Patra et al., 2020; Sandoval et al., 2010). This study aims to explore whether temporal characteristics of verbal fluency performance (1st Response Time, Mean Subsequent Response Time, Intercept, and Slope) can serve as early markers of cognitive decline in older Galician-Spanish bilinguals.

Methods

We used verbal fluency and cognitive assessment data from 53 bilingual participants (39 Control, 14 MCI) aged 65+ from the LINDEGALES (ED431F 2024/018) and SECONDSGALES (PID2023-151673OB-I00) projects. Participants were evaluated, in their preferred language, across two verbal fluency tasks: *animals* for semantic fluency (SF), and the *letter* <*s*> for letter fluency (LF). We used Audacity to isolate relevant audio segments and PRAAT to transcribe and time-stamp responses. Groups were determined using age- and education-adjusted MoCA scores (Ojeda et al., 2016). To analyze temporal measures, we divided correct responses into 12 five-second bins, and fit a logarithmic curve to see how performance changed over time (Patra et al. 2020).

Results

Preliminary results show that participants with MCI were significantly faster than controls to start word production in the *letter* <*s*> task ($\beta = -0.61$, 95% CI = [-0.82 – -0.41], $p < .001$), but not in the *animals* task. Healthy participants showed significantly longer mean subsequent response times in both tasks ([animals] $\beta = 22.63$, 95% CI = [21.47 – 23.80], $p < .001$; [letter <s>] $\beta = 2.79$, 95% CI = [2.51 – 3.07], $p < .001$).

The MCI group showed significantly lower intercepts than controls ($\beta = -1.19$, 95% CI = [-2.21 – -0.18], p 0 .023). Healthy controls exhibited a significantly steeper slope in the

semantic task (β = -0.75, 95% CI = [-0.97 – -0.54], p <0.001). This pattern held in the letter task with controls displaying significantly higher intercepts (β = 2.67, 95% CI = [2.29 – 3.05], p <0.001) and a significantly steeper slope (β = -0.32, 95% CI = [-0.43 – -0.20], p <0.001) compared to MCI participants.

Discussion

We expected participants with cognitive decline to show longer first-response times since delayed initiation indicates compromised word retrieval during the earlier, semi-automatic stages of VFTs (Demetriou & Holtzer, 2017). Surprisingly, the MCI group was faster than controls at initiating exemplar generation in the LF task. MCI participants also showed lower intercepts than controls across tasks, and had a flatter slope that showed a strong tendency towards significance in the letter task.

Faster initiation and a flatter slope could indicate impaired ability to strategically plan, organize, and monitor responses. This would result in an initially faster, but overall, less efficient output. Lower intercepts are thought to reflect the fewer lexical resources available at the start of the task while flatter slopes indicate a more effortful use and monitoring of those resources over time (Friesen et al., 2015; Luo et al., 2010; Patra et al., 2020).

These findings coincide with our previous preliminary explorations of clustering and switching patterns where MCI participants of the same data set produced fewer clusters and fewer switches than controls in the *letter <s>* task (Estrada Pineda et al., 2025). Jointly, temporal and qualitative results point to an unsystematic, effortful word retrieval process in a task that poses significant executive control demands. These findings may indicate a distinct early manifestation of executive dysfunction in bilinguals with cognitive decline.

Acknowledgements

SECONDSGALES (PID2023-151673OB-I00) was funded by MICIU/AEI/10.13039/501100011033 and by “ERDF/EU" and LINDEGALES (ED431F 2024/018) by Xunta de Galicia. SMF also acknowledges support of RYC2020-028927-1, MICIU/AEI/10.13039/501100011033 and ESF investing in your future. The Gerontology and Geriatrics Research Group (UDC) aided in recruitment and data collection.

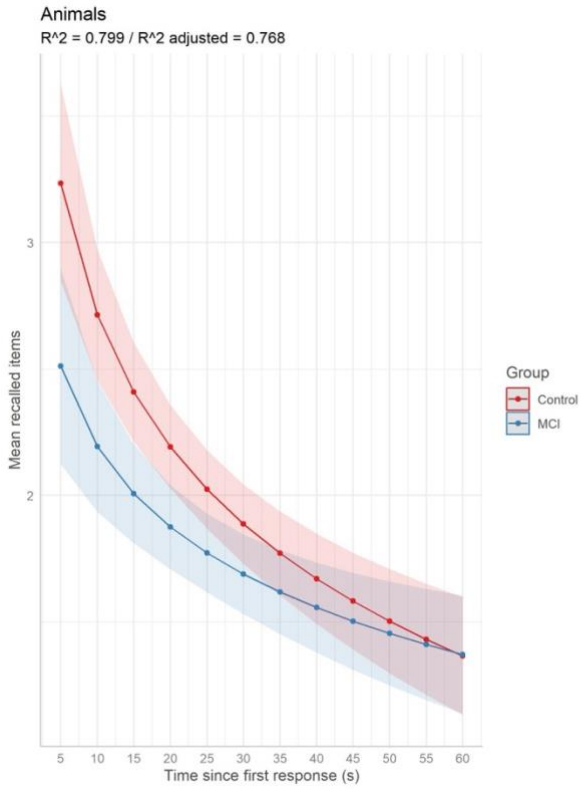
References

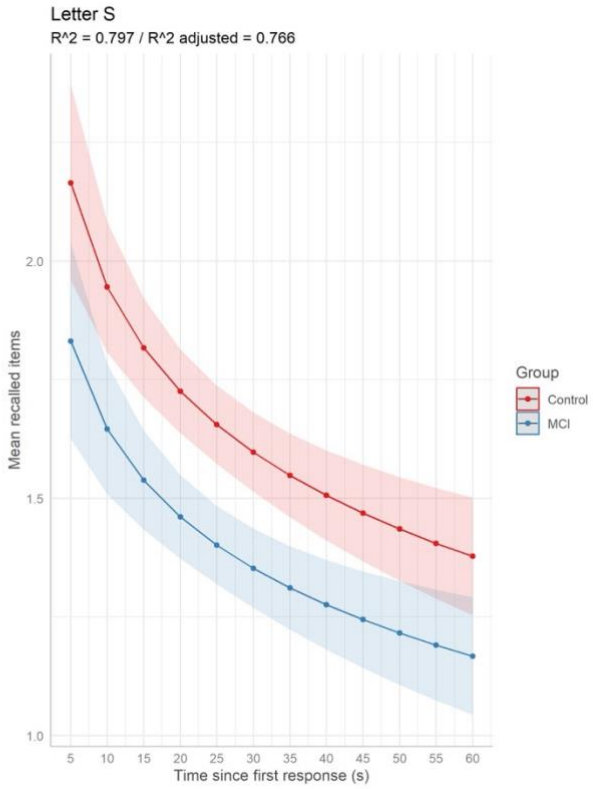
Cintoli, S., Favilli, L., Morganti, R., Siciliano, G., Ceravolo, R., & Tognoni, G. (2024). Verbal fluency patterns associated with the amnestic conversion from mild cognitive impairment to dementia. *Scientific Reports*, *14*(1), 2029. https://doi.org/10.1038/s41598-024-52562-x

Demetriou, E., & Holtzer, R. (2017). Mild Cognitive Impairments Moderate the Effect of Time on Verbal Fluency Performance. *Journal of the International Neuropsychological Society : JINS*, *23*(1), 44–55. https://doi.org/10.1017/S1355617716000825

Estrada Pineda, A. M., Millán Calenti, J. C., Rießler, M., & Martínez Ferreiro, S. (2025, October). *Clustering and Switching in Verbal Fluency Tasks: Potential Indicators of Cognitive Degradation in Elderly Galician-Spanish Bilinguals [Poster]*. 63rd Annual Meeting of the Academy of Aphasia, San Diego, USA.

Friesen, D. C., Luo, L., Luk, G., & Bialystok, E. (2015). Proficiency and control in verbal fluency performance across the lifespan for monolinguals and bilinguals. *Language, Cognition and Neuroscience*, *30*(3), 238–250. https://doi.org/10.1080/23273798.2014.918630
Luo, L., Luk, G., & Bialystok, E. (2010). Effect of language proficiency and executive control on verbal fluency performance in bilinguals. *Cognition*, *114*(1), 29–41. https://doi.org/10.1016/j.cognition.2009.08.014
McDonnell, M., Dill, L., Panos, S., Amano, S., Brown, W., Giurgius, S., Small, G., & Miller, K. (2020). Verbal fluency as a screening tool for mild cognitive impairment. *International Psychogeriatrics, Issue Theme: Enhanced Assessment Instruments for Older Adults with Major Neurocognitive Disorders*, *32*(9), 1055–1062. https://doi.org/10.1017/S1041610219000644
Ojeda, N., Del Pino, R., Ibarretxe-Bilbao, N., Schretlen, D. J., & Pena, J. (2016). Montreal Cognitive Assessment Test: Normalization and standardization for Spanish population. *Revista De Neurologia*, *63*(11), 488–496. https://doi.org/https://doi.org/10.33588/rn.6311.2016241
Patra, A., Bose, A., & Marinis, T. (2020). Performance difference in verbal fluency in bilingual and monolingual speakers. *Bilingualism: Language and Cognition*, *23*(1), 204–218. https://doi.org/10.1017/S1366728918001098
Quaranta, D., Piccininni, C., Caprara, A., Malandrino, A., Gainotti, G., & Marra, C. (2019). Semantic Relations in a Categorical Verbal Fluency Test: An Exploratory Investigation in Mild Cognitive Impairment. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.02797
Salari, N., Lotfi, F., Abdolmaleki, A., Heidarian, P., Rasoulpoor, S., Fazeli, J., Najafi, H., & Mohammadi, M. (2025). The global prevalence of mild cognitive impairment in geriatric population with emphasis on influential factors: A systematic review and meta-analysis. *BMC Geriatrics*, *25*(1), 313. https://doi.org/10.1186/s12877-025-05967-w
Sandoval, T. C., Gollan, T. H., Ferreira, V. S., & Salmon, D. P. (2010). What causes the bilingual disadvantage in verbal fluency? The dual-task analogy. *Bilingualism: Language and Cognition*, *13*(2), 231–252. https://doi.org/10.1017/S1366728909990514
Sociedad Española de Neurología (SEN). (2024, September 19). *Las demencias ya suponen el 8% del total de defunciones que se producen cada año en España*. Sociedad Española de Neurología (SEN). https://www.sen.es/saladeprensa/pdf/Link451.pdf
Tröger, J., Lindsay, H., Mina, M., Linz, N., Klöppel, S., Kray, J., & Peter, J. (2022). Patients with amnestic MCI Fail to Adapt Executive Control When Repeatedly Tested with Semantic Verbal Fluency Tasks. *Journal of the International Neuropsychological Society*, *28*(6), 620–627. https://doi.org/10.1017/S1355617721000849


Animals
R^2 = 0.799 / R^2 adjusted = 0.768
Mean recalled items
3
2
5 10 15 20 25 30 35 40 45 50 55 60
Time since first response (s)
Group
Control
MCI


Letter S
R^2 = 0.797 / R^2 adjusted = 0.766
Mean recalled items
2.0
1.5
1.0
5 10 15 20 25 30 35 40 45 50 55 60
Time since first response (s)
Group
Control
MCI

A study of the impact of speech rhythm and semantic context on comprehension and production in Parkinson's disease

Anaëlle Filisetti | Vitoria Piai | Antoine Petit | Laurent Ott | Anahita Basirat

Introduction and aims

Although language comprehension and production are often studied separately, several studies suggest that they mutually influence each other. In particular, the sentence's semantic context in comprehension appears to affect word production (e.g., Piai et al., 2014). To investigate this context effect, the Concise Language Paradigm (ClaP, Roos et al., 2024) was proposed as a framework to study comprehension and production in the same experimental trial. Previous findings indicate that increasing the semantic constraint of a sentence context facilitates comprehension and production in healthy participants and in participants with language disorders (Huck et al., 2017; Roos et al., 2024). This effect can also be used in speech-language therapy practice for patients with naming disorders. Beyond semantic context, speech prosody also carries important information, and speech rhythm appears to be essential for speech intelligibility and for prediction (Aubanel & Schwartz, 2020; Huettig, 2015). Therefore, since speech rhythm impacts perception and prediction, investigating whether it can influence language production seems relevant. In addition, as Parkinson's disease (PD) involves both rhythm impairments and lexical access difficulties, it could be valuable to investigate the impact of rhythm and semantic context on language production in this population. To do so in our planned studies, we will use the ClaP paradigm where participants listen to a sentence missing its final word, and then name an image representing that word.

Methods:

Given that no databank exists in French with materials systematically manipulating sentence constraint, in the first study, we created a corpus of 269 sentences varying in their degree of constraint. Sentences are organised into pairs, each comprising a constrained sentence and an unconstrained sentence, both linked to the same target word. All target words exist in the Multipic database (Duñabeitia et al., 2018), allowing the sentences to be used in paradigms such as the CLaP. Sentence constraint was calculated using cloze probability based on a sentence completion task involving 160 participants.

Results

Sentences' mean cloze probability, which reflects the probability to produce the target word to complete the sentence, was 0.77 (SD=0.21) for constrained sentences and 0.04 (SD=0.07) for unconstrained sentences. The mean modal completion, which represents the probability of the most frequent response, was 0.78 (SD=0.19) for constrained sentences and 0.22 (SD=0.15) for unconstrained sentences *(Figure 1)*.

Figure 1: Distribution of the target cloze probability (A) and modal response cloze probability (B) for constrained (C) and unconstrained (U) sentences.

Discussion

For our future study, we plan to recruit 30 native French speakers with PD and a matched control group. To ensure a high degree of constraint for constrained sentences and a low degree of constraint for unconstrained sentences, we will select 81 sentence pairs from our established databank where constrained sentences had a target cloze probability above 0.70 and unconstrained sentences had a modal response below 0.30. In our study, the sentence rhythm will also vary, contrasting natural rhythm with isochronous rhythm, in order to also explore the impact of rhythm on word production. To quantify the impact of rhythm and semantic context on language production, we will compare naming latency and accuracy between the groups across conditions. While designed to be tested on PD participants, this databank of French language materials could also be used in other pathologies such as aphasia, both for experimental studies and for speech-language therapy.

References:

Aubanel, V., & Schwartz, J.-L. (2020). *The role of isochrony in speech perception in noise | Scientific Reports*. https://www.nature.com/articles/s41598-020-76594-1

Duñabeitia, J. A., Crepaldi, D., Meyer, A. S., New, B., Pliatsikas, C., Smolka, E., & Brysbaert, M. (2018). MultiPic : A standardized set of 750 drawings with norms for six European languages. *Quarterly Journal of Experimental Psychology*, *71*(4), 808-816. https://doi.org/10.1080/17470218.2017.1310261

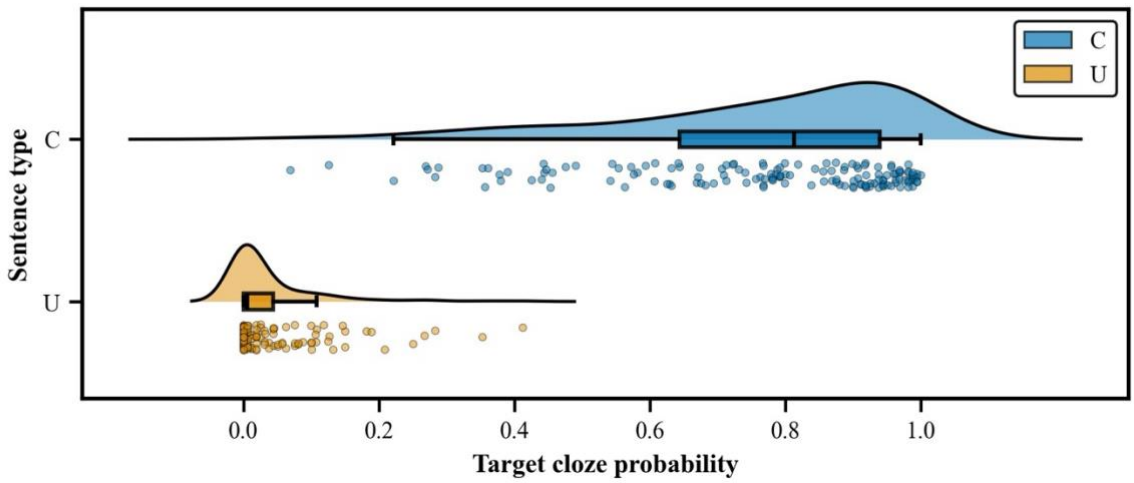
Huck, A., Thompson, R. L., Cruice, M., & Marshall, J. (2017). Effects of word frequency and contextual predictability on sentence reading in aphasia : An eye movement analysis. *Aphasiology*, *31*(11), 1307-1332. https://doi.org/10.1080/02687038.2017.1278741

Huettig, F. (2015). Four central questions about prediction in language processing. *Brain Research, Predictive and Attentive Processing in Perception and Action*, *1626*, 118-135. https://doi.org/10.1016/j.brainres.2015.02.014

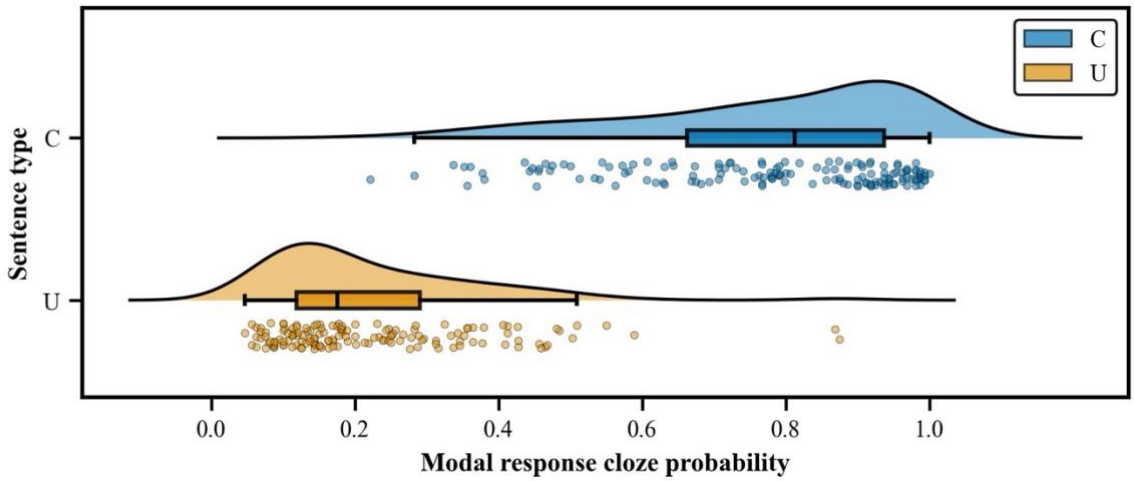
Piai, V., Roelofs, A., & Maris, E. (2014). Oscillatory brain responses in spoken word production reflect lexical frequency and sentential constraint. *Neuropsychologia*, *53*, 146-156. https://doi.org/10.1016/j.neuropsychologia.2013.11.014

Roos, N. M., Chauvet, J., & Piai, V. (2024). The Concise Language Paradigm (CLaP), a framework for studying the intersection of comprehension and production : Electrophysiological properties. *Brain Structure and Function*, *229*(9), 2097-2113. https://doi.org/10.1007/s00429-024-02801-8

A


C
U
Sentence type
C
U
0.0
0.2
0.4
0.6
0.8
1.0
Target cloze probability

B


C
U
Sentence type
C
U
0.0
0.2
0.4
0.6
0.8
1.0
Modal response cloze probability

Tradeoff Patterns in Connected Language Production in Post-Stroke Aphasia Following Speech-Language Therapy

Elizabeth Galletta, Nicole Gerlak and Mira Goral

Introduction & Aim

Connected speech is a critical component of functional communication and is frequently impaired in individuals with post-stroke aphasia (Azios et al., 2022). Measures such as Correct Information Units (CIUs) capture lexical informativeness, while AS-units reflect sentence-level complexity and completeness (Dipper et al., 2021). Improvements in connected speech may occur across multiple dimensions, including lexical informativeness and sentence-level complexity (Stark et al., 2023). Examining these measures together allows for the identification of potential tradeoffs between efficiency and sentence complexity in language production (Altman & Goral, 2025). However, limited research has examined these relationships in various elicitation tasks as measures of treatment efficacy.

The present study aims to examine changes in connected speech following speech-language therapy (SLT), and explore potential tradeoffs between lexical efficiency/ informativeness (CIU/TVU ratio) and sentence-level complexity (AS-units) across two elicitation tasks.

Methods

Data were collected from two individuals with post-stroke aphasia (P1: 58-year-old male, 2 years post left MCA infarct; P2: 53-year-old male, 3 years post left MCA infarct). Both had previously received SLT and presented with moderate aphasia severity based on Aphasia Quotients from the Western Aphasia Battery–Revised (WAB-R) obtained at baseline (P1: 69.2; P2: 64.4). Written informed consent was obtained in accordance with institutional review board approval. The study utilized two parallel single-case multiple baseline experimental designs. Repeated measures were collected at pre-treatment, post-treatment, and follow-up across two 5-week intervention blocks targeting functional communication. For each participant, one treatment block included SLT alone, and the other included SLT combined with aerobic exercise (AE), with the order of conditions counterbalanced across participants as part of the original study design. Connected speech samples were elicited at each time point using two discourse contexts: (1) responses to WH-questions and (2) personal narrative production. Samples were audio-recorded, transcribed, and coded.

Outcome measures included (1) CIUs and total verbal units (TVUs), used to calculate a raw CIU score as well as the CIU/TVU ratio (lexical informativeness and efficiency); (2) AS-unit analysis to determine sentence complexity and completeness (e.g., proportion of simple vs. complex utterances/total utterances). Analyses focused on within-participant changes over time and relationships between lexical and syntactic measures across contexts. No condition-specific (SLT vs. SLT+AE) effects were observed; therefore, the present analysis focuses on changes associated with SLT and connected speech outcomes.

Results

Results varied across participants in WH-question responses; narrative data are currently being analyzed. For P1, a pattern suggestive of a potential tradeoff was observed in WH-question responses: sentence-level complexity decreased, while lexical efficiency (CIU/TVU

ratio) remained relatively stable or showed slight improvement. In contrast, P2 demonstrated a pattern suggestive of a tradeoff pattern in which increases in sentence-level complexity co-occurred with decreases in the CIU/TVU ratio, indicating reduced lexical efficiency alongside greater syntactic complexity.

Discussion

The present findings suggest that SLT-related changes in connected speech may reflect both shared and individual-specific patterns of language recovery in post-stroke aphasia. While both participants demonstrated gains in one of the measures, either sentence complexity or efficiency following SLT, the relationship between lexical informativeness and syntactic complexity differed. P1 demonstrated a tradeoff pattern in which slight gains in efficiency came at a cost of slightly decreased structural complexity and completeness. In contrast, P2 exhibited a tradeoff pattern in which increased structural complexity was accompanied by reduced efficiency. These findings highlight that improvements in connected speech do not follow a uniform trajectory and may instead reflect individual differences in how linguistic resources are distributed across domains. Examining lexical informativeness and sentence complexity together provides a more comprehensive understanding of treatment-related change than either measure alone (Altman & Goral, 2025).

Future research with larger samples is needed to determine the stability of these patterns and whether task type systematically influences tradeoff dynamics. Clinically, these results support individualized treatment goals that consider potential interactions between efficiency and complexity, rather than assuming parallel improvement across domains.

References

Altman, C., & Goral, M. (2025). Towards an integrated narrative analysis in nonfluent aphasia. Aphasiology, 1–25. https://doi.org/10.1080/02687038.2025.2479793

Azios, J. H., Archer, B., Simmons-Mackie, N., Raymer, A., Carragher, M., Shashikanth, S., & Gulick, E. (2022). Conversation as an outcome of aphasia treatment: A systematic scoping review. *American Journal of Speech-Language Pathology, 31*(6), 2920–2942. https://doi.org/10.1044/2022_AJSLP-22-00011

Dipper, L., Marshall, J., Boyle, M., Hersh, D., Botting, N., & Cruice, M. (2021). Creating a Theoretical Framework to Underpin Discourse Assessment and Intervention in Aphasia. Brain sciences, 11(2), 183. https://doi.org/10.3390/brainsci11020183

Stark, B. C., Bryant, L., Themistocleous, C., den Ouden, D.-B., & Roberts, A. C. (2023). Best practice guidelines for reporting spoken discourse in aphasia and neurogenic communication disorders. Aphasiology, 37(5), 761–784. https://doi.org/10.1080/02687038.2022.2039372

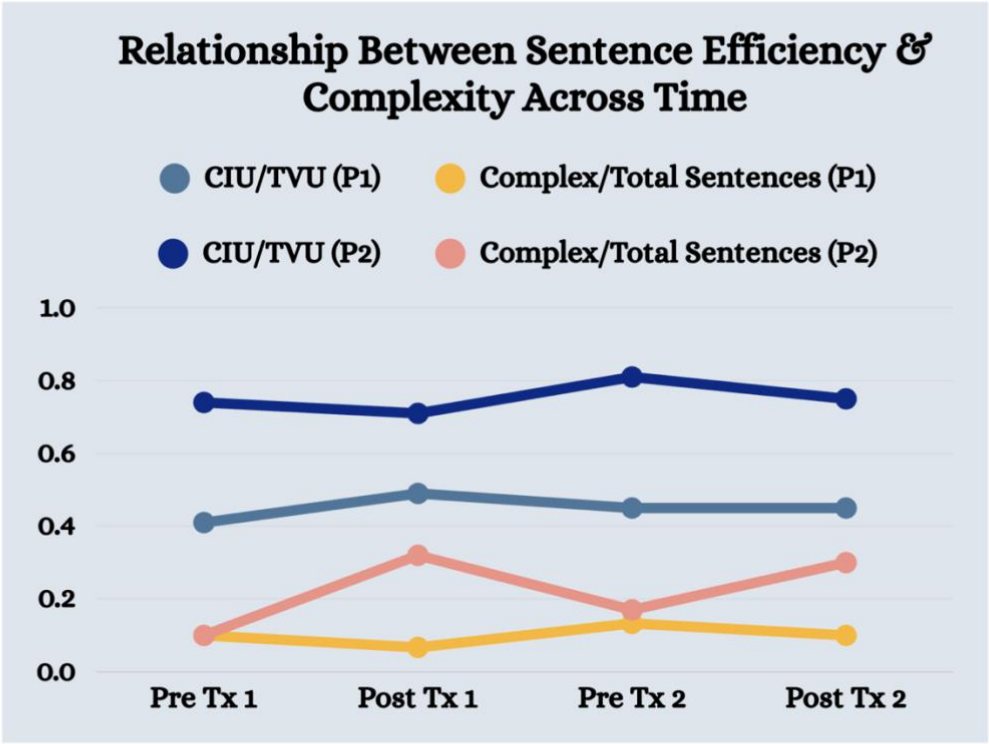


Figure 1. Relationship between lexical efficiency (CIU/TVU ratio) and the proportion of complex and complete utterances (AS-units of 3-5) out of total utterances for both participants (P1 and P2).

Syntactic impairment in Huntington's disease: a case series study

Katerina Gkatzioni | Elpida Grouiou | Vasiliki Koukoulioti | Panagiotis Ioannidis | Stavroula Stavrakaki

Introduction and aims

Huntington's disease (HD) is a fatal, inherited neurodegenerative disorder which affects the striatum of the basal ganglia. Symptoms include motor, cognitive, but also language impairments. Concerning language, patients with HD are assumed to fail to inhibit the application of morphosyntactic rules. Previous studies suggest not only lower complexity in terms of number of embedded clauses (Hinzen et al., 2018), but also a high frequency of errors in sentence connection of both main and embedded clauses (Tovar et al., 2020). Comprehension of complex sentences, including object relative clauses, is also impaired (Teichmann et al., 2008). These deficits are related to damage in the striatum, which is involved in procedural memory, the language computational system, affecting aspects of both morphology and syntax, whereas language abilities associated with declarative memory, such as vocabulary, remain unaffected (Ullman, 2004). The present study aims to investigate morphosyntactic abilities in HD by examining the case of three Greek-speaking individuals. In particular, the study addresses the question whether HD patients show difficulties in the production and comprehension of syntactically complex sentences.

Methods

The participants were recruited at the B' Department of Neurology of the AHEPA University Hospital in Thessaloniki (Board meeting number of the Committee of Bioethics of the School of Medicine of the Aristotle University of Thessaloniki: 220/2024) and were diagnosed with HD on the basis of a genetic test and classified, with clinical consensus, as being in the early phase of the disease. Their language abilities were assessed with the Boston Diagnostic Aphasia Examination (BDAE)^{GR} (Messinis et al., 2013) and the Mini Linguistic State Examination (Papadopoulou et al., 2025) and their cognitive abilities with Addenbrooke's Cognitive Examination (ACE-R^{GR}, Konstantinopoulou et al., 2011). P1 produced short sentences and had low performance in repetition and comprehension of complex material. P2 had low performance in several tasks in the BDAE and, importantly, sentence comprehension. Finally, P3 performed in the normal range in all sub-tests of BDAE. The assessment of linguistic abilities was performed by means of the Test of Production and Comprehension of Relative Clauses (henceforth PCRC, see for example, Talli & Stavrakaki, 2020). The PCRC consists of two parts: 1) production part which aims at the elicitation of subject and object relative clauses by means of elicited answers to questions referring to pictures with animal figures, 2) the comprehension part, which is a picture-sentence matching task, testing the same structures.

Results

Table 1 presents the performance of the three patients (%). All patients were impaired in the production of object relative clauses. P1 and P2 were impaired also in the production of subject relative clauses. P1 and P3 exhibited the dissociation between subject and object relative clauses, whereas P2 was severely impaired in both. Concerning comprehension, P1 had overall better performance, with the dissociation between subject and object relative clauses being relatively small. P2 was severely impaired also in comprehension. Finally, P3

had lower performance in comprehension than in production, with no dissociation between the two structures.

Overall, P2 showed impaired performance across the board, regardless the type of the relative (subject vs. object) and modality (production vs. comprehension). By contrast, P1 and P3 showed selective impairments with object relatives to be significantly affected in the production task.

Discussion

The performance of these patients with HD was highly variable. P2 seems to be impaired across structures and modalities, although her scores in ACE-R and MLSE were not particularly low. This finding points to a specific language vulnerability in HD, before other cognitive abilities are affected. P1 exhibited a modality effect with production being more impaired than comprehension and a structure effect with object relative clauses being more impaired than object relative clauses. The structure effect points to an impairment of syntactic rule application, which manifests more in the most challenging structure in the most challenging modality. The performance of P3 confirms the manifestation of language impairment in the early phases of the disease, although he exhibited a more complex pattern with better performance in production than in comprehension for subject relative clauses and better performance in comprehension than production for object relative clauses. These results reveal vulnerability in the syntactic domain even at the early stages of patients with HD. The generalizability of the results is limited by the small sample; nevertheless, they provide evidence for a syntactic deficit in the early stages of HD, which deserves further investigation.

Selected references

Hinzen, W., Rosselló, J., Morey, C., Camara, E., Garcia-Gorro, C., Salvador, R., & de Diego-Balaguer, R. (2018). A Systematic Linguistic Profile of Spontaneous Narrative Speech in Pre-Symptomatic and Early-Stage Huntington's Disease. *Cortex (100)*, 71–83. https://doi.org/10.1016/j.cortex.2017.07.022

Talli, I., & Stavrakaki, S. 2020. “Short-Term Memory, Working Memory and Linguistic Abilities in Bilingual Children with Developmental Language Disorder”, *First Language* 40(4), 437–460. https://doi.org/10.1177/0142723719886954

Teichmann, M., Gaura, V., Demonet, J.-F., Supiot, F., Delliaux, M., Verny, C., Renou, P., Remy, P., & Bachoud-Levi, A.-C. (2008). Language Processing Within the Striatum: Evidence from a PET Correlation Study in Huntington’s Disease. *Brain*, *131*(4), 1046–1056. https://doi.org/10.1093/brain/awn036

Tovar, A., Garí Soler, A., Ruiz-Idiago, J., Mareca Viladrich, C., Pomarol-Clotet, E., Rosselló, J., & Hinzen, W. (2020). Language Disintegration in Spontaneous Speech in Huntington's Disease: A More Fine-Grained Analysis. *Journal of Communication Disorders*, *83*, 105970. https://doi.org/10.1016/j.jcomdis.2019.105970

Ullman, M. T. (2004). Contributions of Memory Circuits to Language: The Declarative/Procedural Model. *Cognition*, *92*(1-2), 231-270. https://doi.org/10.1016/j.cognition.2003.10.008

	P1	P2	P3
Sentence Type	**Production**		
SubjRel	70	20	100
ObjRel	20	30	50
	Comprehension		
SubjRel	100	25	75
ObjRel	81	38	69

Table 1: percentage correct responses. SubjRel: subject relative clauses, ObjRel: object relative clauses

Where is the locus of language mixing in bilingual aphasia? A preliminary comparison of two cases.

Solène Hameau | Myriam Khoury | Mathilde Iserentant | Jennifer Martin | Mareike Moormann | Joana Cholin

Introduction and aims

Language mixing, referring to intrusions from the non-target language within utterances, is a common phenomenon in bilingual aphasia. Mixing can occur during picture naming tasks, where speakers may produce words from the non-target language, alternate between languages across attempts, or combine elements from both languages within a single response. Such behaviours are commonly described as pathological mixing and can be contrasted with normal, socially constrained code-switching in unimpaired bilinguals.

Pathological mixing in aphasia is often attributed to either a disruption of the language control network, leading to unwanted activation or selection of the non-target language (Green & Abutalebi, 2008), or to compensatory strategies related to lexical retrieval difficulties, whereby words from the stronger language are used to fill gaps in the weaker language (e.g., Goral et al., 2019; Hameau et al., 2023). Although language mixing at the single-word level has been documented in numerous case studies of bilingual aphasia, and different authors have attributed mixing to distinct underlying impairments (e.g., impaired control vs lexical retrieval difficulty), to our knowledge, no studies have directly contrasted these mechanisms based on single-word mixing behaviour across multiple bilinguals with aphasia within the same theoretical framework.

To address this gap, we examine two bilingual individuals with aphasia who both exhibit language mixing during picture naming, but who differ markedly in the form of their mixed responses, their awareness and monitoring of these responses, and their broader aphasic impairment profiles. To better understand whether their mixing behaviours stem from distinct impairments at the single word level, we analyse how different types of mixing behaviours are predicted by different psycholinguistic factors. By contrasting these two profiles, this study aims to examine whether language mixing can arise from qualitatively different disruptions within the word production system.

Methods

Participants

Data from two bilingual participants with aphasia (that we call CA and CB) was used for the present study. Relevant participant information can be found in Table 1. Data from the first participant (CA) was taken from an existing published study (Hameau et al., 2023). CA was a Spanish-English late bilingual, whereas CB was a French-English-Spanish late trilingual. Both participants presented with chronic aphasia following a stroke, and both exhibited language mixing when naming pictures. While CA's behaviours suggested he was very aware of his mixing, for CB the picture was much less clear.

Materials and procedure

Participants (see Table 1 for background information) underwent extensive background testing in all their languages (e.g., Bilingual Aphasia Test, bilingualism questionnaires, and a Simon task).

Picture naming was assessed for both participants on a set of 250 coloured pictures with >85% name agreement (Duñabeitia et al., 2018), in both Spanish and English for CA, and in French for CB (English and Spanish picture naming tasks will be completed before the conference presentation). Pictures were presented on a computer screen for both participants (through a Powerpoint presentation for CA and using OpenSesame for CB). Both participants were instructed to name the pictures in the target language only. To ensure correct language mode, the task was administered by a native speaker for each language.

For CB, each picture naming trial was immediately followed by a response certainty judgment (indicating the extent she thought she produced the correct word on a 1-100 scale).

For each target, several item-specific variables were obtained (see Analyses section).

Error coding

Within the errors produced by each participant in the naming task, we considered full word mixing errors (translation equivalent, formal or semantic error in the other language), neologisms consisting of parts of the words in two languages, and words produced (in L1) with an L2/L3 accent.

Analyses

For each participant, we will examine the influence of a number of psycholinguistic variables on the probability of producing each type of mixing error as compared to a correct response, using GLME : semantic variables (imageability, concept familiarity), lexical variables (word frequency, name agreement, phonological overlap of the translation equivalent – a measure of "cognateness", phonological neighbourhood density) and sublexical variables (syllable frequency, whether the first phoneme is shared across translation equivalents).

Results

Full results will be presented at the conference.

Discussion

The present study is expected to contribute to ongoing debates on language mixing in bilingual aphasia by showing how similar surface behaviours may arise from different underlying processes. By adopting a fine-grained, psycholinguistically informed analysis of single-word mixing errors, this approach has the potential to clarify how different forms of mixing relate to distinct components of the word production system, including lexical access, downstream phonological or phonetic processes, and monitoring. Within control-based frameworks of bilingual language production, the study aims to refine interpretations of pathological mixing by emphasizing the need to specify which linguistic processes are affected, rather than equating mixing with a global control impairment. Clinically, the work

underscores the importance of interpreting language mixing in relation to its form and functional context, rather than treating it as a unitary symptom.

References

Duñabeitia, J. A., Crepaldi, D., Meyer, A. S., New, B., Pliatsikas, C., Smolka, E., & Brysbaert, M. (2018). MultiPic: A standardized set of 750 drawings with norms for six European languages. Quarterly Journal of Experimental Psychology (2006), 71(4), 808–816. https://doi.org/10.1080/17470218.2017.1310261

Goral, M., Norvik, M., & Jensen, B. U. (2019). Variation in language mixing in multilingual aphasia. *Clinical Linguistics & Phonetics, 33*(10–11), 915–929. https://doi.org/10.1080/02699206.2019.1584646

Green, D. W., & Abutalebi, J. (2008). Understanding the link between bilingual aphasia and language control. *Journal of Neurolinguistics, 21*(6), 558–576. https://doi.org/10.1016/j.jneuroling.2008.01.002

Hameau, S., Dmowski, U., & Nickels, L. (2023). Factors affecting cross-language activation and language mixing in bilingual aphasia: A case study. *Aphasiology, 37*(8), 1149–1172. https://doi.org/10.1080/02687038.2022.2081960

Table 1. Participants' background information

	CA	**CB**
Age	77	54
Sex	Male	Female
Handedness	Forced right-hander	Left-hander
Education	15 years	16 years
Languages	L1 Spanish, L2 English (L1-dominant but very proficient in L2)	L1 French, L2 English, L3 Spanish (L1-dominant, good pre-stroke proficiency in L2 and L3)
Aetiology	Left MCA stroke five years prior to the study	Right hemisphere stroke consecutive to a meningioma resection 3 years prior to the study
Aphasia type and severity	Moderate non-fluent aphasia in L1 and L2, with agrammatism	Mild anomic aphasia in L1, L2 and L3 (with some phonological difficulties)

Dissociation of Retrieval and Inflection of Verbs and Adjectives in Bilingual Speakers with Aphasia

Silvia Hargrove | Philomena Elom | Ioulia Agrotou | Mira Goral

Introduction and Aims

Language production relies on the interaction between lexical retrieval and morphosyntactic encoding, two processes that are closely related yet dissociable. Psycholinguistic models posit multiple stages of word production, including lexical selection and grammatical encoding, which may be independently disrupted in aphasia (Benetello et al., 2016; Faroqi-Shah & Friedman, 2015). Empirical evidence suggests that individuals with aphasia may retrieve lexical items without correctly inflecting them, and impairments at different levels of lexical retrieval could result in distinct profiles, reflected in specific error types (Biran et al., 2024).

Research has focused predominantly on verbs due to their role in argument structure and tense marking. However, adjectives, particularly in morphologically rich languages also pose significant morphosyntactic demands through agreement marking. Despite this, adjectives remain under-investigated, especially in bilingual aphasia (Meltzer-Asscher & Thompson, 2014). Cross-linguistic differences between English and Spanish provide a testing ground for examining how lexical and morphosyntactic processes interact under impairment (Kroll & Bialystok, 2013).

The present study examines the relationship between lexical retrieval and morphosyntactic encoding across grammatical categories (verbs vs. adjectives) and languages (Spanish vs. English) in bilingual individuals with aphasia. Specifically, we aim to determine whether retrieval and inflection are dissociable within each category, whether this relationship differs across categories, and whether observed patterns reflect language-specific constraints or general aphasia-related impairments.

Methods

Eight bilingual individuals with chronic non-fluent aphasia were recruited in New York City. Participants ranged in age from 49 to 74 years and demonstrated mild-to-moderate aphasia severity (WAB AQ = 62.9–94.1). All reported high pre-stroke proficiency in both languages, with post-stroke impairment observed across both languages, typically with greater preservation of the first language (see Table 1).

Participants completed tasks assessing lexical retrieval and morphosyntactic encoding for verbs and adjectives in both languages. Verb retrieval was assessed using an action-naming task, while verb inflection was evaluated using a task requiring tense and agreement marking (Faroqi-Shah, unpublished). Adjective inflection and retrieval were assessed using the Adjective Noun Agreement Task (ANAT; Goral & Borodkin, unpublished). Verb and adjective retrieval and inflection in connected language production was assessed using a semi-spontaneous narrative task.

Data were analyzed using generalized linear mixed-effects models (GLMMs) with binomial distributions. Fixed effects included TaskType (retrieval vs. inflection), Language, and Category (verb vs. adjective), with participant as a random intercept. Contrast between retrieval and inflection of adjectives was analyzed in Spanish only because English adjectives

do not inflect for gender or number. Estimated marginal means were computed, and detailed error analyses were conducted for both lexical and morphosyntactic responses.

Results

A dissociation between lexical retrieval and morphosyntactic encoding emerged, with a significant Category × TaskType interaction ($\beta = 1.31$, $p < .001$), confirming a crossover effect. Namely, for verbs, retrieval exceeded inflection (OR = 1.87, $p = .010$), whereas for adjectives, inflection exceeded retrieval (OR = 0.50, $p = .008$).

For verbs, retrieval accuracy was significantly higher than inflection accuracy across languages ($\beta = -0.87$, $p < .001$). Retrieval accuracy reached 73% in English and 58% in Spanish, compared to inflection accuracy of 53% and 45%, respectively, reflecting a consistent retrieval advantage. At the individual level, the majority of participants (5/7 in English; 4/7 in Spanish) showed this pattern. Exceptions revealed meaningful clinical variability: one participant (ES12) showed the reverse pattern (inflection > retrieval), while another (SE05) showed preserved inflection in Spanish only.

In contrast, the ANAT showed the opposite pattern of the verb findings. In Spanish, inflection accuracy (86%) was higher than retrieval accuracy (74%; $\beta = 0.74$, $p = .006$), indicating an inflection advantage. This pattern was variable across participants, but no participant showed a reliable retrieval advantage. Individual profiles further supported this dissociation, with some participants (e.g., ES12, SE02, and SE09) showing higher inflection accuracy than retrieval and others performing near ceiling on both tasks.

Error analyses provides additional insight into the nature of the verb and adjective deficit. Action naming errors revealed greater lexical difficulty in Spanish, with higher rates of verb substitutions and paraphrases. Cross-linguistic interference was observed, particularly in Spanish responses influenced by English. In the verb inflection task, in English, tense and inflection errors were comparable (M = 5.3 vs. 4.5), whereas in Spanish, tense errors predominated (M = 8.7 vs. 3.1). Across both languages, past-tense overproduction was the dominant substitution pattern, accounting for 50% of English and 61% of Spanish tense errors. Language-specific error forms also emerged: infinitive substitutions occurred only in Spanish, while progressive tense substitutions occurred only in English.

Discussion

These findings provide evidence that lexical retrieval and morphosyntactic encoding are partially dissociable processes in bilingual people with aphasia, with the nature of this dissociation varying by grammatical category. For verbs, retrieval supports but does not guarantee successful inflection, consistent with models that posit hierarchical dependencies between lexical and morphosyntactic processing (Benetello et al., 2016). The predominance of tense errors, particularly past-tense overproduction, suggests that temporal encoding is vulnerable, likely due to increased cognitive and discourse-level demands.

In contrast, adjectives exhibited a reversed pattern, with relatively preserved inflection despite impaired retrieval. The interaction between verbs and adjectives demonstrates that retrieval and inflection are not globally impaired processes but are differentially affected depending on grammatical category. This supports a multi-component view of language production, in

which distinct linguistic domains (e.g., tense vs. agreement) impose different processing demands (Chen & Bates, 1998).

Cross-linguistic differences highlight the role of morphological complexity. Lower performance in Spanish and the presence of language-specific error patterns indicate that morphologically rich systems amplify encoding demands while also shaping the form of errors. These findings align with the Competition Model, suggesting that language-specific cue strength influences processing under impairment (MacWhinney, 2022).

Finally, individual variability underscores the heterogeneity of aphasia. Distinct profiles indicate that deficits may arise at different levels of the production system. Clinically, this highlights the importance of assessing both lexical and morphosyntactic processes across multiple categories and languages.

References

Benetello, A., Finocchiaro, C., Capasso, R., Capitani, E., Laiacona, M., Magon, S., & Miceli, G. (2016). The dissociability of lexical retrieval and morphosyntactic processes for nouns and verbs: A functional and anatomoclinical study. *Brain and Language*, *159*, 11–22.
Biran, M., Ben-Or, G., & Yihye-Shmuel, H. (2024). Word retrieval in aphasia: From naming tests to connected speech and the impact on well-being. *Aphasiology*, *38*(4), 738–757.
Chen, S., & Bates, E. (1998). The dissociation between nouns and verbs in Broca's and Wernicke's aphasia: Findings from chinese. *Aphasiology*, *12*(1), 5–36.
Faroqi-Shah, Y., & Friedman, L. (2015). Production of Verb Tense in Agrammatic Aphasia: A Meta-Analysis and Further Data. *Behavioural Neurology*, *2015*, 1–15.
Kroll, J. F., & Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and cognition. *Journal of Cognitive Psychology*, *25*(5), 497–514.
MacWhinney, B. (2022). The Competition Model: Past and Future. In J. Gervain, G. Csibra, & K. Kovács (Eds.), A Life in Cognition (Vol. 11, pp. 3–16). Springer International Publishing.
Meltzer-Asscher, A., & Thompson, C. K. (2014). The forgotten grammatical category: Adjective use in agrammatic aphasia. *Journal of Neurolinguistics*, *30*, 48–68.

Table 1. Participants' information

Participants	Age	Sex	Years of education	Time post-stroke (years)	WAB AQ (English)	Etiology	L1	Age of immigration to the US	Age of L2 acquisition	Overall proficiency pre-stroke (self-rated)	Overall proficiency post-stroke (self-rated)
SE02	51	F	14	27	82.9	Left CVA	Spa	n/a	5	L1: 10/10 L2: 10/10	L1: 7/10 L2: 7/10
SE04	49	F	11	1st: 3 2nd: 1	79.2	Left CVA	Spa	n/a	4	L1: 10/10 L2: 10/10	L1: 7/10 L2: 5/10
SE05	58	F	17	8 months	92.4	Left CVA	Spa	45 y	31	L1: 10/10 L2: 7/10	L1: 5/10 L2: 4/10
SE06	54	F	18	5	94.1	Left CVA	Eng	n/a	3	L1: 10/10 L2: 10/10	L1: 5/10 L2: 3/10
SE07	74	M	18	1	87.5	Left CVA	Spa	36 y	14	L1: 10/10 L2: 9/10	L1: 6/10 L2: 6/10
SE09	54	F	14	1	74.8	Left CVA	Spa	13 y	13	L1: 10/10 L2: 10/10	L1: 5/10 L2: 3/10
SE10	55	F	12	9	62.9	Left CVA	Spa	17 y	18	L1: 10/10 L2: 7/10	L1: 3/10 L2: 2/10
ES12	70	F	19	1	87.5	Left CVA	Eng	n/a	3	L1: 10/10 L2: 7/10	L1: 6/10 L2: 3/10

Novelty meets Effectiveness: Creative Naming Strategies in Persons with Aphasia

Mareike Hartmann | Lotta Heidemann | Carola de Beer | Martina Hielscher-Fastabend | Kerstin Richter

Introduction and aims

Linguistic creativity is a fundamental cognitive ability, enabling the formation and/ or interpretation of complex words and expressions (Körtvélyessy et al., 2022). Variability in creative linguistic potential is shaped by executive functions such as cognitive flexibility, inhibitory control, and working memory (Khalil et al., 2019). An expression is generally considered creative if it is both *original* and *effective.* In this context, originality refers to the novelty of the expression, and effectiveness determines whether and to what extent it is successful in communication (Runco & Jaeger, 2012).

Semantic-lexical deficits in persons with aphasia (PWA) negatively affect their communication. Difficulties in retrieving complex word forms, such as compound nouns and verbs, are well documented in PWA (e.g., Lorenz et al., 2014; Rofes et al., 2015) and may foster compensatory behavior, including creative use of language. While semantic paraphrasing is widely accepted as a communicative strategy, the production of creative expressions – such as semantic deviations and neologisms (e.g., *map ball* for *globe*, example from Liederman et al., 1983) – are predominantly treated as errors in research and clinical contexts (e.g., Howard & Gatehouse, 2006). Up to date, only a few approaches highlight the importance of linguistic creativity for communication in PWA (Holland, 2021; Liederman et al., 1983).

Against this background, the aim of this ongoing study is to investigate the extent to which PWA demonstrate creative language use in naming, and how this creativity contributes to their success in communication.

Methods

At the time of abstract submission, participants included 10 PWA (4 women, 6 men, age range: 55-74 years) with a moderate to mild chronic aphasia following a left-hemispheric stroke. The control group consists of 40 neurologically unimpaired controls (NUC), matched for age and gender (20 women, 20 men, age range: 19-73 years). Participants were asked to verbally name 20 objects and 20 activities, requiring the retrieval of German compound nouns and infinitive verbs. Each target image was presented alongside with a distractor image. For compound nouns, the distractors shared the same formal head (e.g., *pufferfish* vs. *surgeonfish* or *safety vest* vs. *life vest*), while the verbs were semantically similar (e.g., *sail* vs. *row* or *leash* vs. *walk*). This design was developed to elicit and to ensure the production of target compounds and verbs. Due to the specificity and rarity of the materials, it places considerable demands on the production of German compounds and verbs in both PWA and NUC. Therefore, it is expected that misnaming or alternative (potentially novel) designations may occur. According to the task criteria, an answer is deemed effective if it enables a clear distinction between the target image and the distractor image, for instance by incorporating distinctive semantic features into the response.

A special annotation scheme was developed to analyze the responses, integrating morphological aspects (e.g., compound structure, verb construction and inflection) and semantic aspects. In this light, semantic features were defined based on the Semantic Feature Analysis (SFA, Boyle & Coelho, 1995; Wambaugh & Ferguson, 2007). These features proposed a framework for identifying the semantic dimensions participants employed when producing linguistic variation and creative expressions. As it is known from the SFA, semantic features capture core conceptual attributes of the targets, including category membership, function, and perceptual properties (for objects), as well as typical tools, locations or purposes (for actions).

Results

Data collection and analysis are ongoing at the time of abstract submission. Initial analyses of response patterns indicate that both PWA as well as NUC frequently rely on existing, semantically related words when the exact term for the target cannot be retrieved or is unfamiliar (e.g., *sawfish* for *swordfish* or *forming* for *to do pottery*). In PWA and NUC, the production of semantic neologisms is observed in some cases. Comparing compound and verb production, semantic neologisms occur more often in the naming of objects (e.g., *wing carousel* for *chain carousel*), but can also be found on occasion when activities are named (e.g., the novel German verb *anspangen*, which denotes the action of putting one's hair up, and is derived from the German noun *Spange* (English: hairclip)). Overall, PWA show marked difficulties in naming single words and also in compound word formation. In these cases, they produce semantic approximations and phrase-like structures (e.g., */potato .. fried . in a pan/* for *pan-fried potatoes* or /*spoon with dots inside*/ for *skimmer*). Across all response types (single-word responses and phrases), visual characteristics and functional attributes of the objects emerge as the primary semantic reference points, whereas actions are often named in the context of a sentence along with their corresponding subject. At the same time, it should be noted that PWA demonstrate effective responses that allow the target image to be distinguished from the distractor image.

Discussion

Taken together, these findings suggest that when precise lexical retrieval is not possible, PWA employ creative lexical strategies and exhibit linguistic variation, which may support communicative success. Even individuals in the control group (NUC) show linguistic variation in approximately 10 – 15% of responses. Participants from both groups produce novel word forms, although the proportion of semantic neologisms is lower among PWA than among NUC. Due to their difficulty in generating single words, PWA often use phrase- and sentence-like structures and semantic approximations, which can be considered effective because they allow for a distinction between the target image and the distractor image.

It remains to be investigated which specific cognitive and linguistic skills lead to creative language use and linguistic variability in PWA. At present, the effectiveness of creative expressions is evaluated by enabling the distinction between the target image and the distractor. In future evaluation studies, both the comprehensibility and originality of the verbal expressions will additionally be rated by independent evaluators.

References

Holland, A. L. (2021). The value of “communication strategies” in the treatment of aphasia. *Aphasiology*, *37*(7), 984-994.
Howard, D. & Gatehouse, C. (2006). Distinguishing semantic and lexical word retrieval deficits in people with aphasia. *Aphasiology*, *20*(9), 921–950.
Khalil, R., Godde, B. & Karim, A. A. (2019). The Link Between Creativity, Cognition, and Creative Drives and Underlying Neural Mechanisms. *Frontiers in Neural Circuits*, *13*, Artikel 18.
Körtvélyessy, L., Štekauer, P. & Kačmár, P. (2022). *Creativity in Word Formation and Word Interpretation: Creative Potential and Creative Performance*. Cambridge University Press.
Liederman, J., Kohn, S., Wolf, M. & Goodglass, H. (1983). Lexical Creativity during Instances of Word-Finding Difficulty: Broca's vs. Wernicke's Aphasia. *Brain and Language*, *20*(1), 21–32.
Lorenz, A., Heide, J. & Burchert, F. (2014). Compound naming in aphasia: effects of complexity, part of speech, and semantic transparency. *Language, Cognition and Neuroscience*, *29*(1), 88–106.
Rofes, A., Capasso, R. & Miceli, G. (2015). Verb production tasks in the measurement of communicative abilities in aphasia. *Journal of Clinical and Experimental Neuropsychology*, *37*(5), 483–502.
Runco, M. A. & Jaeger, G. J. (2012). The Standard Definition of Creativity. *Creativity Research Journal*, *24*(1), 92–96.

Beyond words: Capturing Creativity in Speech-Replacing gestures in people with aphasia

Lotta Heidemann, Mareike Hartmann, Kerstin Richter,, Martina Hielscher-Fastabend & Carola de Beer

Introduction and aims

Despite the common misconception, that creativity is limited to a few brilliant artists, it is an integral part of the cognitive function exhibited by every human being. Creativity can be conceptualised as a process of problem solving and idea generation (Nijstad et al., 2010). Problem solving is also needed when encountering communicative challenges such as misunderstandings or time pressure.

People with aphasia (PWA) face significant challenges in their daily living, as their language impairments affect their communicative abilities. To overcome their communicative challenges (e.g. word retrieval difficulties) PWA need communicative strategies. One possible strategy is the utilisation of creative verbal utterances since PWA have been observed to effectively produce creative expressions in order to replace a non-retrievable target word (Liederman et al., 1983). The communication of PWA is not only limited to verbal utterances. Studies have shown that gestures of PWA provide supplementary information in addition to their verbal output (e.g. De Beer et al., 2017) and therefore gestures can gain relevance in the communication of PWA. In cases of communicative pressure some PWA are reported to produce *creative gestures* as a strategy to deal with communicative challenges (Poggi, 2002). Poggi (2002) defines creative gestures as being created spontaneously when no readily accessible or conventionalised gesture is available. Studies also examined the forms of creative gestures and claim that these gestures are more complex and effortful than non-creative gestures (e.g. use of a more extensive gesture space, more dynamic movement) (Cienki & Mittelberg, 2013).

Despite the importance of gestures in communication of PWA only a few studies focused on creative gesture use in PWA. Yet, the coding and description of creative gesture production in PWA remains unclear.

We therefore aim to investigate the forms of creativity in gesture production and potentially relevant underlying cognitive abilities (e.g. executive functioning) of PWA. Additionally, we aim to develop a new coding scheme to capture the key aspects of creativity in gesture production. In order to ensure the production of complex gestures in our participants, we decided to examine the production of speech-replacing gestures. We employed a task that necessitates the creation of speech-replacing gestures to represent objects without readily accessible gestures (e.g. hammerhead shark, German: *Hammerhai),* enhancing creative problem-solving processes in gesture production, in accordance with Poggi's (2002) definition of creative gestures.

Method

People with mild to moderate aphasia (n=9) and neurologically unimpaired controls (NUC; n= 34) saw 20 photographs of objects (e.g., animals, tools), representing German compound nouns. Each target picture (e.g., hammerhead shark) was presented alongside a picture of a cohyponym (e.g., tiger shark). Participants were asked to solely use gestures to illustrate the

target object with clear distinction from the cohyponym. The data collection of PWA is still ongoing.

To describe creativity in speech-replacing gesture production, we developed a new coding

scheme. Previous literature has drawn attention to formal aspects of gestures and communicative strategies in creative gesture production (Cienki & Mittelberg, 2013; Poggi, 2002). According to this literature our coding scheme incorporates the two dimensions (i) **expression strategies** and (ii) **formal aspects** to systematically describe creativity in gesture production. We first coded **expression strategies** to describe utilisation and combination of different gestures employed by participants for represent the target objects in distinction to the cohyponym. Table 1 provides an overview of the strategies with explanations and examples. The employment of effective strategies is predicated upon the capacity to distinguish the target item from the cohyponym.

To operationalise the aspects of form and meaning of each speech-replacing gesture produced, we focused on semantic properties and coded **semantic features** identifiable in the gestures (e.g. object-shape, object-size, movement-intensity; movement-type; following Beattie & Shovelton (2011)).

Results

Preliminary results of the strategy use show that all three different strategies were represented in our sample. The strategy *Distinctive feature* was the most frequently employed strategy, accounting for 92.6% of all strategies produced by the PWA group and 79.1 % of all strategies in the control group. For PWA, *Storytelling* was the least frequently employed strategy (2.1% of all cases of strategy use) and was only adopted by three PWA. This strategy was also the least frequently employed strategy of NUC's, however with a higher proportion compared to the PWA (23.3 % of total strategies of NUC). It was observed that only three PWA employed all three strategies during the whole experiment, while this applied to 27 of the NUC. More detailed group comparisons on the semantic features identified in participants' gestures will be presented at the conference.

Discussion

PWA show inter-individual variation in their selection of strategy use to distinguish the target object from its cohyponym only using gestures. The results suggest that some PWA have the capacity to flexibly change the employment of the different types of expression strategies with respect to the content to be expressed in a similar way to the NUC. Some other PWA stick to one or two of the strategies, without adjusting those to the content to be expressed. In summary, our preliminary results suggest a higher inter-individual variability in the employment of expression strategies in PWA than in NUC, suggesting underlying inter-individual differences in creativity in PWA.

References

Beattie, G., & Shovelton, H. (2011). An exploration of the other side of semantic communication: How the spontaneous movements of the human hand add crucial meaning to narrative. *Semiotica*, *2011*(184).

Cienki, A., & Mittelberg, I. (2013). Creativity in the forms and functions of spontaneus gestures with speech. In T. Veale, K. Feyaerts, & C. Forceville (Eds.), *Creativity and the agile mind: A multi-disciplinary study of a multi-faceted phenomenon* (pp. 231–252). De Gruyter Motion
De Beer, C., Carragher, M., van Nispen, K., Hogrefe, K., De Ruiter, J. P., & Rose, M. (2017). How Much Information Do People With Aphasia Convey via Gesture? AJSLP, *26*(2), 483–497.
Liederman, J., Kohn, S., Wolf, M., & Goodglass, H. (1983). Lexical creativity during instances of word-finding difficulty: Broca's vs. Wernicke's aphasia. *Brain Lang*, *20*(1), 21–32.
Nijstad, B. A., De Dreu, C. K. W., Rietzschel, E. F., & Baas, M. (2010). The dual pathway to creativity model: Creative ideation as a function of flexibility and persistence. ERSP, *21*(1), 34–77.
Poggi, I. (2002). From a Typology of Gestures to a Procedure for Gesture Production. In I. Wachsmuth & T. Sowa (Eds.), *Gesture and Sign Language in Human-Computer Interaction* (pp. 158–168). Springer

Table 1 Different strategies in creative gesture production based on the example item *chain carousel* (German: *Kettenkarussell*) with the cohyponym *children's merry-go-round*

Strategy	Definition	Examples from the NUC participants
Decomposition	Participant separates the reference to a) the head and b) the modifier of the target compound	Tracing gesture for the shape of a necklace (reference to chain) followed by an upward facing circular movement (reference to carousel)
Storytelling	Participant "tells a story" via gestures, i.e., provides context information	Participant starts his gestural expression by imitation to buy a ticket from a ticket counter
Distinctive feature	Participant is depicting the on distinctive feature of the target object (in opposition to the co-hyponym)	Performing a circular movement above the head (reference to high position in space of the chain carrousel)

Embedding-Based Discourse Coherence in Spanish Dementia Speec

Dilan Joseph García Inclán, Silvia Martinez Ferreiro, Alexandre Nikolaev

Introduction and Aims

Discourse in dementia often shows early semantic decline that may be more prominent than impairments at other linguistic levels. At the discourse level, this is reflected in reduced informational content, weaker thematic coherence, and difficulty maintaining reference across utterances. Literature suggests that these features can help distinguish between pathological and typical speech (Kim & Lee, 2021; Valles, 2006) as they likely reflect impairments in semantic memory and executive control, which limit the ability to plan, organize, and sustain coherent communication (Burke et al., 2023).

Embedding-based measures represent utterances as high-dimensional vectors and quantify semantic similarity across discourse units, providing scalable and objective indices of coherence loss. Both local and global coherence metrics are sensitive to dementia-related decline (Burke et al., 2023). They also help detect mild cognitive impairment (MCI), a transitional state with elevated risk of progression to dementia (Zhang et al., 2021; Liao et al., 2026; Clarke et al., 2021), highlighting their promise as early, non-invasive markers.

However, most studies using these tools have focused on English, with fewer conducted in other languages (Jiang et al., 2025). This creates an important methodological concern as embeddings are trained on language-specific corpora, and their behavior depends on a language's morphological, syntactic, and distributional properties (Nikolaev et al., 2025). Filling this gap can both extend sensitive, reproducible tools to an underserved clinical population and clarify how embedding-based coherence measures perform across structurally distinct languages.

This study addresses this gap by testing embedding-based discourse coherence measures in a Spanish-speaking sample, using two elicitation tasks: occupational recall and a hypothetical scenario. We aim to: (1) test whether coherence metrics differentiate Spanish-speaking individuals with dementia from healthy controls across both tasks; (2) determine whether coherence scores vary by task within and between groups; and (3) assess the psychometric performance of these measures in Spanish relative to benchmarks from English-language research.

Methods

Participants were recruited from the SECONDS-GALES project, a longitudinal study of language-based markers of neurocognitive disorders in Galician-Spanish speakers. The sample includes ~60 adults (65 years): a dementia group (DG: ~30) with a formal clinical diagnosis and an age- and healthy control group (HC: ~30). The tasks were chosen to contrast cognitive demands: occupational recall draws mainly on episodic and semantic long-term memory, whereas the hypothetical scenario relies more on executive control, prospective reasoning, and semantic abilities often affected in dementia. This design tests whether task type modulates the sensitivity of coherence measures, an issue rarely examined in prior work (Clarke et al., 2021; Segkouli et al., 2025).

Speech samples were elicited with the ALEA semi-structured interview, transcribed using CHAT/CLAN conventions, and segmented into utterances before computing coherence. We applied two complementary embedding-based measures: multilingual BERT (mBERT) for sentence-level coherence and fastText for local lexical coherence. This choice reflects the distinction between global and local coherence (Burke et al., 2023) and the models differing sensitivity to morphological structure across languages (Nikolaev et al., 2025).

Results and Discussion

This study presents a methodological framework for embedding-based analysis of discourse coherence in Spanish dementia speech. The proposed pipeline has been fully implemented, including transcription processing, utterance segmentation, and computation of sentence-level (mBERT) and lexical (fastText) coherence measures.

The system is designed to produce quantitative coherence indices at both local and global levels, enabling comparison across clinical groups (dementia vs. healthy controls) and task conditions (occupational recall vs. hypothetical scenario).

While full statistical analysis is ongoing, the framework is validated in terms of feasibility for Spanish clinical speech data and will be applied to the SECONDS-GALES dataset.

Acknowledgements

The SECONDSGALES project PID2023-151673OB-I00 was funded by MICIU/AEI/10.13039/501100011033 and by “ERDF/EU".

References

Alyahya, R. S. W., Lambon Ralph, M. A., Halai, A., & Hoffman, P. (2022). The cognitive and neural underpinnings of discourse coherence in post-stroke aphasia. *Brain Communications*, *47*. https://doi.org/10.1093/braincomms/fcac147

Burke, E., Gunstad, J., & Hamrick, P. (2023). Comparing global and local semantic coherence of spontaneous speech in persons with Alzheimer's disease and healthy controls. Applied Corpus Linguistics, 3(3), 100064.

Clarke, N., Barrick, T. R., & Garrard, P. (2021). A comparison of connected speech tasks for detecting early Alzheimer's disease and mild cognitive impairment using natural language processing and machine learning. Frontiers in Computer Science, 3.

Jiang, H., Chen, Z., Liu, Y., Yang, C., Yuan, X., & He, R. (2025). The structure of spontaneous speech changes in Alzheimer's disease: Crosslingual evidence from English and Greek. PLOS One, 20(5), e0324270.

Kim, B. S., & Lee, M. S. (2021). Discourse performance and related cognitive function in mild cognitive impairment and dementia: A preliminary study. Applied Neuropsychology: Adult, 30(2), 194–203.

Liao, Y.-S., Wai, T., Liao, T.-Y., Chang, H.-L., Chang, Y.-L., & Fu, L.-C. (2026). Mild cognitive impairment detection system based on unstructured spontaneous speech: Longitudinal dual-modal framework. JMIR Medical Informatics, 14, e80883.

Nikolaev, A., Chuang, Y.-Y., & Baayen, R. (2025). Case-number dissociation in Finnish noun embeddings: fastText vs. BERT layer effects.

Pistono, A., Jucla, M., Bézy, C., Lemesle, B., Le Men, J., & Pariente , J. (2019). Discourse macrolinguistic impairment as a marker of linguistic and extralinguistic functions decline in early Alzheimer’s disease. *Int J Lang Commun Disord*, *54*(3), 390–400. https://doi.org/10.1111/1460-6984.12444

Segkouli, S., Gkioka, M., Kokkas, S., Votis, K., Valero, S., Miguel, A., Antoniades, A., Charalambous, E., & Manias, G. (2025). Linguistic markers in spontaneous speech: Insights into subjective cognitive decline. Healthcare, 13(22), 2888.

Valles, B. (2006). La cohesión y la coherencia en la conversación del paciente con demencia: un estudio discursivo. Rev. Asoc. Colomb. Gerontol. Geriatr, 20(1).

Zhang, Y., Natale, G., & Clouston, S. (2021). Incidence of mild cognitive impairment, conversion to probable dementia, and mortality. American Journal of Alzheimer's Disease & Other Dementias, 36.

Dynamic Scaling of Sentence Processing Effort in Hungarian Speakers with Aphasia: A Pupillometry Study

Tamás Káldi

Introduction and aims

Despite the critical role of attention in language processing and the high prevalence of attentional deficits in persons with aphasia (PWA) (Hula & McNeil, 2008), real-time attention allocation during sentence comprehension remains underexplored. While it is established that distinct linguistic phenomena—such as linguistic prominence, non-canonical word order, and syntactic complexity—demand varying degrees of attentional resources in neurotypical adults (Mishra, 2015), there is a dearth of knowledge regarding how these structural variations impact attention allocation in PWA. The current study addresses this gap in Hungarian—a language characterized by rich case-marking and flexible word order—using pupillometry to continuously track processing effort across three distinct linguistic conditions: sentences with narrow Focus and without it (F vs. NF), canonical and non-canonical word order (SVO vs. OVS), and subject and object center-embedded relative clauses (SR vs. OR).

We formulate several group-specific predictions for these conditions. Regarding Focus, given that prosodic prominence triggers attention allocation (Káldi & Babarczy, 2021) and that PWA remain sensitive to prolonged prosodic cues (Schneider et al., 2025), we predict increased attention allocation in F sentences for both neurotypical (NT) controls and PWA. For word order, while we expect NT adults to exhibit the standard cost associated with non-canonical OVS processing, we predict this effect will be absent in the PWA group due to established disruptions in case-marking processing (MacWhinney et al., 1991; Hanne et al., 2015). Finally, regarding syntactic embedding, we follow Chapman and Hallowell (2021) in predicting that while NT adults will exert greater cognitive effort for OR relative to SR sentences, this effect will be absent in PWA, potentially reflecting a failure to dynamically scale attention to increasing syntactic complexity.

Methods

Participants

22 PWA (M_{age} = 63.45, SD = 12.25, 11 women), and 20 older NT (M_{age} = 57.8, SD = 12.53, 14 women) participated. PWA were mild-to-moderate non-fluent patients in the chronic post-stroke phase.

Procedure & materials

Participants listened to Hungarian sentences across the three conditions while maintaining gaze on a central fixation circle. Stimuli included: (i) Focus vs. No-Focus (e.g., *A kertből egy kutya meg-szökött/szökött-meg* [‘A dog escaped from the garden’]), (ii) Word order (SVO: *A fiú meglátta az apát* vs. OVS: *Az apát meglátta a fiú* [‘The boy saw the father’]), and (iii) Relative clauses (SR: *A kislány, aki a fiút megölelte...* vs. OR: *A kislány, akit a fiú megölelt...* [‘The girl who/whom the boy hugged...’]). To ensure engagement (Chapman & Hallowell 2020), a comprehension question followed 50% of the trials after a short delay. Pupil size was continuously recorded during sentence presentation and the subsequent inter-stimulus interval.

Results

Time-course changes in pupil size were analyzed using Growth Curve Analysis (GCA). Generalized Additive Mixed Models were also used and yielded consistent results. For conciseness, only the GCA findings are reported here. The non-linear pupillary response was modeled using third-order orthogonal polynomials (linear, quadratic, and cubic) to capture dynamic effort allocation. The mixed-effects models included fixed effects for Group, Condition, and their interactions with the time terms. The random effect structure included maximal random slopes for participants across all time terms and random intercepts for items.

Focus (F vs. NF)

When processing NF sentences, the NT group shows a steady increase in pupil size over time. This increase is further enhanced in F sentences, as evidenced by a steeper and more pronounced curve. PWA match the neurotypical pattern in overall pupil size, successfully scaling up their effort for F sentences. However, their cognitive effort trajectory differs significantly from controls; when shifting from NF to F sentences, the PWA group displays a uniquely altered, sharply peaked trajectory of effort over time.

Word-order (SVO vs. OVS)

In the NT group, processing SVO sentences produces a curved, non-linear hump in pupil size, typically associated with attention allocation during sentence processing. When processing OVS sentences, this group exhibits a significantly larger overall pupil size and a steeper rate of growth over time. In contrast, PWA fail to show this distinct response to word order variation. Although their change in pupil size reflects the attentional effort generally associated with linguistic processing, their pupillary response to OVS sentences does not significantly differ from their response to SVO sentences.

Relative clauses (SR vs. OR)

When processing SR sentences, the NT group displays a steady increase in pupil size. In contrast, PWA process these same sentences with a significantly lower overall pupil size and a flatter growth rate. Interestingly, when sentences become structurally more complex with OR clauses, the NT group does not change their overall average pupil size. Conversely, the PWA group responds to this increased complexity: when shifting from SR to OR sentences, the PWA group exhibits a significantly larger overall pupil size and a steeper rate of initial growth.

Discussion

Overall, these findings demonstrate that PWA retain the fundamental capacity to allocate attentional effort during auditory sentence processing. However, this modulation is contingent upon whether the linguistic features of the stimulus can be successfully integrated. Specifically, when prosodic prominence triggers attention (Focus), PWA recruit additional resources in a manner similar to neurotypical speakers, supporting evidence of preserved prosodic sensitivity (Schneider et al., 2025).

In contrast, the absence of an attentional shift for non-canonical word order suggests that impairments in morphosyntactic integration—specifically the inability to process case-marking (MacWhinney et al., 1991)—preclude the dynamic scaling of effort. Regarding

syntactic complexity, while a "plateau effect" has been suggested for high-complexity structures (Chapman & Hallowell, 2021), our results reveal intriguing cross-linguistic contradictions. The lack of an OR-specific cost in the NT group, alongside the unexpected effort increase observed in PWA, diverges from English-based literature. We speculate that these discrepancies arise from language-specific properties of Hungarian, where the interaction of an overt case system and free word order modulates cognitive load differently than in more rigid-order languages. These results underscore the necessity of cross-linguistic perspectives in modeling the interaction between attention and syntax in aphasia. Interpretation of these results is ongoing.

References

Chapman, L. R., & Hallowell, B. (2020). Expecting Questions Modulates Cognitive Effort in a Syntactic Processing Task: Evidence From Pupillometry. *Journal of Speech Language and Hearing Research*, *64*(1), 121–133.

Chapman, L. R., & Hallowell, B. (2021). The Unfolding of Cognitive Effort During Sentence Processing: Pupillometric Evidence From People With and Without Aphasia. *Journal of Speech, Language, and Hearing Research*, *64*(12), 4900–4917.

Hanne, S., Burchert, F., De Bleser, R., & Vasishth, S. (2015). Sentence comprehension and morphological cues in aphasia: What eye-tracking reveals about integration and prediction. *Journal of Neurolinguistics*, *34*, 83–111.

Hula William D. & McNeil Malcolm R. (2008). Models of attention and dual-task performance as explanatory constructs in aphasia. *Seminars in Speech and Language*, *29*(3), 169–187.

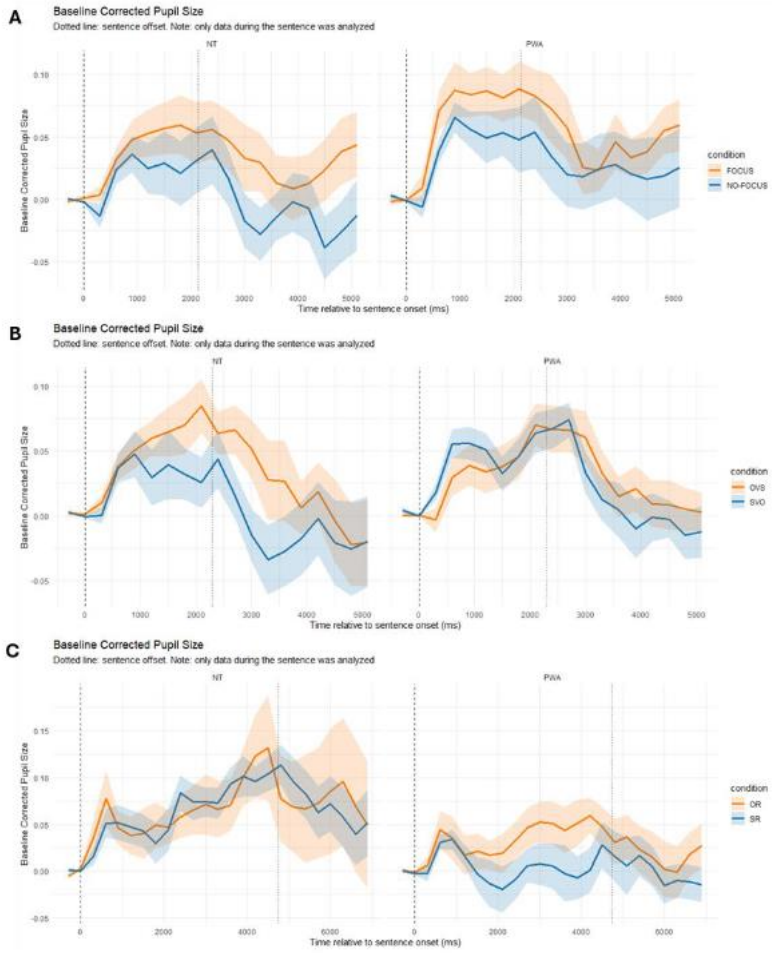
Káldi, T., & Babarczy, A. (2021). Linguistic focus guides attention during the encoding and refreshing of Working Memory content. *Journal of Memory and Language*, *116*, 104187.

MacWhinney, B., Osmán-Sági, J., & Slobin, D. I. (1991). Sentence Comprehension in Aphasia in Two Clear Case-Marking La nguages. Brain and Language, 41(2), 234–249.

Mishra, R. K. (2015). Attention and the Processing of Sentences. In R. K. Mishra (Ed.), *Interaction Between Attention and Language Systems in Humans: A Cognitive Science Perspective* (pp. 89–104). Springer India.

Schneider, K., Tuomainen, O., Wartenburger, I., & Hanne, S. (2025). Prosodic cue use for identifying interrogative and declarative structures with wide and narrow focus in individuals with and without aphasia. *Frontiers in Communication*, *10*.

Figure 1. Time-course of baseline-corrected pupil size for persons with aphasia (PWA) and neurotypical (NT) controls. Panel A: Linguistic focus; Panel B: Word order; Panel C: Embedding. Ribbons represent ±1 SE.



Hidden in the treetop: Moved and non-moved prepositions in agrammatism

Yuval Z. Katz | Naama Friedmann

Introduction

Aphasiological research has historically focused on two primary symptoms associated with agrammatism, often treating them as one: errors in functional morphemes, and errors in word order, particularly role-substitution in sentences with non-canonical word order. This study aims to provide a syntactic characterization of the agrammatic impairment in relation to these two symptoms and their interaction by testing the grammaticality judgement of moved and non-moved prepositions.

Methods

Ten Hebrew-speakers with agrammatic aphasia (five post-CVA, five PPA) and ten non-impaired controls completed a grammaticality judgment task consisting of 48 sentences, half grammatical and half ungrammatical, manipulated for whether the prepositional phrase appeared in its original position (in situ) or had undergone movement to the CP zone of the sentence (question formation, topicalization, or relativization), and whether the preposition was syntactically selected or unselected by the verb. To minimize effects of memory and other processing limitations, participants were allowed to reread the sentences or ask the experimenter to repeat them.

Results

Controls performed at ceiling across all conditions. In contrast, individuals with agrammatism were able to identify incorrect prepositions when they were in situ (M = 94%), but showed significantly reduced accuracy (M = 67%) when the prepositional phrase was moved to the CP layer ($t(9) = 5.79$, $p < .001$). This pattern was also consistent across participants individually. Performance in grammatical sentences was high both without movement (M = 98%) and with movement (M = 93%), with a marginally significant difference between them ($t(9) = 2.27$, $p = .049$). There was no difference in error rate between selected and unselected prepositions in detecting the ungrammatical sentences with movement ($t(9) = 0.39$, $p = .7$).

Discussion

Participants detected ungrammaticality when the preposition remained in situ but failed to do so when it was moved. Most previous research syntactic movement in perception has focused on the sentence-picture matching task and a specific error type: role substitution between the moved element and the intervening subject DP (Caplan & Futter, 1986; Caramazza & Zurif, 1976; Grodzinsky, 1989; Piñango, 2000; Schwartz et al., 1980). The fact that participants cannot detect the ungrammaticality of the moved preposition suggests that the CP zone in agrammatism is radically inaccessible, not only to thematic assignment.

Participants' success in detecting the ungrammaticality of prepositions in situ challenges the view that agrammatism necessarily entails a general impairment in functional morphemes, as in some non-syntactic accounts of agrammatism (e.g., Kean, 1977), or that agrammatism necessarily involves an impairment to argument structure.

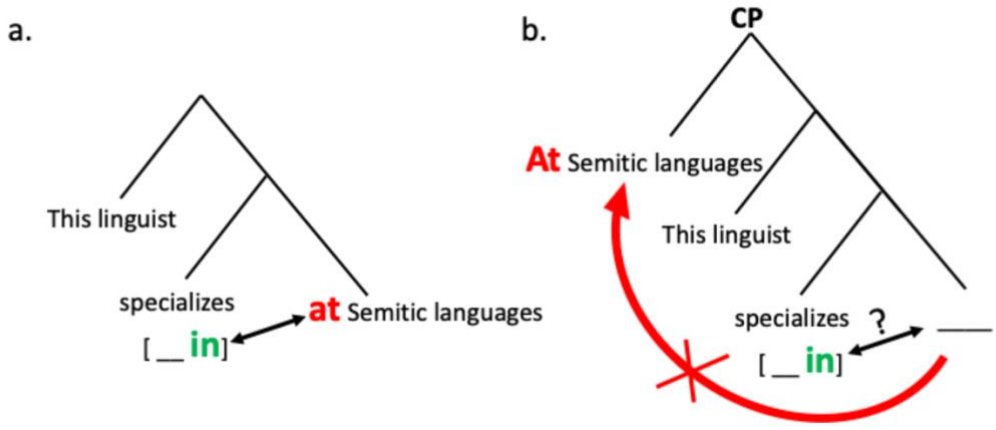
These findings suggest that the difficulty lies in constructing syntactic dependencies involving the CP, as predicted under the Tree Pruning Hypothesis (Friedmann, 2002, 2006; Friedmann & Grodzinsky, 1997). Selected and unselected prepositions require licensing within the verb phrase. When in situ, agrammatic participants without an argument structure impairment can check the preposition's congruence against the verb's subcategorization information (Figure 1a). When displaced to the CP, individuals with agrammatism cannot establish a link between the preposition and its original post-verbal position, which prevents them from evaluating the preposition's congruence with the verb (Figure 1b).

References

Friedmann, N., & Grodzinsky, Y. (1997). Tense and agreement in agrammatic production: Pruning the syntactic tree. *Brain and Language, 56*(3), 397–425. https://doi.org/10.1006/brln.1997.1795
Grodzinsky, Y. (1989). Agrammatic comprehension of relative clauses. *Brain and Language, 37*(3), 480–499. https://doi.org/10.1016/0093-934X(89)90031-X
Kean, M.-L. (1977). The linguistic interpretation of aphasic syndromes: Agrammatism in Broca's aphasia, an example. *Cognition, 5*(1), 9–46. https://doi.org/10.1016/0010-0277(77)90015-4
Schwartz, M. F., Saffran, E. M., & Marin, O. S. (1980). The word order problem in agrammatism: I. Comprehension. *Brain and language, 10*(2), 249-262. https://doi.org/10.1016/0093-934X(80)90055-3

Figure 1
Detection of incorrect prepositions in situ and moved, in agrammatic aphasia.

In (a) the prepositional phrase follows the verb and is therefore available in agrammatism. Individuals with agrammatism can compare the preposition at used in the sentence to the preposition in listed in the lexical entry of specialize, and therefore they can identify that the sentence is ungrammatical. In (b), on the other hand, the prepositional phrase moves to the CP zone. Patients with agrammatism do not establish a chain between the prepositional phrase and its original position following the verb, and therefore they do not know that the preposition used in the sentence is not the one listed in the lexical entry of the verb.



The multilingual Token Test App

Dörte de Kok

Introduction and aims

The shortened version of the Token Test (De Renzi & Faglioni, 1987) is one of the most widely used tests to diagnose aphasia. Recently, we developed a tablet-based version that comprises more than 40 languages (Bastiaanse et al., 2020). This tool allows clinicians to test multilingual people with suspected aphasia, even if they don't speak (all of) the languages that need to be investigated.

However, this means that normative data for all included languages need to be provided. For many of the languages, norms for paper-and-pencil-versions do exist. These offline tests include the exact same items but are administered in a different manner. A pilot of Akinina et al. (2019) implied that the versions are comparable, but a Dutch pilot (Visser, 2022) found a trend to higher scores for the app version in a group of aphasic participants. The current study explores the difference between tablet-based and paper-pencil testing outcomes for Dutch by comparing the existing Dutch paper-pencil norms (Van den Berg et al., 2021) to newly collected norms for the tablet application.

Due to the large number of languages, collecting normative data for all languages is hardly possible. Therefore, we also explore the comparability of non-brain-damaged speakers' scores across languages, with a Dutch, a Russian (Akinina et al., 2019) and a mixed group.

Finally, the normative data of native speakers cannot be used for second-language (L2) speakers. We, therefore, also explore the scores of non-brain-damaged healthy L2-speakers of Dutch in the Dutch version of the test and in their respective native version.

Methods

The Token Test app consists of the same 36 items as described in the original version of De Renzi and Faglioni (1987). Items consist of short commands, ranging from simple (e.g. "Touch a circle") to complex (e.g. "Instead of the white square touch the yellow circle"). Depending on the subpart, participants have to choose between 10 or 20 shapes: squares and circles in five different colors and (for the parts with 20 shapes) small vs. large shapes.

For the current study, we collected data from 75 native Dutch participants without aphasia, 5 participants with (mild to severe) aphasia and 15 people with varying first languages and Dutch as a second language (all demographic data is provided in Table 1). All three groups carried out the Dutch version of the Token Test app. The third group was also tested in their respective native version, starting with Dutch and having at least one week in-between test moments. These datasets were collected as parts of masters' thesis projects of Elise Corba, Anne Verhoeven, Carla Visser and Noah van den Hoeven.

Results

Scores per group can be found in Table 1.
An independent t-test revealed that the Dutch native speakers without aphasia scored

significantly higher on the app than the group of Van den Berg and colleagues (2021) on the paper-pencil-version, $t(146) = 2.30$, $p = 0.023$, $d = .38$.

A one-way ANOVA was used to compare the results of the three non-brain damaged groups tested with the app: the Dutch group, the Russian group (Akinina et al., 2019) and the mixed group. There was no significant difference between these three groups, $F(2, 196) = 0.6590$, $p = .519$, $\eta^2p = 0.007$. Post-hoc Tukey comparisons verified that there were no differences between any groups (Dutch-Russian: $p = .486$, Dutch-mixed group: $p = .9233$, Russian-mixed group: $p = .9701$). An independent t-test comparing the app results from Russian aphasia group and the Dutch aphasia group showed no significant differences, $t(70) = -0.1937$, $p = .847$, $d = 0.009$.

Taking a closer look at the multilingual group, Mann-Whitney-U tests revealed that the multilingual group scored significantly lower in Dutch than the native Dutch group, $U = 121$; $p < .001$. Their performance was significantly correlated to their self-rated listening skills in Dutch (Pearson-correlation), $r(13) = 0.663$, $p = .007$.

Discussion

Comparing the results of the Dutch app data with existing normative data for the paper-and-pencil-version (Van den Berg et al., 2021) showed that the norms from the paper and pencil version cannot be directly applied to the app version. The differences were small but significant. It should be noted that the Van den Berg group is on average somewhat older, which could be at least a partial explanation for this difference. A more direct comparison between the app and the paper-and-pencil-version is needed to verify the difference we found. This should include a group of people with aphasia, who will show less of a ceiling effect.

Comparing the scores of people without aphasia in different language versions of the app showed that all three groups (Dutch, Russian, mixed languages) scored comparably. Also, the Dutch and Russian aphasia groups did not differ significantly. This last result should be taken cautiously, as the Dutch aphasia group is very small. To confirm these promising results, the Dutch aphasia group needs to be extended, and additional languages should be added. If we can replicate the current pilot data, we can provide language-independent norms for the Token Test app. This would be of great value, as performing norming studies for all languages is time-consuming and for some of the languages not possible.

As expected, the L2-speakers (without aphasia) score significantly lower than L1-speakers. The norms therefore cannot be applied to them. However, the strong correlation of their scores with their self-rated language level provides an interesting opening. With a larger group, an algorithm could be devised that can help to generate relative norms for different (pre-aphasia) language levels.

The Token Test app can help clinicians to diagnose people in a language they do not speak themselves. The comparable scores across languages can facilitate this. Even if there are no norm scores available for the L1 of a client, clinicians can compare their L1 score to the (to be developed) “universal norm score” and determine whether the client presents with aphasia.

References

Akinina Yu., Buivolova O., Soloukhina O., Bastiaanse R. (2019) Psychometric properties of the Token Test App. *Stem-, Spraak- en Taalpathologie, 24,* 74-76.
Bastiaanse, R., Akinina, Y., Satoer, D., & De Kok, D. (2020). *The Token Test App* (Second Edition) [Software]. University of Groningen.
De Renzi, E., & Faglioni, P. (1978). Normative Data and Screening Power of a Shortened Version of the Token Test. *Cortex*, *14*(1), 41–49. https://doi.org/10.1016/S0010-9452(78)80006-9
Van den Berg, T. C., El Hachioui, H., Heijenbrok-Kal, M. H., Van Harskamp, F., & Visch-Brink, E. G. (2021). Nederlandse normering van de 36-item versie van de Token Test: Bepaling van de cut-off score, ernstindeling en kritieke waarde. *Nederlands Tijdschrift Voor Logopedie*, *93*(6), 20–27.
Visser, C.-H. (2022). *De Token Test app: vergelijkbaar met de papieren versie?* [Unpublished masters' thesis]. University of Groningen.

	Gender	Age	Education	Score Dutch (max. 36)	Score native language (max. 36)
No aphasia, Dutch = L1 (n=75)	32 male 43 female	18 - 94 M = 46.3 SD = 19.6	Low: 12 Middle: 21 High: 36	28 - 36 M = 34.55 SD = 1.26	NA
Aphasia, Dutch = L1 (n=5)	3 male 2 female	68 - 86 M = 75 SD = 7.3	unknown	9 - 33 M = 20.25 SD = 12.53	NA
No aphasia, Dutch = L2, mixed L1 (n=15)	6 male 9 female	30 - 58 M = 41.8 SD = 7.33	unknown	23 - 35 M = 31.1 SD = 3.06	32 - 36 M = 34.7 SD = 1.1
Van den Berg et al. (2021), Dutch, no aphasia, paper-pencil (n=73)	25 male 48 female	19 – 93 M = 66.96 SD = 15.74	Low: 27 Middle: 16 High: 30	M = 34.04 SD = 1.43	NA
Akinina et al. (2019), Russian, no aphasia (n=109)	45 male 64 female	18-75 M = 41 SD = 18.39	unknown	NA	27 - 36 M = 34.79 SD = 1.51
Akinina et al. (2019), Russian, aphasia (n=67)	unknown	unknown	unknown	NA	0.5 - 36 M = 19.32 SD = 10.21

Note: For age and score, range, mean (M) and standard deviations (SD) are provided. Level of education according to Dutch Central Bureau of Statistics: https://www.cbs.nl/nl-nl/longread/discussion-papers/2021/invulling-praktisch-en-theoretisch-opgeleiden/3-indeling-van-opleidingen-op-basis-van-niveau-en-orientatie

Table 1

Psycholinguistic effects on word-level comprehension: results from a novel web-based word-picture matching task

Alycia Laks | Yeonwoo Kim | Catherine Kelly | Andrew DeMarco | Peter Turkeltaub

Introduction and aims

Aphasia occurs in 21-40% of stroke survivors (Cichon et al., 2021), with around 68% of those individuals also having alexia (Brookshire et al., 2014). Both disorders can greatly impact quality of life. Access to well-designed assessment measures is essential for clinicians to determine patterns of deficits needed for treatment planning. Word-to-picture matching (WPM) tasks are often used to assess single-word comprehension deficits in standardized assessments. Despite being commonplace, available tests are limited by different factors. Most available tests do not have paired auditory and written versions, which is necessary to determine modality specific impairments versus difficulty with task demands alone. Standardized assessments only provide accuracy, which can miss subtle impairments captured in response times (RT), particularly in those with very mild aphasia (Laks et al., 2025). Finally, these tasks do not have careful psycholinguistic manipulations, which can impact performance in typical and impaired readers, and can be used to diagnose alexia syndromes (Coslett & Turkeltaub, 2016).

Here we aimed to develop a novel WPM task that addressed those concerns. We created a web-based version that can be deployed clinically. We collected a normative sample to determine expected performance and a sample of left-hemisphere stroke survivors (LHSS) to determine patterns of impairment seen in aphasia and alexia.

Methods

Participants

Participants were recruited as part of the ReadMap (ClinicalTrials.gov ID: NCT06700005) study. Participants learned English before age eight and had no major psychiatric or neurological disorders. Participants included 63 left hemisphere stroke survivors (LHSS; age: 59.73(15.53); education: 16.42(2.70) race: Asian/Black/White/More than one 3/25/35/0) and 62 age and education-matched controls (age: 59.21(14.00); education: 16.89(2.11); race: Asian/Black/White/More than one 2/24/33/3).

Procedure

As part of an assessment battery, participants completed the paired written and auditory WPM match tasks on a touchscreen computer. The WPM tasks were four-alternative forced-choice design. Participants were presented with a prime word and asked to select the prime from four pictures. There were 48 trials per input modality (written or auditory). The monosyllabic targets were manipulated for frequency (Brysbaert & New, 2009) and regularity (Chee et al., 2020), resulting in four groups of word types: irregular/high frequency, irregular/low frequency, regular/high frequency, and regular/low frequency. The groups of foils were matched to the targets within their word type on frequency, length in phonemes and graphemes, and semantic distance.

Statistical Analysis

Chi-square tests of independence and two-tailed independent-samples t-tests were utilized to determine group differences in demographic variables. We examined accuracy, RTs, and

efficiency (accuracy/medianRT on correct trials) in controls and LHSS. A repeated measures ANOVA compared efficiency between groups with age and education as covariates. We used Pearson's correlations to compare performance between WPM modalities and partial correlations to determine the relationship between written WPM, oral reading, and lexical decision of the same 48 words, controlling for auditory WPM. Next, we z-scored LHSS efficiencies against controls, after regressing out age and education level. Finally, a k-means clustering analysis was used on mean-centered z-scores for each word type, in each modality, to determine patterns of errors in the LHSS.

Results

There were no group differences in demographic variables of age ($t(123) = -0.20$, $p = .84$), education ($t(122) = 1.08$, $p = .28$), or race (X^2 (3, N = 125) = 3.27, $p = .35$). Controls were at ceiling (mean written accuracy=98%, auditory=98%), while performance was impaired in LHSS (mean written=88%, auditory=91%). WPM performance was modulated by modality ($F(1,121) = 4.82$, $p = .03$, auditory>written), frequency ($F(1,121) = 16.8$, $p < .001$, high > low), and regularity ($F(1,121) = 10.6$, $p = .001$, regular>irregular in controls). Written and auditory versions were correlated in both controls and LHSS (controls: $r = 0.85$, $p < .001$; LHSS: $r = 0.85$, $p < .001$). Written WPM was correlated to written lexical decision ($r = 0.49$, $p < .001$) and oral reading ($r = 0.32$, $p = .011$) in LHSS. LHSS z-scores had a mean of -2.02 on auditory and -1.92 on written WPM. 54 (86%) LHSS group scored at least one standard deviation below the control mean in either modality, including 16/23 LHSS with very mild aphasia (≥93.8 AQ on WAB-R). The k-means clustering analyses resulted in four clusters of deficit patterns (Figure 1): modestly greater deficits on auditory vs. written, modestly greater deficits on written vs. auditory, severe deficits in auditory and not written, and high frequency regular words spared in both modalities. Qualitative review of z-scores demonstrated that some LHSS had other performance patterns suggesting process-specific comprehension deficits (Figure 1).

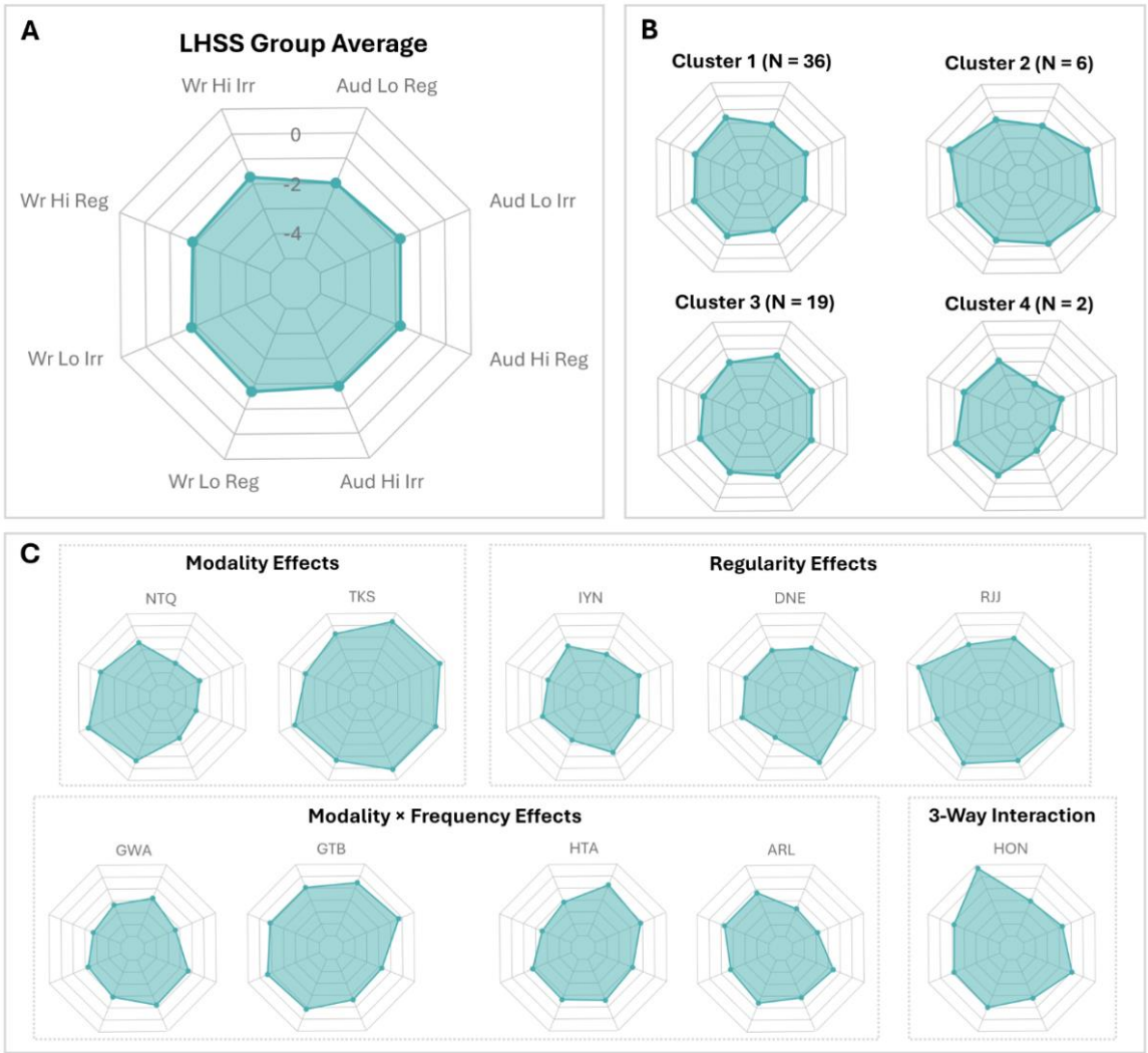
Discussion

Our novel web-based word-to-picture matching task captures group and individual patterns of word comprehension deficits in written and auditory modalities. By incorporating both RTs and accuracy, efficiency scores are able to detect even subtle impairments, which are missed by ceiling effects on common standardized tests. We found that modality, frequency, and regularity affect word comprehension on WPM. These patterns inform the process-level deficits leading to word comprehension difficulties in individual people, which can be used to determine treatment planning. Limitations include the use of imageable monosyllabic words only, which was required for psycholinguistic manipulations, and selection of picture targets. Additionally, RTs will be impacted by the administration device, which may impact the interpretation of results. Future directions include the deployment of this task for widespread use, and the use of other comprehension tasks, such as semantic relatedness judgment, to assess low imageability and multisyllabic words.

Selected References

Brookshire, C. E., Wilson, J. P., Nadeau, S. E., Gonzalez Rothi, L. J., & Kendall, D. L. (2014). Frequency, nature, and predictors of alexia in a convenience sample of individuals with chronic aphasia. *Aphasiology*, *28*(12), 1464–1480. https://doi.org/10.1080/02687038.2014.945389

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*(4), 977–990. https://doi.org/10.3758/BRM.41.4.977
Chee, Q. W., Chow, K. J., Yap, M. J., & Goh, W. D. (2020). Consistency norms for 37,677 english words. *Behavior Research Methods*, *52*(6), 2535–2555. https://doi.org/10.3758/s13428-020-01391-7
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), 3778. https://doi.org/10.3390/jcm10173778
Coslett, H. B., & Turkeltaub, P. (2016). Acquired Dyslexia. In *Neurobiology of Language* (pp. 791–803). Elsevier. https://doi.org/10.1016/B978-0-12-407794-2.00063-8
Laks, A. B., DeMarco, A. T., Hurwitz, S. S., & Turkeltaub, P. E. (2025). Rapid Lexical Retrieval and Executive Functioning in Very Mild Aphasia. *American Journal of Speech-Language Pathology*, 1–13. https://doi.org/10.1044/2025_AJSLP-25-00345



Disfluencies in the production of Czech prepositional phrases

Michal Láznička

Introduction and aims

Apart from their pragmatic functions, disfluencies in speech reflect cognitive demands of lexical and constructional retrieval, grammatical processing, or utterance planning. This is evidenced by the fact that words that are more frequent or words that are more predictable given the context generally trigger fewer disfluencies (Goldman-Eisler, 1958; Beattie & Butterworth, 1979; Schneider, 2016). The occurrence of disfluencies in connected speech of persons with aphasia has been studied from two main perspectives: One line of research has focused on the distributional properties of pause durations and their occurrence with regard to different structural factors, such as part of speech or clausal structure (Angelopoulou et al., 2018; DeDe et al., 2025), while other researchers focused on the role of word frequency (Goral et al., 2010; Mack et al., 2015). In a recent analysis, Bello-Lepe et al. (2024) extended the perspective beyond single word level by using collocational strength to explain pause duration in the speech of English speakers with and without aphasia. Together with proportional and combinatorial frequency, association (collocational) strength is one of corpus-based measures introduced to the study of aphasia by usage-based analyses (Gahl & Menn, 2016).

The aim of this paper is to extend this approach by introducing dispersion (distribution of words across contexts) as an important usage-based measure. We examine these effects in Czech prepositional phrases (PPs), comparing how usage-based measures predict disfluencies in aphasia and typical speech.

Methods

We analysed PPs with common noun complements from corpora of Czech aphasia (487 tokens produced by 10 speakers) and spoken Czech (122k tokens produced by 2,641 speakers). PPs were annotated for fluency (fluent or disfluency before or within PP), presence of premodifiers, and presence of inflectional marker on complement. Additionally, usage-based measures included: frequency and dispersion of PP and complement, and association strength between preposition and complement. Dispersion and association were quantified with adjustments for frequency to tackle the interdependence of traditional measures (Gries, 2024). The corpus measures were orthogonalised using PCA, revealing “commonness” as a composite of frequency and dispersion. Mean number of fluent word trigrams per utterance was included for the aphasia corpus data to account for baseline fluency in aphasia and individual differences. Multinomial logistic regression was used to assess the influence of these variables on fluency.

Results

Both datasets showed main effects of commonness and association, and modification. PPs that are less common trigger more disfluencies. Additionally, disfluencies are more likely to occur within PPs, when the complement is premodified and when the association between preposition and complement is weak. Importantly, these lexical effects in aphasia held even when controlling for individual baseline fluency. Furthermore, modification interacted with the usage-based variables in the spoken corpus, reducing their effects. The effects of the

usage-based variables were comparable for both groups, particularly for disfluencies within PPs. Morphological complexity of the complement (difference from citation form) did not have any effect on disfluency patterns. See Table 1 for full summary.

Discussion

Consistent with usage-based linguistics, PPs that are used frequently and in diverse contexts cause fewer disfluencies overall, pointing to a higher level of entrenchment (e.g. Schmid, 2016). PPs with high degree of association may be stored and retrieved as chunks rather than being assembled compositionally (Bybee, 2010), allowing less room for being interrupted internally. When a modifier is introduced into a PP, its internal complexity and length increases and the connection between the preposition and the complement is "loosened", resulting in diminished effects of the usage variables.

Despite much higher disfluency rates in aphasia, these effects were highly similar across the two datasets, suggesting that the same domain-general mechanisms are involved in language representation and processing with differences being quantitative rather than qualitative, as expected under a usage-based analysis (Gahl & Menn, 2016). This analysis shows how integrating usage-based principles in aphasia research provides a more nuanced picture of linguistic behaviour, complementing rule-based, structure-oriented approaches.

References

Angelopoulou, G., Kasselimis, D., Makrydakis, G., Varkanitsa, M., Roussos, P., Goutsos, D., Evdokimidis, I., & Potagas, C. (2018). Silent pauses in aphasia. *Neuropsychologia*, *114*, 41–49. https://doi.org/10.1016/j.neuropsychologia.2018.04.006

Beattie, G. W., & Butterworth, B. L. (1979). Contextual Probability and Word Frequency as Determinants of Pauses and Errors in Spontaneous Speech. *Language and Speech*, *22*(3), 201–211. https://doi.org/10.1177/002383097902200301

Bello-Lepe, S., Mahmood, S., Varley, R., & Zimmerer, V. (2024). Speech pauses in speakers with and without aphasia: A usage-based approach. *Cortex*, *178*, 287–298. https://doi.org/10.1016/j.cortex.2024.06.012

Bybee, J. (2010). *Language, Usage and Cognition.* Cambridge University Press. http://public.eblib.com/choice/publicfullrecord.aspx?p=534754

DeDe, G., Salis, C., & Bose, A. (2025). Discourse planning and production in post-stroke aphasia: The role of executive function on pauses and clauses. *Aphasiology*, *0*(0), 1–24. https://doi.org/10.1080/02687038.2025.2534987

Gahl, S., & Menn, L. (2016). Usage-based approaches to aphasia. *Aphasiology*, *30*(11), 1361–1377. https://doi.org/10.1080/02687038.2016.1140120

Goldman-Eisler, F. (1958). The Predictability of Words in Context and the Length of Pauses in Speech. *Language and Speech*, *1*(3), 226–231. https://doi.org/https://doi.org/10.1177/002383095800100308

Goral, M., Levy, E., Swann-Sternberg, T., & Obler, L. (2010). Frequency and Word-Length Factors and Lexical Retrieval in Sentence Production in Aphasia. *Procedia - Social and Behavioral Sciences, AoA2010*, *6*, 107–108. https://doi.org/10.1016/j.sbspro.2010.08.054

Gries, S. T. (2024). Frequency, Dispersion, Association, and Keyness. In *Scl.115*. John Benjamins Publishing Company. https://benjamins.com/catalog/scl.115

Mack, J. E., Chandler, S. D., Meltzer-Asscher, A., Rogalski, E., Weintraub, S., Mesulam, M.-M., & Thompson, C. K. (2015). What do pauses in narrative production reveal about the

nature of word retrieval deficits in PPA? *Neuropsychologia*, *77*, 211–222. https://doi.org/10.1016/j.neuropsychologia.2015.08.019
Schmid, H.-J. (Ed.). (2016). *Entrenchment and the psychology of language learning: How we reorganize and adapt linguistic knowledge* (First Edition). American Psychological Association.
Schneider, U. (2016). Hesitation placement as evidence for chunking: A corpus-based study of spoken English. In H. Behrens & S. Pfänder, *Experience Counts: Frequency Effects in Language* (pp. 61–90). De Gruyter. https://doi.org/10.1515/9783110346916-004

Predictor	disfluency before PP	disfluency within PP	comment
modification	0.16 / 0.16***	1.28*** / 1.44***	modified PPs likely to be interrupted
commonness	-0.26** / -0.06***	-0.55*** / -0.33***	common PPs more fluent
association	-0.07 / -0.06***	0.30* / 0.28***	strong association protects against interruption of PP
baseline fluency	-0.39† / NA	-0.61* / NA	more fluent participants produce fewer disfluencies within PPs
modification : commonness	na / 0.07***	na / 0.12***	modification attenuates usage-based variables
modification : association	na / 0.06*	na / -0.19***	

Table 1: Effect sizes (β) across datasets and disfluency types; format: aphasia / spoken corpus, *** p<0.001, ** p<0.01, * p<0.05, † p<0.1

Behavioral and neurophysiological effects of transcranial direct current stimulation to the left inferior frontal gyrus in patients with expressive aphasia

Silvia Marchesotti | Naz Doganci | Camille Farcy | Marion Corre | Cynthia Dery | Radek Ptak | Adrian G. Guggisberg

Introduction and aims

Transcranial direct current stimulation (tDCS) has shown promising results in alleviating word-finding difficulties in patients with aphasia (Marangolo, 2020; Matar et al., 2022; Ilardi et al., 2026). In healthy individuals, previous evidence from our group has shown that anodal high-definition tDCS applied to the left inferior frontal gyrus (IFG) facilitates verb learning (Farcy et al. 2024). Crucially, this improvement was associated with enhanced functional connectivity (FC) between the left IFG and the temporo-parietal junction (TPJ), two key regions involved in language processing, whose damage often leads to aphasia. Based on this evidence, we aimed to investigate whether such improvement in naming following tDCS and its link with network interactions could also be observed in patients with aphasia.

Methods

For this, we conducted a crossover, randomized, sham-controlled clinical trial in a group of fourteen chronic post-stroke patients with expressive aphasia. Each patient underwent both sham and real tDCS stimulation to the left IFG in two separate sessions, spaced one month apart (washout period). Each stimulation condition was administered twice over two consecutive days and was coupled with computerized verb learning. Resting-state electroencephalography (EEG) activity was recorded, and naming abilities were assessed using a verb-naming task and a connected speech task (Nicholas & Brookshire 1993) on the day before, the day after, and one month after the stimulation sessions.

Results

In the naming task, tDCS significantly improved verb naming performance, regardless of whether the verbs had been trained during the computerized task. This effect was observed immediately after stimulation but was not maintained at the one-month follow-up. In contrast, sham stimulation led to improvement only for trained verbs. The magnitude of improvement following tDCS showed a modest, yet non-significant, correlation with baseline performance, suggesting a possible greater effect in patients with milder impairments. For narrative speech, verb fluency improved over time independently of stimulation condition, likely reflecting a training effect. In parallel, global fluency (number of words/min) decreased during the day after the stimulation sessions. No significant effects were observed for other connected speech measures, such as the percentage of Correct Information Units (%CIU) and informative fluency (CIU/min).

Preliminary resting-state analyses of global FC indicate that tDCS, compared to sham stimulation, induced changes in regions beyond those directly stimulated. These included an increase in FC in the alpha band (8–13 Hz) over posterior regions and a decrease in FC in the theta band (4–7 Hz) over the right frontal region (near the homologue of Broca's area).

Discussion

Overall, these findings support the effectiveness of tDCS applied to the IFG in improving verb naming in patients with chronic expressive aphasia and suggest that the underlying mechanism of action may involve widespread, interhemispheric changes in functional connectivity.

References

Ilardi, C.R., Federico, G., Alfano, V., Basso, L., De Luca, M., Caltagirone, C. and Marangolo, P., 2026. Enhancing verb naming with anodal tDCS over the left inferior frontal gyrus: Insights from lesion mapping and structural connectivity estimation. Cortex.

Farcy, C., Chauvigné, L.A., Laganaro, M., Corre, M., Ptak, R. and Guggisberg, A.G., 2024. Neural mechanisms underlying improved new-word learning with high-density transcranial direct current stimulation. NeuroImage, 294.

Marangolo, P., 2020. The potential effects of transcranial direct current stimulation (tDCS) on language functioning: Combining neuromodulation and behavioral intervention in aphasia. Neuroscience Letters, 719.

Matar, S.J., Newton, C., Sorinola, I.O. and Pavlou, M., 2022. Transcranial direct-current stimulation as an adjunct to verb network strengthening treatment in post-stroke chronic aphasia: a double-blinded randomized feasibility study. Frontiers in Neurology, 13.

Nicholas, L.E. and Brookshire, R.H., 1993. A system for quantifying the informativeness and efficiency of the connected speech of adults with aphasia. Journal of Speech, Language, and Hearing Research, 36(2).

Elaboration of events in a picture description task: a case of Croatian and Norwegian PWA

Ana Matić | Monica I. Norvik | Jelena Kuvač Kraljević | Hanne Gram Simonsen

Introduction and aims

Production of a coherent discourse is a complex task that involves the interaction of language, cognitive and social-emotional skills. The integration of these skills is also essential for introducing and maintaining the main events in a discourse, which are critical proxies for functional communication. In this study we mainly focus on the event elaboration, as the first step into deeper analyses in the future.

A previous study that compared the performance of Croatian and Norwegian PWA (Matić Škorić et al., 2026) found differences in performance between the two groups on the adapted versions of the Comprehensive Aphasia Test (CAT-HR and CAT-N; Swinburn et al., 2020; Swinburn et al., 2021), with the differences mainly being caused by differences in item difficulty, which reflected the different structures of the two languages and socio-demographic differences between the two groups. That study excluded the Picture description task, so it was decided to analyze it in a separate step. Therefore, the aim of this study is to examine the elaboration of the four events in a picture description task in Croatian and Norwegian PWA.

Methods

For the purpose of this study, we used the Picture description task from the CAT. Picture description represents a consistent referent and situation, which is highly suitable for comparable analyses among various groups regarding different demographic and clinical variables (MacKenzie et al., 2007). We analyzed the discourse production of 40 PWA (Nper language = 20) matched in severity and demographic factors as much as possible between the two groups. The samples have been collected as part of the test standardization process in both countries. We adopted the four main events from the CAT test (the father is sleeping/ the girl is waking up the father/ the cat is catching the fish or making a mess/ the books are falling off the shelf). In the CAT test they are only analyzed on the level of their presence. In the current study we analyzed how PWA elaborate these events, i.e. how many concepts that they produce are related to each of the four events, as well as their completeness and accuracy.

Results

This is an ongoing study, with analyses being conducted in the moment of abstract submission. So far, it seems that PWA with milder symptoms mention at least two out of these four main events, but the frequency, completeness and accuracy of the concepts related to these events varies greatly. These preliminary findings demonstrate the relative complexity of detailed elaboration of events in naturalistic speech for PWA, but the impact of education, level of severity and language is yet to be analyzed.

Discussion

For the story to be complete and relevant, it has to include all the events in the picture. In this sense, the elaboration of events is one of the key components for establishing cause-

consequence relationships. This study addresses the elaboration of events as the first step towards a broader investigation into other components of narrative production analyses (e.g. linguistic and propositional levels from the LUNA framework; Dipper et al., 2021) and how they converge to support and facilitate (successful) communication in aphasia. By comparing Croatian and Norwegian PWA, this research builds upon the benchmarks established by different authors under different theoretical frameworks (e.g. Kuvač Kraljević et al., 2023; MacKenzie et al., 2007…). Testing this across different languages allows us to discern whether discourse deficits are universal or influenced by specific language typologies.

Ultimately, moving beyond the word-level to discourse-level analysis is vital for clinical practice, as it aligns speech-language therapy with the realistic demands of functional communication. The cross-linguistic approach adopted in the current study ensures that the research findings in the first step, and intervention strategies in the next, are robust, culturally sensitive, and focused on real-world communicative contexts.

Keywords: people with aphasia; discourse; elaboration of events; concepts; cross-linguistic and cross-cultural study

References

Dipper, L., Marshall, J., Boyle, M., Hersh, D., Botting, N., & Cruice, M. (2021). Creating a Theoretical Framework to Underpin Discourse Assessment and Intervention in Aphasia. *Brain sciences*, *11*(2), 183. https://doi.org/10.3390/brainsci11020183.

Mackenzie, C., Brady, M., Norrie, J., & Poedjianto, N. (2007). Picture description in neurologically normal adults: Concepts and topic coherence. *Aphasiology*, *21*(3-4), 340-354. https://doi.org/10.1080/02687030600911419

Matić Škorić, A., Norvik, M. I., Kuvač Kraljević, J. K., Røste, I., & Simonsen, H. G. (2026). Comprehensive Aphasia Test (CAT): Comparability of the Croatian and Norwegian versions. *Aphasiology, 40*(2), 301–327. https://doi-org.mime.uit.no/10.1080/02687038.2023.2250516

Richardson, J. D., & Dalton, S. G. (2016). Main concepts for three different discourse tasks in a large non-clinical sample. *Aphasiology*, *30*(1), 45-73. https://doi.org/10.1080/02687038.2015.1057891

Swinburn, K., Porter, G., Howard, D., Høeg, N., Norvik, M., Røste, I., & Simonsen, H. G. (2021). *CAT-N – Comprehensive Aphasia Test*. Novus Forlag.

Swinburn, K., Porter, G., Howard, D., Kuvač Kraljević, J., Lice, K., & Matić, A. (2020). *Sveobuhvatni test za procjenu afazije CAT-HR. [Comprehensive Aphasia Test - Croatian]*. Naklada Slap.

The complementiser que in Spanish connected speech: a morphosyntactic signature of dementia

Noor-E-Zannat Nusha | Silvia Martínez-Ferreiro

Introduction and aims

Spanish *que* (that) is a small but unusually busy word: it can link a main clause to a subordinate one (*creo que viene* – I think that s/he is coming), introduce a relative clause (*el libro que leí* – the book that I read), or appear in comparisons (*más alto que yo* – taller than me). The focus of this study is on the complementiser *que* — the version that introduces a finite complement clause after verbs like *creer* (to believe), *saber* (to know), or *decir* (to say). Unlike English *that*, this *que* is not freely droppable, which makes its rate in spontaneous speech a clean window onto whether speakers are still building embedded clauses at all.

Llinàs-Grau and Martínez-Ferreiro (2014) studied the case of *that* in English, comparing 100 adults with aphasia and 100 healthy controls, finding that the aphasia group produced fewer than half as many *that*-introduced subordinate clauses (about 2.2 vs. 5.1 per speaker). The presence-vs-omission pattern itself looked similar; what differed was how often subordination happened at all.

Whether the same thinning happens in dementia, and whether it tracks individual cognition, has not been explored in Spanish. This study asks (1) Do speakers with dementia produce *que* less often than healthy older adults once speech volume is controlled? (2) Is any difference explained by general sentence shortening, or is it specific to *que*? (3) Do MoCA (Nasreddine et al., 2005) scores reflect varying use of complementiser *que* reliably in both people with dementia and controls?

Methods

Participants. Forty Spanish-speaking older adults took part: 20 cognitively healthy controls (age 73.8, SD 7.7; 17F/3M; MoCA 25.2, SD 3.2) and 20 with a clinical diagnosis of dementia (age 83.4, SD 11.2; 17F/3M; MoCA 10.7, SD 4.6). Groups were matched on gender (Fisher $p = 1.00$) but the dementia group was older (Welch $t = -3.17$, $p = .003$); age is therefore included as a covariate. Spontaneous speech data come from the LINDEGALES (ED431F 2024/018) and SECONDSGALES (PID2023-151673OB-I00) projects where participants underwent both cognitive and linguistic evaluation.

Task and transcription. Participants completed a semi-structured connected-speech task. Recordings were transcribed in CHAT format and segmented into utterances following CHILDES/TalkBank conventions (MacWhinney, 2000). Two trained linguists transcribed and reviewed the data in two steps, resolving disagreements by consensus.

Coding. Only complementiser *que* (the finite-clause introducer, e.g., *Creo que viene*) was counted. Relative *que* (*el libro que leí*) and comparative *que* (*más alto que yo*) were excluded. Cases where *que* appeared without a completed embedded clause (false starts, fragments) were excluded from the numerator but kept in the word and utterance denominators. For each speaker, total *que* count, *que* per 100 words, and *que* per analyzable utterance (comp_que_per_utt) were computed.

Analysis. Group differences in the three complementiser *que* measures (raw complementiser que count, complementiser que per 100 words, and complementiser que per analyzable utterance) were tested with Mann–Whitney U tests (rank-biserial effect sizes, 10 000-sample bootstrap CIs, Holm correction across measures). The headline model was a negative-binomial GLM predicting *que* counts from group with log (total words) as an offset; age and MLU-w were added in pre-specified sensitivity models. Within-group associations with MoCA used Spearman ρ with 10 000-permutation *p*-values. Analyses were conducted in Python.

Results

Group difference. Controls produced more *que* than speakers with dementia on every measure (medians: count 12.5 vs. 5.0; per 100 words 1.57 vs. 1.10; per utterance 0.22 vs. 0.13). Holm-corrected Mann–Whitney tests were significant for count and per-utterance (p_{Holm} = .009 and .005) and borderline for per-100-words (p_{Holm} = .054). The negative-binomial GLM gave IRR = 0.55 (95 % CI 0.39–0.77, $p < .001$): once speech volume is controlled, the dementia group uses *que* at about 55 % of the control rate.

Sensitivity analyses. Adding age reduced but did not remove the effect (IRR = 0.61, CI 0.42–0.88, p = .008). Adding MLU-w — which is itself lower in dementia and naturally co-varies with complement-clause use — further attenuated the estimate (IRR = 0.63, p = .031); because MLU and *que* share variance by construction, this is a sensitivity check, not a clean test of "targeted" simplification.

Within-group gradient. The most theoretically interesting result comes from the control group: *que* per 100 words correlated strongly with MoCA (Spearman ρ = .59, CI .22–.82, permutation p = .006) — healthier cognition within the normal range went with more *que*. Within the dementia group, the correlation ran the other way but was not significant (ρ = −.34, p = .14), plausibly a floor effect once finite embedding becomes rare.

Discussion

Complementiser *que* behaves like a graded marker of cognitive status. Between groups, speakers with dementia produce roughly half as many *que* tokens per word as matched controls; within healthy older adults, *que* use scales with MoCA. The same small word both separates the groups and tracks individual cognition in the healthy range; a similar pattern was observed by Llinàs-Grau and Martínez-Ferreiro (2014) for *that* in English. This study demonstrates the same pattern holds in Spanish speakers with dementia.

Caveats: The ∼10-year age gap means the group result must be read with the age-adjusted model; the effect survives but is smaller. The Holm-corrected per-100-words contrast (.054) is borderline and should be read as suggestive. Finally, it is a pilot study (n = 40).

These are initial findings of a broader thesis on morphosyntactic markers of cognitive decline in Spanish connected speech. Further analyses on the same cohort — a subordination index, clause density, function-word proportion, finiteness, and grammaticality of elicited subordinate structures — are currently being coded and are expected by June 2026. Together they will show whether *que* belongs to a coherent syntactic profile of dementia or is an isolated effect.

Acknowledgement

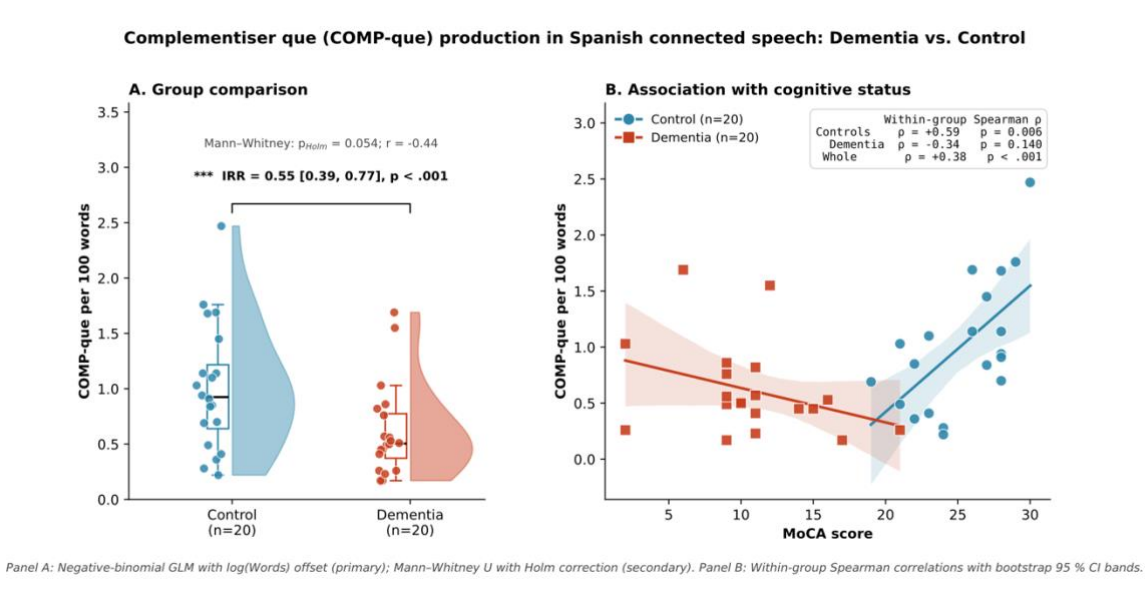
This work is a part of the SECONDSGALES (PID2023-151673OB-I00 funded by MICIU/AEI/10.13039/501100011033 and by "ERDF/EU") and the LINDEGALES (ED431F 2024/018; Xunta de Galicia) research projects led by Silvia Martinez-Ferreiro. SMF also acknowledges support of RYC2020-028927-1, MICIU/AEI/10.13039/501100011033, ESF investing in your future, and the Gerontology and Geriatrics Research Group.

References

Llinàs-Grau, M., & Martínez-Ferreiro, S. (2014). On the presence and absence of *that* in aphasia. *Aphasiology, 28*(1), 62–83. https://doi.org/10.1080/02687038.2013.828345.

MacWhinney, B. (2000). *The CHILDES Project: Tools for Analyzing Talk* (3rd ed.). Lawrence Erlbaum. https://talkbank.org/manuals/CHAT.pdf.

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society, 53*(4), 695–699. https://doi.org/10.1111/j.1532-5415.2005.53221.x



Grammatical Tense Impairment in Aphasia: a usage-based investigation of verb forms in spontaneous language production

Ludovica Onofri | Nathalie Kendrick | Vitor Zimmerer

Introduction and aims

Grammatical tense impairment is commonly observed in people with aphasia (PwA), yet its underlying causes remain debated. Existing accounts have generally attributed these difficulties to syntactic (Friedmann & Grodzinsky, 1997), morphological or semantic factors (Faroqi-Shah & Thompson, 2007). However, relatively little attention has been paid to the role of frequency – particularly beyond single-word level – or to the influence of task type.

The present study replicates previous research (Onofri & Zimmerer, 2024) and adopts a usage-based construction grammar framework (Goldberg, 2003). This framework emphasises the role of usage-frequency effects at the multi-word level while also introducing the notion of lexical bias (i.e., the probability of a word to occur more frequently in a specific syntactic structure). Although these approaches have been adopted in aphasia to investigate passives (Gahl et al., 2003) and plurals (Hatchard & Lieven, 2019), their application to tense remains limited.

In addition, whereas most previous studies have relied on constrained experimental tasks (e.g., sentence completion or sentence-picture matching tasks), the present study examines tense production in samples of semi-spontaneous language, an underexplored context in this domain. We compared the frequency-based, morphological, phonological, and semantic properties of verbs produced by PwA and control participants to assess which factors influence production.

Methods

Participants

A secondary analysis was conducted on the NEURAL database available on AphasiaBank (Stark; MacWhinney et al., 2011). We analysed the responses of 10 PwA (mean age = 62.2, SD = 11.9; mixed aphasia types) and 10 controls (mean age = 56.9, SD = 8.9). Participants were American English native speakers.

Materials and procedure

Data were in the form of transcribed semi-structured interviews of which we selected one question – i.e., whether participants recalled their stroke (PwA) or any illness/injury (controls). All produced verb forms were classified according to their phonological, morphological, and semantic properties using custom R scripts. In addition, four frequency-based variables (i.e., *lexeme frequency, form frequency, tense bias, time bias*) were computed.

Bias was operationalised as the proportion of occurrences in which a given verb appears in a specific tense (tense bias) or generic time reference (time bias). These proportions were automatically calculated from 500 randomly sampled instances per verb drawn from the online *Corpus of Contemporary American English* (COCA; Davies, 2008-), the largest online database of American English encompassing around one billion words collected from 1990 to

2019. For this automated recognition using customised R scripts, we acknowledge an approximate error rate of 15%, which is on a par with other comparable systems.

Lexeme frequency values per verb were also extracted from COCA (Davies, 2008–). Finally, estimated form frequency was derived by tense bias and lexeme frequency. This methodology replicates previous research (Onofri & Zimmerer, 2024).

Results

The analysis found that PwA produced significantly fewer verbs in the past (β = -1.25, p < .001) compared to controls. Additionally, verbs produced by PwA were significantly more aligned to their biased tense (β = 0.10, p < .001) and time reference (β = 0.13, p < .001) compared to controls. Number of phonemes (NoP) was inversely related to all frequency variables (lexeme frequency: β = -0.59, p < .001; form frequency: β = -0.97, p < .001; tense bias: β = -0.04, p < .001; time bias: β = -0.01, p < .001) and there was a significant interaction between Group and NoP for all frequency variables (lexeme frequency: β = 0.12, p = .002; form frequency: β = 0.13, p < .001; tense bias: β = -0.01, p = .037; time bias: β = -0.01, p = .025).

No significant group differences were found for quantity or lexical diversity of verbs, phonological and morphological complexity, semantic weight, or regularity.

Discussion

This study generally replicates previous findings (Onofri & Zimmerer, 2024), confirming that lexical bias and usage-frequency can shape verb production profiles, particularly with respect to tense use, and help differentiate between PwA and control participants. At the same time, some inconsistencies with prior work emerged. In particular, the present data suggest a past tense impairment in this sample of PwA, supporting accounts that propose greater difficulty with past compared to non-past reference.

These results must be interpreted while keeping in mind the limitations of the current study, such as the reduced sample size and the high variability within PwA given by the heterogenous group composition.

Nevertheless, the growing evidence base for lexical bias has clinical implications, indicating that usage-frequency and bias should be considered in aphasia assessments and interventions. Additionally, results underscore the importance of accounting for phonological complexity when researching frequency effects in aphasia.

References

Davies, Mark. (2008-) The Corpus of Contemporary American English (COCA). Available online at https://www.english-corpora.org/coca/.
Faroqi-Shah, Y., & Thompson, C. K. (2007). Verb inflections in agrammatic aphasia: Encoding of tense features. *Journal of Memory and Language*, *56*(1), 129–151.
Friedmann, N., & Grodzinsky, Y. (1997). Tense and Agreement in Agrammatic Production: Pruning the Syntactic Tree. *Brain and Language*, *56*(3), 397–425.

Gahl, S., Menn, L., Ramsberger, G., Jurafsky, D. S., Elder, E., Rewega, M., & Audrey, L. H. (2003). Syntactic frame and verb bias in aphasia: Plausibility judgments of undergoer-subject sentences. *Brain and Cognition*, *53*(2), 223–228.
Goldberg, A. E. (2003). Constructions: a new theoretical approach to language. *Trends in Cognitive Sciences*, *7*(5), 219–224.
Hatchard, R., & Lieven, E. (2019). Inflection of nouns for grammatical number in spoken narratives by people with aphasia: how glass slippers challenge the rule-based approach. *Language and Cognition*, *11*(3), 341–372.
MacWhinney, B., Fromm, D., Forbes, M. & Holland, A. (2011). AphasiaBank: Methods for studying discourse. Aphasiology, 25,1286-1307. Grant support NIH-NIDCD R01-DC008524 (2022-2027).
Onofri, L., Zimmerer, V., (2024, September 9-12). *Grammatical Tense Impairment in Aphasia: A usage-based analysis* [Poster session]. Science of Aphasia Conference, Geneva, Switzerland.

Acute white matter integrity differentially predicts early connected speech recovery in post-stroke aphasia

Alberto Osa García, Simona Maria Brambati, Maélys Pajon Roux, Amélie Brisebois, Bérengère Houzé, Christophe Bedetti, Alex Desautels and Karine Marcotte

Introduction and Aims

Connected speech impairments are highly prevalent in post-stroke aphasia (PSA) and reflect disruptions across multiple linguistic domains. Although connected speech is increasingly used to characterize aphasia recovery, the neural correlates supporting early recovery remain poorly understood. The present study aims to investigate whether acute white matter integrity predicts subacute connected speech performance in PSA.

Methods

Twenty-five individuals with PSA (mean age = 66 years, SD = 14; 10 women) were recruited from an acute care stroke unit. Participants underwent diffusion MRI and language assessment within the first 3 days post-stroke (T1; mean : 2 ± 1 day), followed by a language reassessment during the subacute stage (T2; mean : 10 ± 3 days).

Connected speech samples were elicited using the picture description task of the Western Aphasia Battery – Revised (Kertesz, 2006). Speech samples were transcribed and analyzed to extract measures of fluency (words per minute [WPM], mean length of utterance [MLU]) and informativeness (Thematic Units [TU], Information Content Units [ICU]). TU and ICU measures were additionally normalized for speaking time (TU/min; ICU/min). Microstructural measures were derived using the EVAL program from the CLAN software suite (MacWhinney, 2000). TU and ICU scoring procedures were based on previously validated Laurentian French protocols (Brisebois et al., 2020; Marcotte et al., 2024). Discourse measures were calculated at both timepoints, and change scores (T2–T1) were computed for all variables. Previously published analyses from this dataset demonstrated high inter-rater reliability for discourse measures (Brisebois et al., 2021).

Diffusion MRI preprocessing and tractography were performed using the TractoFlow pipeline (Theaud et al., 2020). Diffusion tensor imaging metrics, including fractional anisotropy (FA), mean diffusivity (MD), and axial diffusivity (AD), were extracted bilaterally from the arcuate fasciculus (AF), inferior fronto-occipital fasciculus (IFOF), and frontal corpus callosum (Fr-CC). These tracts were selected based on successful reconstruction in all participants. Acute lesion segmentation was performed semi-automatically on MD maps using Clusterize (De Haan et al., 2015), and lesion volumes were normalized to intracranial volume.

Multiple regression analyses examined whether diffusion metrics predicted subacute connected speech measures while controlling for age, education, and lesion volume. Initial aphasia severity was also included to compare the predictive contribution of behavioral and neuroimaging variables.

Results

Initial aphasia severity significantly predicted subacute measures related to productivity, fluency, and informativeness, but not efficient informativeness measures (TU/min, ICU/min). In contrast, diffusion metrics from the frontal corpus callosum (Fr-CC) significantly predicted all informativeness measures and were the strongest predictors of efficient informativeness. For example, FA from the Fr-CC significantly improved prediction of ICU/min beyond demographic and lesion covariates (adj. R^2 = .50 vs .29 in the covariate-only model), whereas initial aphasia severity did not significantly improve the model further.

Speech fluency (WPM2) was associated with diffusion metrics from the left arcuate fasciculus (AF) and left inferior fronto-occipital fasciculus (IFOF). FA from the left IFOF increased explained variance from Adj. R^2 = .14 in the covariate-only model to adj. R^2 = .53, and to adj. R^2 = .67 after inclusion of initial aphasia severity. Similar but weaker effects were observed for homologous right-hemisphere tracts.

Finally, left-hemisphere diffusion metrics (particularly MD and AD from the left IFOF) were the only significant predictors of change measures (ΔWPM), suggesting a dissociation between predictors of achieved performance and predictors of recovery-related improvement.

Discussion

Together, these findings suggest that diffusion-derived biomarkers capture aspects of early recovery not reflected by initial aphasia severity alone. Interhemispheric frontal connectivity may contribute specifically to efficient informativeness, whereas dorsal and ventral language pathways appear more strongly associated with speech fluency and recovery-related gains. The results further suggest that diffusion-derived biomarkers may capture aspects of recovery not reflected by initial aphasia severity alone.

References

Brisebois, A., Brambati, S. M., Désilets-Barnabé, M., Boucher, J., Osa García, A., Rochon, E., Leonard, C., Desautels, A., & Marcotte, K. (2020). The importance of thematic informativeness in narrative discourse recovery in acute post-stroke aphasia. *Aphasiology*, *34*(4), 472-491.

Brisebois, A., Brambati, S. M., Boucher, J., Rochon, E., Leonard, C., Désilets-Barnabé, M., Desautels, A., & Marcotte, K. (2021). A longitudinal study of narrative discourse in post-stroke aphasia. *Aphasiology 36(7), 805-830*

De Haan, B., Clas, P., Juenger, H., Wilke, M., & Karnath, H. O. (2015). Fast semi-automated lesion demarcation in stroke. *NeuroImage: Clinical*, *9*, 69–74.

Kertesz, A. (2006). *Western Aphasia Battery- Revised*. Pearson.

Marcotte, K., Roy, A., Brisebois, A., Jutras, C., Leonard, C., Rochon, E., & Brambati, S. M. (2024). Reliability of the picture description task of the Western Aphasia Battery – revised in Laurentian French persons without brain injury. *The Clinical Neuropsychologist*, *38*(8), 1980–2008.

Theaud, G., Houde, J.-C., Boré, A., Rheault, F., Morency, F., & Descoteaux, M. (2020). TractoFlow: A robust, efficient and reproducible diffusion MRI pipeline leveraging Nextflow & Singularity. *NeuroImage*, *218*, 116889.

The effect of lingustic complexity on temporal aspects of speech production in normal Aging: Data from Greek

Ioannis Papakyritsis | Michaela Nerantzini | Emilia Michou

Introduction and aims

Healthy aging is associated with functional/ anatomical changes, i.e. reduction of strength, speed, endurance, flexibility and accuracy of the speech production system (Xue & Hao, 2003), and cognitive/ linguistic changes, including reported declines in processing speed, attention, memory, language processing and word retrieval (see e.g., Harada et al., 2013). These changes have a well-documented effect on temporal aspects of speech production, namely speaking rate, pause and disfluency patterns (Tucker, Ford & Hedges, 2021), although variability of findings across studies has been reported. The degree of age-related speaking rate reduction has been found to depend on experience and sociolinguistic variables (Quené, 2013). Although several studies have reported shorter phrases (Jacewicz et al., 2010), longer pauses (Bóna, 2014) and more frequent pauses and increased disfluency (Martins and Andrade, 2011) in older adults, Bóna (2011) found a comparable rate of disfluent speech in younger and older adults that tended to increase with more demanding speakings tasks, while Taschenberger et al. (2019) report a declination in filled pause frequency with older age, specifically in adverse speaking conditions, including the presence of background noise. Taken together, the available findings suggest a more complex relationship between healthy aging changes and temporal speech organization. It is plausible that experience and cognitive decline collectively affect speech performance in older speakers, especially during more difficult speaking tasks.

The aim of this study is to investigate the effect of linguistic complexity on temporal aspects of speech production in the Greek older population. We seek to find out how rate of speech, pause patterns and speech tempo flexibility, are conditioned by age and phonological/ syntactic complexity. This question carries clinical significance given that speaking rate is a crucial, modifiable variable that is often the focus of attention in intervention protocols for patients with neurogenic communication disorders, including aphasia and apraxia of speech, that are usually older adults. Such protocols are typically hierarchically organized, taking linguistic complexity variables into account.

Methods

Two groups of participants, 60 young adults (20-25 years old) and 60 older adults 65-90 years old) carried out a reading task consisting of two set of sentences of different complexity. 10 short simple sentences, 10 syllables in length with an SVO syntactic structure, and 10 long, complex sentences, 20-25 syllables in length, adapted from Arvaniti (2012) initially devised to measure speech rhythm. Speaking rate was modified on request; three different rate conditions were elicited, “habitual”, “slow” and “fast” rate of speech. A set of acoustic and other temporal measures, including speaking rate, articulation rate, relative speaking rate increase and decrease, breath group size and duration, pause duration and position, were used to compare performance across the two sentence sets for both groups.

Results

The older group spoke at a slower tempo and used a higher number of pauses than young adults across all rate conditions. In the increased complexity condition, the difference in performance between the two groups tended to be more marked, especially in terms of pause frequency patterns. Interestingly, both groups tended to articulate faster in the complex sentence set. Regarding speaking rate flexibility, several older speakers experienced considerable difficulties modifying, especially increasing, their rate on request. Linguistic complexity did not appear to have a significant impact on speech tempo modification performance. As a matter of fact, preliminary analysis indicated that older speakers were more likely to successfully increase their speaking rate when producing the longer, more complex set of sentences.

Discussion

Overall, findings show that increased sentence complexity in the context of a reading task, can have an effect on the temporal organization of speech output, despite significant individual variation. However, age-related decline in speech performance was not clearly conditioned by linguistic complexity; when producing the complex sentence set older speakers tended to use more pauses but, at the same time, they did not slow their overall speaking rate compared to the reduced complexity condition, similar to that of the younger speakers. These results can have possible implications for the design of speaking rate modification protocols for older patients with neurogenic communication disorders.

References

Arvaniti, A. (2012). *Journal of Phonetics*, 40(3) 351-373.
Bóna, J. (2011). Govor, 28, 95–115.
Bóna, J. (2014). The Journal of the Acoustical Society of America, 136, EL116–EL121.
Harada, C. et al. (2013). Clinics in Geriatric Medicine, 29, 737–752.
Jacewicz, E. et al. (2010). The Journal of the Acoustical Society of America, 128, 839–850.
Quené, H. (2013). The Journal of the Acoustical Society of America, 133, EL452–EL457.
Martins, V. D. O. & Andrade, C. R. F. D. (2011). Rev. Soc. Bras. Fonoaud. 16(3), 344–349.
Taschenberger, L. et al. (2019). Proceedings of DiSS 2019 (pp. 55–58). Budapest, Hungary: ELTE Faculty of Humanities.
Tucker BV et al. (2021). WIREs Cogn Sci. 2021;12:e1557.
Xue, S. A., & Hao, G. J. (2003). Journal of Speech, Language, and Hearing Research, 46, 689–701.

Learning and maintenance of L3 novel words in a multilingual person with aphasia: Is there a cognate facilitation effect?

Claudia Peñaloza | Brigitte Ribeiro | Andrea Escoté | Alejandro Villanueva

Introduction and aims

Past research has demonstrated that cognates (translation pairs sharing meaning and word form across languages) show a facilitation effect relative to non-cognates (translations pairs only sharing meaning but not word form) in neurotypical bilinguals. Cognates elicit more accurate responses and faster reaction times than non-cognates in single word processing (Costa et al., 2000) and novel word learning (Marecka et al., 2021). The effect has been observed across varying language combinations and in different directions (in the native L1 and the second-acquired L2) although it may vary depending on task demands.

The cognate facilitation effect has been demonstrated in multilingual persons with aphasia (MPWA) in picture naming (Marte et al., 2023), although both facilitation and interference have been reported in anomia therapy (Lerman & Malcolm, 2025). Notably, this effect has not been previously examined in verbal learning. Assessing the cognate effect in novel word learning in MPWA is highly relevant as novel word learning may inform anomia therapy (Peñaloza et al., 2022). Further, the ability to learn cognates versus non-cognates can provide insights into how languages influence one another which can impact anomia therapy outcomes in MPWA (Peñaloza et al., 2025). This study aimed to assess the cognate effect in novel word learning and long-term maintenance in a MPWA across expressive recall, recognition and translation.

Methods

Participant. AB (43 years old) was a multilingual woman (L1 Catalan: native proficiency; L2 Spanish: native proficiency; L3 English: intermediate proficiency; L4 German = intermediate proficiency) with mild fluent chronic aphasia following a left fronto-parietal acquired brain injury (ABI). English was her weakest language pre-ABI.

Stimuli. AB completed an extensive picture naming screener in all her languages (except for German). Twelve L3 English words self-rated as unknown pre-ABI (naming accuracy = 0/3 baselines) of different semantic categories (6 cognates and 6 noncognates matched on lexical frequency, length, L1 Catalan/L2 Spanish naming accuracy ≥ 1/3 baselines) were selected for training.

Associative word learning task. The experiment was completed in two sessions, a first session evaluating learning ability and a second session assessing maintenance of learned words completed 3 days apart. The *learning phase* included 5 training cycles (2 training blocks per cycle, 12 trials per training block). Each trial presented a L3 word-picture pairing and required AB to listen and repeat the trained word. Each training cycle continued with an *expressive recall test* evaluating lexical access for the newly trained word forms (12 trials) followed by a *recognition test* assessing word-meaning recognition (24 trials: 12 correct/12 incorrect word-picture pairs). At the end of training, two *receptive translation tests* assessed AB's knowledge of L3 English- L2 Spanish and L3 English- L1 Catalan translation pairs (24 trials: 12 correct/12 incorrect pairings per translation test). AB's long-term maintenance of newly trained English words was assessed with tests of expressive recall, recognition and

translation. Additionally, AB completed a word learning questionnaire asking about reliance on other languages during the learning task (session 1).

Results

Expressive recall. AB's accuracy increased substantially across items from baseline to test 5 (83.3%; binomial test p= 0.0193). Significant learning was found for cognates at test 5 (100%; binomial test p= 0.0156) with improved accuracy from baseline, McNemar test: χ^2 (1, N= 6)= 4.167, p= 0.031. Non-cognates also improved (83.3%; binomial test p= 0.0193) but this set did not show significant change over time, χ^2 (1, N= 6)= 2.250, p= 0.125.

Recognition. AB's accuracy was at ceiling nearly across all tests. Test 5 achieved maximum accuracy overall (100%) with each item set showing significant above-chance performance (100%; binomial test p= 0.0156).

Translation. AB's accuracy in the L3 English- L2 Spanish translation test was high across items (83.3%; binomial test p= 0.0193), being significantly above chance for cognates (100%; binomial test p = 0.0156) but not for non-cognates (66.7%; p= 0.3438). AB's accuracy in the L3 English- L1 Catalan translation test achieved higher accuracy (91.7%; binomial test p= 0.0032) being significantly above chance for cognates (100%; binomial test p = 0.0156) but not for non-cognates (83.3%; binomial test p= 0.1094).

Maintenance. AB's expressive recall accuracy decayed to 66.7% (binomial test p = 0.1938). Her recognition accuracy was maintained at ceiling (100%), with each item set being significantly above chance (100% binomial test p = 0.0156). Her accuracy in both translation tests improved overall (100%), with each item set showing significant above-chance performance (100% binomial test p = 0.0156).

Word learning questionnaire. AB reported that trained words came to mind more often in her L1 Catalan and she relied on translation during the initial part of the training procedure. Gradually, she thought of the words more often in her L3 English but needed to mentally supress the Catalan translations. She felt that Catalan was helpful rather than interfering during learning.

Discussion

Our preliminary results show successful L3 word learning for AB across expressive recall, recognition and translation tests and significant maintenance at follow-up albeit only in recognition and translation tests. These findings support evidence for spared verbal learning ability in some people with aphasia (Peñaloza et al., 2022). In contrast to non-cognates, cognates showed significant improvement, both relative to baseline and chance-level performance in expressive recall and translation tests at the end of training. These findings support a cognate facilitation effect in L3 word learning in a MPWA and suggest learning words in a weaker language can rely on existing lexical knowledge in a native language.

References

Costa, A., Caramazza, A., & Sebastian-Galles, N. (2000). The cognate facilitation effect: implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(5), 1283.

Marecka, M., Szewczyk, J., Otwinowska, A., Durlik, J., Foryś-Nogala, M., Kutyłowska, K., & Wodniecka, Z. (2021). False friends or real friends? False cognates show advantage in word form learning. *Cognition*, *206*, 104477.
Marte, M. J., Peñaloza, C. & Kiran, S. (2023). The cognate facilitation effect on lexical access in bilingual aphasia: Evidence from the Boston Naming Test. *Bilingualism: Language and Cognition, 26*(5), 1009–1025.
Lerman, A., & Malcolm, T. R. (2025). Facilitation and Interference Effects During Cognate Retrieval in a Multilingual Person With Aphasia. *American Journal of Speech-Language Pathology*, *34*(3), 1005-1022.
Peñaloza, C., Martin, N., Laine, M., & Rodríguez-Fornells, A. (2022). Language learning in aphasia: A narrative review and critical analysis of the literature with implications for language therapy. *Neuroscience & Biobehavioral Reviews*, *141*, 104825.
Peñaloza, C., Marte, M. J., Billot, A., & Kiran, S. (2025). Cross-language interaction during sequential anomia treatment in three languages: Evidence from a trilingual person with aphasia. *Cortex*, *189*, 107-130.

Silent and Filled Pauses as Markers of Longitudinal Language Change in Mild Primary Progressive Aphasia: A Case Study

Cassiady Perard | Miral Goral

Introduction and aims

Primary Progressive Aphasia (PPA) is a neurodegenerative disorder characterized by gradually worsening anomia and progressive disruption of language production and comprehension (Gorno-Tempini et al., 2011). Symptoms are subtle initially, but as PPA progresses, the deficits increasingly impact functional communication. Identifying early sensitive markers of language change in mild PPA remains challenging, particularly in those who are users of more than one language. Disfluencies, such as filled and silent pauses, are commonly observed across PPA variants and may reflect underlying disruptions in lexical retrieval and speech planning, with prior studies showing variant-specific differences in pause frequency and duration (Potagas et al., 2022; Gorno-Tempini et al., 2008).

Despite this, pauses have been underexplored as longitudinal markers of early change. Examining the frequency and duration of silent and filled pauses during connected language production may offer a sensitive indicator of subtle changes in the progression of PPA.

This study examines longitudinal changes over time in connected language production in a bilingual person with mild PPA, with a focus on disfluencies as potential early markers of change.

Methods

This longitudinal case study followed a 62-year-old Spanish-English bilingual man (SE13P) with mild PPA across three time points over an eight-month period. All testing was conducted in English. Language samples were elicited using three connected language tasks: picture sequence descriptions, narrative storytelling, and responses to WH-questions. Language measures included productivity, grammaticality, sentence complexity, coherence, repetitions, and disfluencies, with a particular focus on the frequency and duration of silent and filled pauses. Frequency was the number of times a pause was counted within a task. Duration of each pause was measured in seconds and then averaged across each task. Silent pauses were defined as intervals greater than 0.25 seconds; shorter pauses were excluded from analysis. Filled pauses were defined as vocalized hesitation sounds (e.g., “um”, “uh”). Pauses occurring at task boundaries (e.g., prior to initiating a response) were excluded to focus on within-utterance speech production. In addition, comparable data from a matched neurologically healthy control individual were analyzed.

Pauses were measured semi-manually using PRAAT (Boersma, & Weenink, 2026). To ensure reliability of pause identification, a third of the data was independently coded by two raters. For each task and time point, mean pause duration and pause frequency was calculated. Preliminary analyses examined within-task changes over time and compared SE13's performance to the neurotypical control data using nonparametric statistical tests (Mann-Whitney U).

Results

Mean silent pause duration increased over time in WH-question responses and picture description tasks, with more modest increases observed in narrative production. From initial to final testing, mean pause duration increased by approximately 60% in WH-Questions, 54% in picture description tasks, and 22% in narratives. The longest pause durations were observed at the final testing time, and were overall longer than the initial time and significantly longer than those of the matched control participant. Pause frequency was variable and did not demonstrate a consistent longitudinal pattern across tasks and sessions. High interrater reliability was observed, with a range of accuracy from 90-98% for silent pauses and 82%-92% for filled pauses.

In contrast to the pause data, most traditional language measures showed relative stability over time. Although overall productivity, as measured by total words and CIUs (Nicholas & Brookshire, 1993), showed a slight downward trend, measures of linguistic efficiency (CIUs/Total words), sentence complexity, coherence, and repetition did not show clear change over time. Both the participant with PPA and control produced coherent discourse with high levels of topic maintenance and structurally complete sentences.

Discussion

Preliminary results suggest that pause duration can be a sensitive measure of subtle language change over time. The findings are consistent with previous literature, suggesting that participants with nonfluent variant PPA exhibited longer silent pause duration compared to healthy neurotypical controls (Mack et al., 2015; Potagas et al., 2022). The observed duration increases over time in both silent and filled pauses in this case study suggest that disfluency-related measures may be sensitive to subtle longitudinal changes in connected speech, even in mild stages of PPA.

In contrast, most other language measures remained relatively stable over time. Although overall productivity, as measured by total words and CIUs, showed a slight negative trend, measures of efficiency (proportions of CIUs out of total words), sentence complexity, and coherence did not change and were comparable to the matched control participant. The only consistent difference observed in language measures was lower grammaticality in the participant with PPA compared to the control, without observable change over time; however, this may be influenced by bilingual language experience rather than impairment alone.

Taken together, these findings suggest that traditional language measures may not be sufficiently sensitive to detect subtle early change in mild PPA. Instead, disfluency measures - particularly silent and filled pause duration and frequency - may provide a more sensitive marker of early progression. These findings are particularly relevant for bilingual speakers, where traditional measures may be less reliable due to variability in language use, proficiency and exposure.

Future research could examine these effects in larger cohorts and across bilingual populations, in both languages, to determine the generalizability of pause-based markers for clinical monitoring and early detection.

References

Boersma, P., & Weenink, D. (2026). *Praat: doing phonetics by computer* (Version 6.4.65) [Computer software]. https://praat.org
Gorno-Tempini, M. L., et al. (2008). The logopenic/phonological variant of primary progressive aphasia. *Neurology, 71*(16), 1227–1234. https://doi.org/10.1212/01.wnl.0000320506.79811.da
Gorno-Tempini, M. L., et al. (2011). Classification of primary progressive aphasia and its variants. *Neurology, 76*(11), 1006–1014. https://doi.org/10.1212/WNL.0b013e31821103e6
Mack, J. E., et al. (2015). What do pauses in narrative production reveal about word retrieval deficits in PPA? *Neuropsychologia, 77*, 211–222. https://doi.org/10.1016/j.neuropsychologia.2015.08.019
Nicholas, L. E., & Brookshire, R. H. (1993). A system for quantifying the informativeness and efficiency of connected speech in adults with aphasia. *Journal of Speech, Language, and Hearing Research, 36*(2), 338–350. https://doi.org/10.1044/jshr.3602.338
Potagas, C., et al. (2022). Silent pauses and speech indices as biomarkers for primary progressive aphasia. *Medicina, 58*(10), 1352. https://doi.org/10.3390/medicina58101352

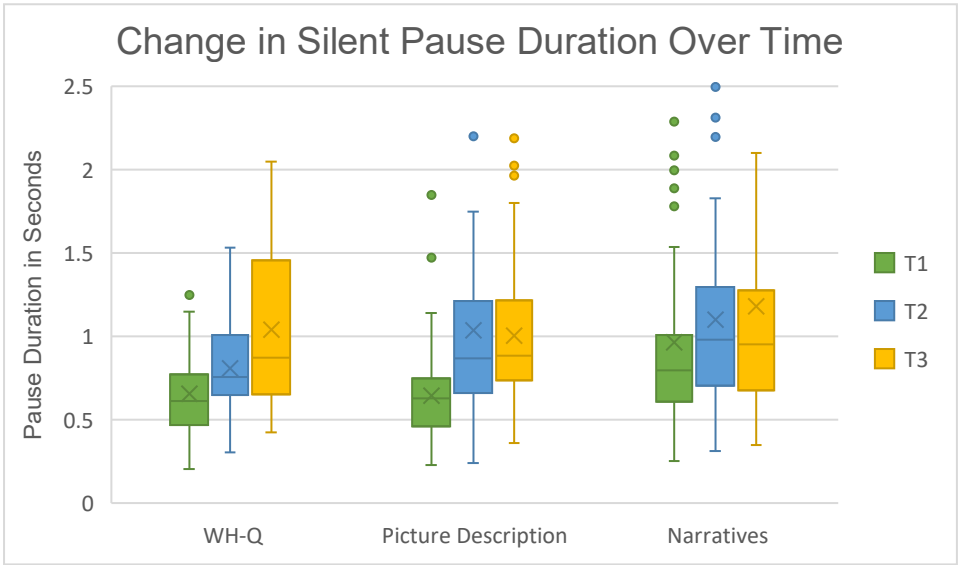


Figure 1

Noun and Verb Association in German: Test Standardisation and Validation of the GB-WoBA

Madeline Philipsen | Manon Winkler | Roel Jonkers | Thomas Picht | Adrià Rofes

Introduction and aims

The comparison between nouns and verbs has been the focus of extensive psycho- and neurolinguistic research[1]. Behavioural accounts suggest a dissociation between the two word categories, while the distinction is less clear-cut at the neural level[1]. These findings have practical implications for awake brain surgery, because thorough mapping of language functions in the brain can lead to a reduction of postoperative deficits and an increase in resected volume of tumour tissue[2]. Commonly used tasks to achieve these purposes are object and action naming[3]. However, these production tasks are not sufficient to investigate language comprehension abilities[4] and may not always be suited to assess mild language impairments, which otherwise prevail and are affected in people with a brain tumour[5].

The aim of this study is therefore to design, pilot, standardise, and validate a new German test to assess people with a brain tumour before, during, and after surgery, called Groningen-Berlin Word-Picture Association Test (German: Groningen-Berlin Wort-Bild-Assoziationstest, GB-WoBA). The test will target mild deficits in noun (e.g., dog-cat, dog-leash) and verb association (e.g., kick-punch, kick-write), and differentiate between thematic and taxonomic relations.

Methods

Each item of the GB-WoBA consists of a written prime in the middle of a computer screen, which is to be compared to the four pictures in the corners of the screen (a target, a semantic, a phonological, and an unrelated distractor). The test was developed in several stages: (1) To assess the unambiguity of the image stimuli, a name agreement study with ten healthy participants was conducted. (2) During the subsequent pilot study, ten other healthy participants completed the preliminary set of object and action association items. Items that did not reach an accuracy of at least 80% or did not meet the criteria regarding semantic and phonological similarity were excluded. (3) Afterwards, the standardisation study was conducted in 51 healthy participants divided into five age groups. Items that did not have an accuracy of at least 80% were excluded. Due to matching between taxonomic and thematic relationships on several psycholinguistic variables (frequency, imageability, age of acquisition, number of graphemes and syllables, semantic similarity, and Levenshtein distance), the number of final items was further decreased. Cut-off scores for each age group were also calculated. Effects of age, education, and the location of the target stimulus on accuracy and reaction times were assessed by fitting (generalised) linear mixed models. (4) The validation study will be conducted in 20 people with post-stroke aphasia to evaluate the suitability of the GB-WoBA to assess language comprehension and associative abilities.

Results

The final version of the test after standardisation includes 92 association items (46 noun associations, 46 verb associations). With an average score of 44.1 out of 46 on the final version of the noun subtest and 43.5 out of 46 on the final version of the verb subtest, there was no significant difference in the accuracy of healthy participants. Different factors

influenced the test performance: there was a significant effect of age on accuracy ($p < .001$), while education and the location of the target stimulus did not significantly affect the test results (education: $p = .714$, target location: $p = .070$). None of the three variables had a significant influence on reaction times (age: $p = .063$, education: $p = .484$, target location: $p = .241$). Concerning the comparison between nouns and verbs, we did not find a significant effect of word category on accuracy ($p = .110$). However, word category significantly predicted reaction times ($p < .001$), as reaction times were slower on verbs compared to nouns.

At the time point of writing this abstract, results for the validation study are not yet available.

Discussion

The new test includes images that are unambiguous, and the relations between stimuli concerning semantic similarity and Levenshtein distance are appropriate. Moreover, the test was feasible in healthy participants of different age groups. The results align with the behavioural literature[1]: performance on nouns was better and response latencies were faster, compared to verbs, even though only the latter showed a significant difference. Perhaps because of the generally high accuracy on both noun and verb association, differences in accuracy in our dataset were not predicted by word category. For reaction times, differences between noun and verb association can be traced to the fact that the word categories in these tests were different. However, it is possible that the results reflect the fact that the two groups of noun and verb stimuli were not matched for psycholinguistic variables such as frequency or imageability. Concerning the effects of age on test performance, it can be said that our results are in line with what has previously been found: participants had lower test scores if they were older.[3] As for the performance of patients with aphasia, we can predict that nouns will have higher accuracy and lower reaction times compared to verbs.[1]

References

[1]de Aguiar, V., & Rofes, A. (2022). Chapter 14—The noun-verb distinction. In G. Miceli, P. Bartolomeo, & V. Navarro (Eds.), *Handbook of Clinical Neurology* (Vol. 187, pp. 245–262). Elsevier. https://doi.org/10.1016/B978-0-12-823493-8.00006-7
[2]De Witt Hamer, P. C., Robles, S. G., Zwinderman, A. H., Duffau, H., & Berger, M. S. (2012). Impact of Intraoperative Stimulation Brain Mapping on Glioma Surgery Outcome: A Meta-Analysis. *Journal of Clinical Oncology, 30*(20), 2559-2565. https://doi.org/10.1200/JCO.2011.38.4818

[3]Rofes, A., de Aguiar, V., & Miceli, G. (2015). A minimal standardization setting for language mapping tests: An Italian example. *Neurological Sciences*, *36*(7), 1113–1119. https://doi.org/10.1007/s10072-015-2192-3
[4]Whitworth, A., Webster, J., & Howard, D. (2014). *A Cognitive Neuropsychological Approach to Assessment and Intervention in Aphasia: A clinician's guide* (2nd ed.). Psychology Press. https://doi.org/10.4324/9781315852447
[5]Rybka, L., Jonkers, R., Burzlaff, M., Rosenstock, T., Vajkoczy, P., Picht, T., Faust, K., & Rofes, A. (2024). Preoperative subjective impairments in language and memory in brain tumor patients. *Frontiers in Oncology*, *14*, 1475860. https://doi.org/10.3389/fonc.2024.1475860

A group therapy approach using generative artificial intelligence and accessibility features of smart devices to support written communication and social participation in aphasia

PYTHON Gregoire | DEMAEGDT Astrid

Introduction and aims

In today's hyperconnected world, interactions through smartphones (e.g., WhatsApp groups, social media) play a crucial role in social participation and might even act as a lifeline for persons with aphasia (PWA) (McMahon et al., 2025). However, written language impairments following brain damage can significantly limit access to digital interactions (Marshall et al., 2018). Built-in accessibility features such as predictive text and speech-to-text offer valuable support. Recent advances in generative artificial intelligence (AI), such as chatGPT, demonstrate additional potential for PWA to create or revise written content (Allen et al., 2025).

This study aimed to explore the feasibility and potential benefits of integrating AI-assisted writing tools into group speech and language therapy (SLT) for PWA. A group format was selected due to evidence suggesting that outpatient group interventions could improve linguistic processes, social networks and community access (Lanyon et al., 2013). Specifically, we investigated the impact of group SLT on participants' written communication speed and strategies, confidence, and engagement in digital interactions.

Methods

Three PWA with post-stroke chronic aphasia (P1 & P2 – moderate Wernicke's aphasia, P3 – moderate Broca's aphasia) participated in a 13-session weekly group therapy program conducted between October 2025 and January 2026. They had prior individual SLT experience for several years. During the sessions, led by two SLTs, PWA used their own smartphones and tablets. A baseline assessment (two pre-tests separated by at least 3 weeks) was conducted followed by a post-therapy evaluation. The first 10 sessions focused on functional written communication in digital contexts, incorporating built-in accessibility features and the free app version of ChatGPT to revise and create written content. PWA practiced ecological situations using several apps (Whatsapp, Mail, Notes, Calendar, Web browser, Facebook, Instagram) and received psychoeducation about ChatGPT/AI. Following a 3-week intervention break, during which the PWA were encouraged to use the trained strategies in their daily lives, the final 3 sessions reviewed difficulties and encouraged strategy use.

Results

Results indicated high levels of participant satisfaction and strong engagement with the group format. Writing speed increased significantly in P3 (from 2.3 words per minute (wpm) pre-therapy to 8.2 wpm post-therapy) and in P2 (5.8 to 9.5 wpm), while remaining stable in P1 (10.6 to 11 wpm). All three participants reported increased confidence in digital written communication (adapted version of the Communication Confidence Rating Scale for Aphasia – mean increases of +5.5%, +6.5% and +27.5% respectively) and reduced subjective complaints in written language (Echelle de Satisfaction de la PArticipation COmmunicationnelle pour les personnes ayant un trouble acquis: -2, -1 and -1 respectively).

PWA developed flexible strategies, using predictive text and speech-to-text alternately. Additionally, all participants learned to use ChatGPT and were able to instruct the AI-agent to draft an informative email in a few minutes (205'', 180'' and 252'' respectively).

Discussion

This preliminary study demonstrated the feasibility of using a group therapy approach, incorporating accessibility features and generative AI, to support written communication and social participation in aphasia. Despite the overall positive outcomes, several challenges emerged, including the need for tailored support due to participant heterogeneity in language and digital skills. For instance, P1 and P2 with Wernicke's aphasia preferred ChatGPT's written output, whereas P3 with Broca's aphasia preferred ChatGPT's oral output. Furthermore, participants required support with fundamental digital competencies, such as connecting to Wi-Fi networks, installing new apps, using copy-paste functions, taking screenshots, *etc.* PWA expressed interest in using the Health app and online games with family, potentially enhancing social participation. These findings highlight the potential and practical considerations for integrating AI tools into aphasia rehabilitation, emphasizing the importance to consider basic digital literacy as an integral component of intervention. Further research with larger groups and control conditions is warranted.

References

Allen, A. K., Brennan, C., Riseman, C., Kleiber, H., & Hilger, A. I. (2025). Harnessing AI for aphasia : A case report on ChatGPT's role in supporting written expression. Frontiers in Rehabilitation Sciences, 6, 1600145. https://doi.org/10.3389/fresc.2025.1600145

Lanyon, L. E., Rose, M. L., & Worrall, L. (2013). The efficacy of outpatient and community-based aphasia group interventions : A systematic review. International Journal of Speech-Language Pathology, 15(4), 359-374. https://doi.org/10.3109/17549507.2012.752865

Marshall, J., Caute, A., Chadd, K., Cruice, M., Monnelly, K., Wilson, S., & Woolf, C. (2018). Technology-enhanced writing therapy for people with aphasia : Results of a quasi-randomized waitlist controlled study: Technology-enhanced writing therapy for people with aphasia. International Journal of Language & Communication Disorders. https://doi.org/10.1111/1460-6984.12391

McMahon, M., Azios, J. H., Archer, B., Douglas, N. F., Strong, K., Simmons-Mackie, N., & Worrall, L. (2025). Using Digital Technology to Stay Connected With Friends After Aphasia. American Journal of Speech-Language Pathology, 34(6S), 3593-3610. https://doi.org/10.1044/2025_AJSLP-24-00395

Aphasia and Neurodegenerative Language Disorders in Creole Speakers: A Scoping Review

Jaya Ratane | Mauro Viganò | Seçkin Arslan

Introduction and aims

Post-stroke aphasia and neurodegenerative language disorders may have differential characteristics depending on the language (Menn & Obler, 1990; Martínez-Ferreiro et al., 2026). However, these variations remain largely unexplored in bilingual Creole-speakers. Creole languages emerged in colonial settings through language contact and developed into fully-fledged languages with their own grammatical systems, often embedded in multilingual environments (Mufwene, 2001). They display specific linguistic features, including reduced morphology and distinct syntactic patterns (preverbal tense-aspect-mood particles, serial verb constructions, and zero copula) differentiating them from coexisting languages (Holm, 2000). Despite their widespread use, Creole languages remain underrepresented in research on aphasia (Matić et al., 2026). Managing aphasia in multilingual individuals raises complex issues, including the selection of the language of intervention and the identification of vulnerable linguistic components (Lorenzen & Murray, 2008). Many speech-language pathologists report insufficient knowledge and lack of guidelines in this area (87.2%; Arslan et al., 2026), alongside persistent linguistic and sociocultural barriers to equitable care (Sung et al., 2024). In this context, this scoping review aims to map existing knowledge on language disorders in Creole-speaking populations in order to inform more inclusive research and clinical practices.

Methods

A systematic search was conducted in PubMed, Scopus, and Web of Science using (*language_name*) AND (aphasia OR dementia OR language disorder) for all Creole languages listed in the APiCS atlas. This was complemented by additional manual searches.

Studies were included if they were published in English or French, involved adult Creole-speaking people with aphasia (PWA) or with other neurocognitive disorders, and reported language-related outcomes. Studies on children or adolescents and reports of patient or caregiver perspectives were excluded. For the included records, data were extracted on study characteristics, objectives, language tasks, and main findings.

Results

The systematic database search yielded 564 records, of which six studies met the inclusion criteria after screening. Additional searches identified five further eligible reports, resulting in a total of 11 included studies: 7 peer-reviewed articles, 1 conference proceeding, and 3 theses. The selection process is summarised in the PRISMA flow-diagram (*Figure 1*).

Four peer-reviewed studies focus on post-stroke aphasia. Penn et al. (2001) analysed spontaneous speech in seven Afrikaans-speaking PWA and seven matched controls, showing reduced lexical diversity, fewer verbs and embeddings in aphasia, alongside increased use of English loanwords as compensatory code-mixing. Mosdell et al. (2010) assessed three bilingual Afrikaans-English PWA and 30 controls using both standard and culturally adapted

pictures for naming (Groote Schuur Naming Test) and discourse tasks (Washing Line Picture Test); adapted versions significantly improved performance in controls and were more sensitive to impairments in patients. Penn et al. (1992) evaluated a translated version of the Communicative Effectiveness Index (CETI) in Afrikaans, Sotho and Zulu with 28 PWA (including 3 Afrikaans-speakers) and found it sensitive to language impairments across languages. Massina et al. (2000) reported a longitudinal case of a Guadeloupean Creole-French bilingual speaker: early post-stroke performance showed greater fluency and organisation in Creole despite therapy and greater improvement in sentence repetition in French. After six months both languages improved, with near-normal performance in Creole.

McDonald (2014) describes a Guyanese Creole-English speaker with aphasia, identifying a hierarchy of grammatical vulnerability: pronouns and articles were relatively preserved, whereas tense-mood-aspect particles were frequently omitted, problematising the distinction between aphasia and bilingual features. Souchette (2020) validated a bilingual assessment tool with 36 Martinican Creole-French PWA, showing that all participants presented impairments in French, but not consistently in Creole, with pragmatic abilities overall preserved. Henery (2016) compared 20 Guadeloupean Creole and 20 French speakers with aphasia in picture description tasks, reporting a preference for culturally appropriate images in both groups and greater fluency for Guadeloupean relevant images in the Creole group.

One study examined the presentation of Primary Progressive Aphasia in a trilingual Haitian Creole-French-English speaker (Kelly & Shehab, 2025), showing nonfluent speech with agrammatism and impairments in naming, fluency, repetition, and reading across all languages. Two studies addressed Alzheimer's Disease (AD). De Picciotto and Friedland (2001) assessed verbal fluency in six English-Afrikaans bilinguals with AD and 30 controls across single-language and switching conditions, showing reduced switching in AD. Kessori et al. (2023) [in French] conducted an intervention study with five Réunion Creole-French bilingual AD participants, showing improved lexical-semantic access and good compliance following a culturally adapted Semantic Feature Analysis protocol in Creole. Malcolm et al. (2018) reports a case of agrammatic aphasia in a Caribbean English-Creole bilingual speaker, with tense-marker omissions in both languages. Similar omissions in the spontaneous English speech of 20 bilingual controls suggest a possible effect of bilingualism.

Discussion

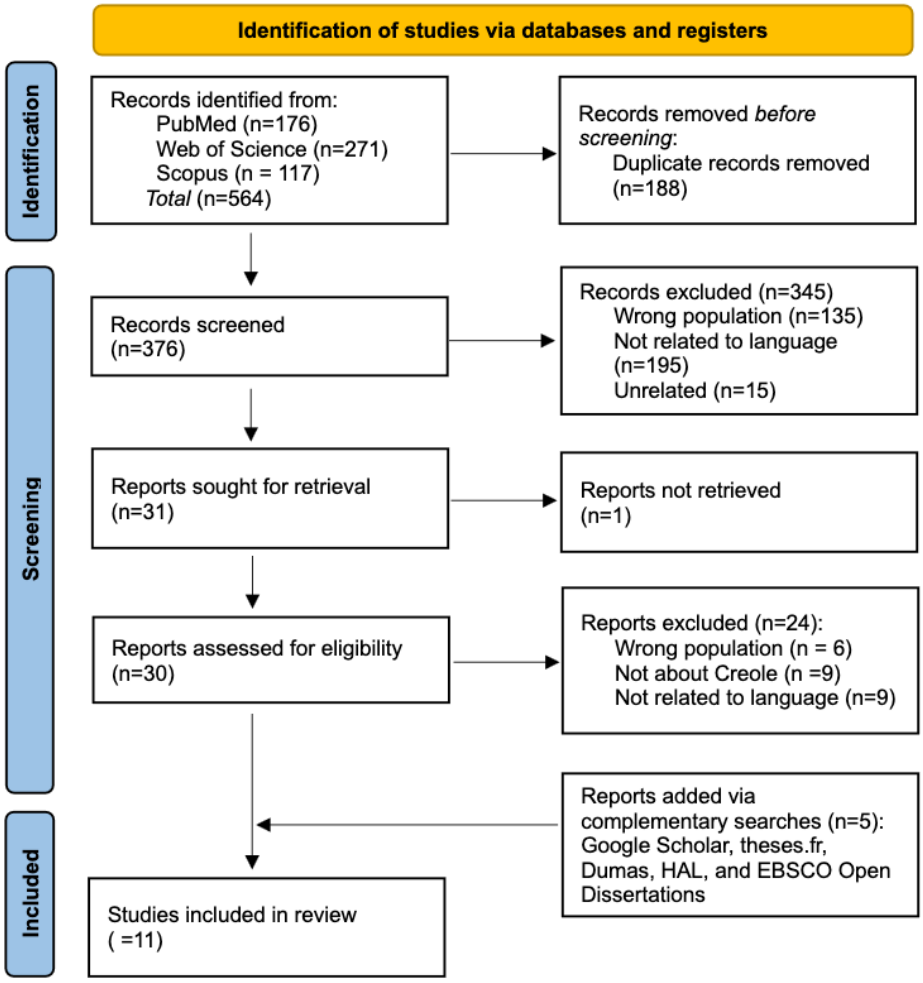
To our knowledge, this is the first scoping review to systematically map research on aphasia and neurodegenerative language disorders in Creole speakers. The small number of studies, their heterogeneity, and the predominance of limited samples reflect broader clinical and research gaps (Arslan et al., 2026). Across studies, findings support the feasibility of assessment and rehabilitation, while emphasizing the importance of culturally adapted tools to improve performance and patient engagement in both assessment (Mosdell et al., 2010; Henery, 2016) and intervention (Kessori et al., 2023). Results on bilingual performance and recovery remain inconclusive, with reports of compensatory code-switching in aphasia (Penn et al., 2001) versus reduced switching in Alzheimer's Disease (De Picciotto & Friedland, 2001).We observe differential models: poorer performance in Creole than in English (Penn et al., 2001), but better outcomes in Creole than in French (Massina et al., 2000; Souchette, 2020). Moreover, omission patterns in English may reflect bilingual interference rather than aphasia alone (Malcolm et al., 2018; McDonald, 2014). Overall, these findings highlight the need for further research on underrepresented languages (Matić et al., 2026), to better understand interactions between linguistic structure, bilingualism, and language disorders

(Lorenzen & Murray, 2008), to account for interlect and mixed forms reflecting Creole sociolinguistic backgrounds (Véronique, 2013), and to support more equitable clinical practices (Sung et al., 2024).

References

(For records included in the review, see *Figure 1*).
Arslan, S., Norvik, M.I.,... & MAP consensus group (2026). Characterizing the linguistic profiles, training needs, and caseloads of speech language pathologists providing clinical services to multilingual people with aphasia: The international Multilingual Aphasia Practices (MAP) consensus group survey. *Plos one*, *21*(4), e0346488.
Holm, J. (2000). *An introduction to pidgins and creoles*. Cambridge University Press.
Lorenzen, B., & Murray, L.L. (2008). Bilingual aphasia: A theoretical and clinical review. *American Journal of Speech-Language Pathology*, *17*(3), 299-317.
Martínez-Ferreiro, S., Arslan, S.,... & Soroli, E. (2026). Guidelines and recommendations for cross-linguistic aphasia assessment: A review of 10 years of comprehensive aphasia test adaptations. *Aphasiology*, *40*(2), 215-239.
Matić, A., Marie, P.,... & Arslan, S. (2026). Setting a Research Agenda for the Assessment and Treatment of Aphasia in Minority Languages. *Cortex 198,* 13-26.
Menn, L. & Obler, L. K. (1990). *Agrammatic aphasia: A cross-language narrative sourcebook*. John Benjamins Publishing Company.
Mufwene, S. S. (2001). *The Ecology of Language Evolution. Cambridge Approaches to Language Contact*. Cambridge University Press.
Sung, J. E., Scimeca, M., Li, R., & Kiran, S. (2024). Cross-linguistic and multicultural considerations in evaluating bilingual adults with aphasia. *American Journal of Speech-Language Pathology*, *33*(6), 2716-2731.
Véronique, G.D. (2013). Émergence des langues créoles et rapports de domination dans les situations créolophones. *In Situ. Revue des patrimoines*, (20).

Figure 1. PRISMA 2020 flow diagram for Scoping Reviews.



Included studies:

De Picciotto, J., & Friedland, D. (2001). Verbal fluency in elderly bilingual speakers. *Folia Phoniatrica et Logopaedica*, *53*(3), 145-152.

Henery, E. (2016). *Influence d'un support de rééducation culturellement adapté sur les productions langagières des patients aphasiques: application à la population guadeloupéenne*. Master's thesis.

Kelly, B.E., & Shehab, A.A.S. (2025). A Case Study of a Multilingual Individual with Primary Progressive Aphasia. *Archives of Clinical Neuropsychology*, *40*(6), 1266-1274.

Kessori, C., Boutin, A. P., & Michalon, S. (2023). KOSA IN SOZ? Création d'un outil de rééducation lexico-sémantique à destination de patients atteints de la maladie d'Alzheimer, créolophones réunionnais et de niveau socio-culturel modeste. *Revue scientifique des travaux de fin d'étude en rééducation et réadaptation*.

Malcolm, T.R., Hejazi, Z., & Obler, L.K (2018). Agrammatic Aphasia in a Bidialectal/Bilingual Speaker of Caribbean English Creole and Standard English. *Academy of Aphasia*.

Massina, C., Le Gall, D., Aubin, G., Mazaux, J.M., Galanthe, E., Sainte-Foie, S., & Emile, J. (2000). Une observation de la récupération différentielle des deux langues chez une patiente aphasique bilingue français-créole guadeloupéen. *Annales de réadaptation et de médecine physique*, 43(8), 450-464.

McDonald, R. (2014). *Characterizing non-fluent aphasia in English-based Caribbean creole languages: A case study*. Master's thesis.

Mosdell, J., Balchin, R., & Ameen, O. (2010). Adaptation of aphasia tests for neurocognitive screening in South Africa. *South African Journal of Psychology*, *40*(3), 250-261.

Penn, C., Milner, K., & Fridjhon, P. (1992). The communicative effectiveness index: Its use with South African stroke patients. *South African Journal of Communication Disorders*, *39*(1).

Penn, C., Venter, A., & Ogilvy, D. (2001). Aphasia in Afrikaans: A preliminary analysis. *Journal of Neurolinguistics*, *14*(2-4), 111-132.

Souchette, L. (2020). *Validation of an oral language assessment tool for bilingual aphasics from Martinique in French and Creole*. Master's thesis.

Language assessment using immersive virtual-reality relative to standard tests: results on neurotypical adults and preliminary validation on people with aphasia

Caroline Rogue | Marina Laganaro

Introduction and aims

To diagnose and qualify language impairments in post-stroke aphasia, speech and language therapists usually use assessments primarily based on standard presentation formats (on paper or on computer screen). The performance of people with aphasia (PWA) may differ substantially between clinical static assessments and real-life communication conditions. For instance, it is well established in the literature that even the mere addition of a static background can influence naming performance and that any type of distractor noise has an impact on performance in particular in PWA (Demierre et al., 2024). Therefore, embedding language assessment in more realistic contexts may have an impact on language processing (Davenport & Potter, 2004) and therefore better highlight the impact of mild impairments in real life. These observations highlight the need to develop assessment approaches that better approximate real-life communication environments.

Reaching a higher degree of immersion can now be achieved through immersive virtual-reality (iVR), a computer-generated 3D interactive environment. Such dynamic environment provides an immersion that closely mimics real-life experience through multi-modality inputs, with distractors and interferences that can be encountered in everyday-life.

Previous research has yielded mixed findings regarding the impact of VR immersion on cognitive task performance. On the one hand, some studies showed that VR can enhance language learning (Ye & Kaplan-Rakowski, 2024). More recently, Franco's et al. showed that learning words through an iVR application led to higher word production accuracy in neurotypical adults, as well as lower rate of lexical errors in PWA, compared to a standard static digital method (Franco et al., 2025).On the other hand, despite evidence supporting the potential cognitive benefit of VR, it is important to consider the additional cognitive cost associated with immersive environments, which may lead to decreased performance (Neguţ et al., 2016). These contrasting findings suggest that traditional assessments tasks may not capture the complexity of linguistic processes as they occur in everyday life situations.

Finally, beyond these advantages, iVR also offers a very high degree of experimental control and replicability (compared to everyday situations variability). This represents a key benefit for balancing ecological validity with methodological rigor.

Taken together, published tools and state of the art indicate that cognitive and language performances might be different in a dynamic environment provided by virtual reality than in the standard static environment.

In order to assess language in a closer to real life communication context, this study presents the elaboration of an immersive virtual reality language assessment in French (Aphasia Assessment), along with its comparison to static assessment and standardization in a neurotypical adult population and preliminary validation in PWA.

Methods

To date, 63 neurotypical native French-speaking adults (29 men and 34 women, M=51.9 years old, SD= 20.12, range:18-89) were recruited, and we plan to recruit 20 PWA with mild to moderate language impairment.

The experiment is carried out over two sessions, during which each participant is exposed to two assessment conditions (a static computer-based and an immersive virtual reality condition) in a counter-balanced order.

In both conditions, participants perform 5 tasks: (task 1) a scene description task, (task 2) a short standardized picture-naming task (Macoir et al., 2016), (task 3) a word comprehension task, (task 4) a blocked-cyclic picture-naming task and (task 5) a narrative task. In the immersive condition the tasks were embedded in a virtual marketplace environment and the participants had to (1) describe the market; buy specific objects by naming/asking them to the avatars (for objects at the market (task 2 and 4); take objects mentioned by the avatar (task 3) and describe at the end their visit to the market (task 5).

For the free speech tasks (1 and 5), words per minute, CIUs per minute and percent CIUs are computed. Accuracy and latencies are collected and analyzed for tasks 2, 3 and 4.

On the neurotypical speakers, analyses are run (i) to compare performance between the two conditions and (ii) to develop normative data stratified by age and gender when these factors significantly influence performance. The performance of people with aphasia will then be compared to neurotypical values in the two conditions.

Results

Preliminary results showed a significant difference in CIUs per minute ($t(61) = 2.00$, $p = .04$), indicating a benefit of the VR condition for speech informativeness. Additionally, productions in the immersive condition tended to exhibit greater syntactic complexity than those in the standard condition, although this difference did not reach statistical significance ($t(61) = -1.84$, $p = .068$).

Naming latencies in the blocked-cyclic picture-naming task did not differ significantly across conditions. The virtual reality presentation did not enhance nor reduce the semantic interference effect relative to the standard static presentation ($\beta = 8.28$, $SE = 31.16$, $t(4154.21) = 0.27$, $p = .79$).

In PWA we expect assessment via iVR to demonstrate greater sensitivity relative to the standard assessments in detecting language impairments, while maintaining high specificity reflected in low rates of false negatives. In addition, assessment outcomes are expected to show strong correlations with established tests measuring the same constructs, supporting convergent validity.

References

Davenport, J. L., & Potter, M. C. (2004). Scene Consistency in Object and Background Perception. *Psychological Science*, *15*(8), 559-564. https://doi.org/10.1111/j.0956-7976.2004.00719.x

Demierre, C., Python, G., Glize, B., & Laganaro, M. (2024). How dual-task interference on word production is modulated by the timing of the secondary task : Evidence from errors in people with aphasia. *Aphasiology*, *38*(6), 1028-1050. https://doi.org/10.1080/02687038.2023.2253567
Franco, J., Glize, B., & Laganaro, M. (2025). Impact of immersive virtual reality compared to a digital static approach in word (re)learning in post-stroke aphasia and neurotypical adults : Lexical-semantic effects? *Neuropsychologia*, *208*, 109069. https://doi.org/10.1016/j.neuropsychologia.2025.109069
Macoir, J., Gauthier, C., Jean, C., & Potvin, O. (2016). BECLA, a new assessment battery for acquired deficits of language : Normative data from Quebec-French healthy younger and older adults. *Journal of the Neurological Sciences*, *361*, 220-228. https://doi.org/10.1016/j.jns.2016.01.004
Neguţ, A., Matu, S.-A., Sava, F. A., & David, D. (2016). Task difficulty of virtual reality-based assessment tools compared to classical paper-and-pencil or computerized measures : A meta-analytic approach. *Computers in Human Behavior*, *54*, 414-424. https://doi.org/10.1016/j.chb.2015.08.029
Ye, Y., & Kaplan-Rakowski, R. (2024). An exploratory study on practising listening comprehension skills in high-immersion virtual reality. *British Journal of Educational Technology*, *55*(4), 1651-1672. https://doi.org/10.1111/bjet.13481

Investigating deverbal nouns in individuals with neurodegenerative syndromes: evidence from Slovene

GEORGIA ROUMPEA | CHRISTINA MANOUILIDOU

Introduction and aims

Alzheimer's disease (AD) is characterized by progressive cognitive decline, including memory loss and language deficits. Previous research on morphologically simple words, like verbs and nouns, has revealed impairment in naming (de Almeida et al., 2021), inflectional morphology (e.g., aspect and tense, Manouilidou et al., 2020), and the processing of pronounceable (**helk*) and non-pronounceable non-words (**atth*) (Synder et al., 1996). While morphologically simple words have been extensively studied, more complex forms, like deverbal nouns (*reader*) remain relatively understudied in AD. Extensive evidence on deverbal nouns (DVNs) comes from healthy adults and other clinical populations. Studies on individuals with semantic dementia (SD) and Parkinson's disease (PD) indicate that the ability to derive DVNs varies by cognitive profile. Kavé et al. (2012) reported preserved DVN morphology in SD, while Silveri et al. (2017) showed that executive dysfunction in PD impairs DVN production from verbal bases (e.g., read → reader). Additional evidence on DVNs comes from studies examining how morphological rules and constraints on their formation affect processing, using acceptability judgment (AJT) and lexical decision (LDT) tasks, as well as pseudowords (i.e., interpretable and pronounceable lexical items that violate word formation constraints). Manouilidou (2006) showed that healthy adults rejected pseudowords as real words and were able to differentiate between different types of word-formation violations in AJT and LDT tasks. Specifically, the study found that constraints related to the lexical category of the base (e.g., **potiritis* 'glasser', the Greek suffix *–is* allows a verbal base but not a nominal one) are stricter to violate than those related to thematic features (**redie,* the prefix *re–* does not allow unaccusative verbs). Continuing with evidence from clinical populations, Manouilidou et al. (2016) investigated Slovene-speaking individuals with mild cognitive impairment (MCI). Cognitive limitations did not affect pseudoword detection in the AJT, with the MCI group performing comparably to the controls, suggesting preserved knowledge of the rules. However, in the LDT, individuals with MCI showed reduced accuracy and slower reaction times (RTs) for pseudowords than controls. Individuals with MCI also differentiated between pseudoword types, supporting previous findings that lexical-category constraints (**čokoladec* 'the one that chokolates') are harder to violate than thematic (**umiralec* 'dier') and aspectual (**preplavalec* 'the one that swam') constraints.

The current study aims to investigate Slovene DVNs in individuals with dementia to better understand how derivational morphology is affected in this population. Moreover, Slovene is an understudied, morphologically rich language compared to Greek and English, offering a fruitful ground for cross-linguistic evidence.

Methods

Participants

Eighteen age-matched and unimpaired controls, four mild-to-moderate AD, and four individuals with undefined dementia were recruited (Slovene native speakers). The

neuropsychological assessment of the clinical groups (MMSE, Stroop task) indicated executive function limitations.

Experimental tasks and materials

The current study followed the paradigm of Manouilidou et al. (2016) and expanded it by including an additional suffix and pseudoword category. Two suffixes that form DVNs in Slovene were used: *–ec* for agent-DVNs and *–je* for process-DVNs.

Three tasks were designed:

1. Derivation task: produce the DVN from a verbal base and vice versa (30 verbs *brati* 'to read', 30 agent-DVNs *bralec* 'reader', 30 process-DVNs *branje* 'reading')

2. Acceptability judgment and lexical decision tasks – same set of materials:

- **real words** (30 non-derived *brati*, 60 derived words – DVNs)

pseudowords:

- 60 categorial violations (**čokoladec*, suffixes *–ec* and *–je* attach to verbal bases not nominal bases)
- 30 thematic violations (**ljubilec*, *–ec* does not allow verbal bases with an experiencer as a subject)
- 54 aspectual violations (**preplavalec, *preplavanje –ec* and *–je* allow imperfective and not perfective verbal bases)
- 30 inherent-aspectual violations (**storjenje*, *–je* does not allow inherently perfective verb bases). Aspectual violations involve an overt aspectual marker on the verbal base (prefix e.g., pre + plavati → preplavati). In inherent-aspectual violations, perfectivity is inherent to the verb itself.
- 30 pronounceable non-words **fekulteta*
- 30 possible words **kuhalec* that are blocked by existing ones *kuhar*.

Results

In the derivation task, no effect of group ($\chi^2 = 3.8605$, df = 2, p = .145) was reported suggesting comparable performance between the clinical groups and the controls. In the AJT, the AD group showed significantly lower accuracy compared to the controls in categorial ($\beta = 1.389$, $p < .001$) and aspectual violations ($\beta = .599$, $p = .035$). Within the mild-to-moderate AD group, performance across different pseudoword types remained comparable. In the LDT, the mild-to-moderate AD group was significantly less accurate and showed longer RTs when detecting pseudowords compared to the controls. Within the mild-to-moderate AD group, participants achieved significantly lower accuracy in rejecting pseudowords compared to real words. Moreover, significantly lower accuracy was observed in the inherent-aspectual violations and the aspectual violations compared to categorial ($\beta = 1.946$, $p = .010$ and $\beta = 1.485$, $p = .007$) and thematic violations ($\beta = 1.476$, $p < .001$ and $\beta = 1.015$, $p < .001$). No dissociation between categorial and thematic violations was reported. The RTs showed no differences among pseudowords. The undefined dementia group showed a pattern of performance similar or more impaired to that of the mild-to-moderate AD group in both tasks.

Figure 1 shows mean accuracy for the clinical groups compared to the controls in the LDT (* indicates significance).

Discussion

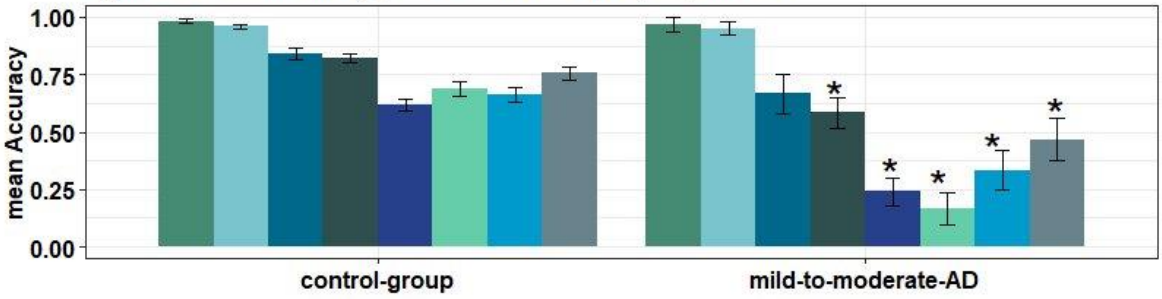
The derivation task results suggest that knowledge of DVN morphology is relatively preserved across groups despite cognitive limitations. However, as task demands increase (detection of violations, time pressure) in AJT and LDT, executive dysfunction (inhibition, processing speed) in the clinical groups potentially leads to reduced performance. Consistent with previous findings, certain morphological constraints are more prominent than others in DVN processing. Categorial constraints are highly sensitive and easily recognized, as has been established in the literature. However, unlike findings from other languages, Slovene-speaking individuals with mild-to-moderate AD and undefined dementia treat thematic and categorial constraints with equal importance. Aspectual restrictions (marked with prefix and inherently), by contrast, are more prone to violation. Crucially, since aspect has not been extensively investigated in AJT and LDT studies, the current findings offer valuable insights into how grammatical aspect interacts with DVN processing.

References

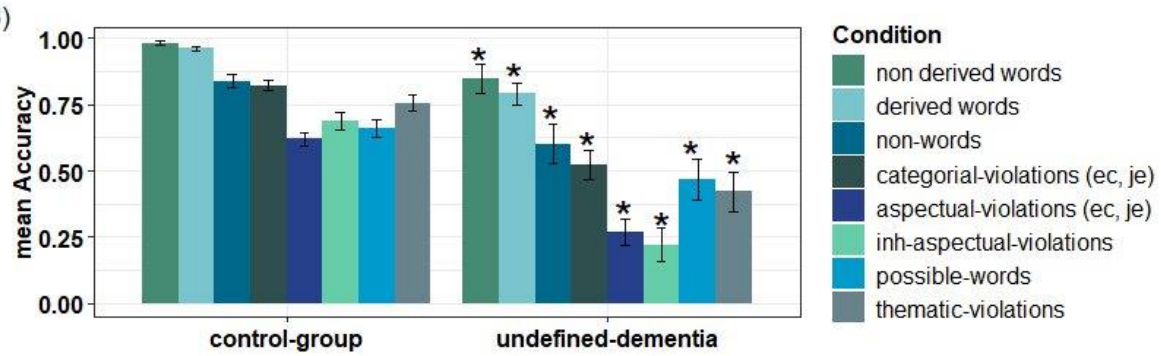
de Almeida, R. G., Mobayyen, F., Antal, C., Kehayia, E., Nair, V. P., & Schwartz, G. (2021). Category-specific verb-semantic deficits in Alzheimer's disease: Evidence from static and dynamic action naming. Cognitive Neuropsychology, 38(1), 1-26.
Kavé, G., Heinik, J., & Biran, I. (2012). Preserved morphological processing in semantic dementia. Cognitive Neuropsychology, 29(7-8), 550-568.
Manouilidou, C. (2006). On the processing of thematic features in deverbal nominals (Doctoral dissertation, University of Ottawa (Canada).
Manouilidou, C., Dolenc, B., Marvin, T., & Pirtošek, Z. (2016). Processing complex pseudowords in mild cognitive impairment: the interaction of preserved morphological rule knowledge with compromised cognitive ability. Clinical Linguistics & Phonetics, 30(1), 49-67.
Manouilidou, C., Roumpea, G., Nousia, A., Stavrakaki, S., & Nasios, G. (2020). Revisiting aspect in Mild Cognitive Impairment and Alzheimer's disease: evidence from Greek. Frontiers in Communication, 5, 119.
Silveri, M. C., Traficante, D., Monaco, M. R. L., Iori, L., Sarchioni, F., & Burani, C. (2017). Word selection processing in Parkinson's disease: When nouns are more difficult than verbs. Cortex, 100, 8-20.
Snyder, L. S., Holland, A. L., & Forbes, M. (1996). Lexical decisions in patients with Alzheimer's disease: Some notes on automatic versus controlled processing. Journal of communication disorders, 29(5), 389-399.

(A)

Figure 1: mean accuracy per condition and per group in LDT



(B)



Dynamic assessment of everyday communication using Virtual Reality (DCOM-VR)

Marina B. Ruiter | Vitória Piai | Hilde Bosschers | Mirjam Otters | Lizet van Ewijk | Edward Bosma | Willemijn J. Doedens

Introduction and aims

Functional communication skills represent important measures in aphasia therapy. For the assessment of functional communication in Dutch-speaking persons with aphasia (PWA), two pen-and-paper tests are currently available for clinical practice: the Scenario Test [1] and the Amsterdam-Nijmegen Everyday Language Test [ANELT, 2]. Both assessments adopt a similar testing approach involving role-playing. In each test, the examiner orally presents an everyday communication scenario and subsequently prompts the PWA with questions such as "What would you say?".

For the resulting scores to be valid indicators of functional communication, participants must - among other things - comprehend and remember each orally presented scenario, switch to the required role, maintain this perspective while generating their response in direct speech, and switch to the nonverbal communication channels (e.g., pointing or gesturing) when language difficulties are encountered. This suggests that responding to scenarios requires not only language processing, but also non-linguistic cognitive functions such as attention, working memory, and executive control [3]. Because PWA commonly frequently exhibit non-linguistic cognitive impairments [4] that particularly hinder pretend role-playing, pen-and-paper based functional communication tests often elicit indirect rather than direct speech. As a result, task impurity is a particular concern in the functional communication assessment in this population.

In addition to concerns about task impurity, limitations in ecological validity may comprise construct validity. In daily communication, PWA are often less able to use their remaining language skills because they face additional non-linguistic cognitive challenges, such as time pressure to communicate information quickly and interference of background noise [5], as well as limited effort and patience on the part of communication partners [6]. In current standardized, pen-and-paper tests of functional communication - typically administered in highly controlled and quiet clinical environments - these personal, social and environmental demands are either absent or present to a much lesser extent than in everyday communication. As a result, the outcome measures obtained from these tests may not be fully valid indicators of daily functioning.

We hypothesised that immersive Virtual Reality (iVR) may be equipped to mitigate the task impurity and ecological validity issues discussed above. Immersion - defined as the psychological state in which individuals perceive themselves as being enveloped by, included in, and interacting with the VR-environment - could support PWA in assuming a role during the functional communication tests, leading to the production of direct speech and thereby improving construct validity. iVR might also support language comprehension during scenario introduction, help PWA to keep the scenario active in working memory, and allow for controlled manipulation of real-world communication demands, for example by adding environmental distractions, such as background noise.

We therefore examined whether iVR can provide a realistic simulation of everyday communication, focusing on its technical feasibility and user acceptability, and whether iVR offers added value for functional communication assessment by enabling manipulation of communicative complexity, and eliciting direct speech.

Methods

A total of 101 Dutch-speaking participants took part in one or more of four project stages, following the design thinking process. Phase 1, 'Understand', focused on identifying the needs, current practices, and challenges experienced by speech and language therapists (SLTs, $n = 38$) in the pen-and-paper–based assessment of functional communication. The second phase, 'Explore', focused on deepening the understanding and prioritisation of the needs and requirements raised by SLT in Phase 1, alongside exploring the perspectives of PWA regarding everyday communication, its challenges, and their priorities regarding their implications for assessment. To that aim we organised focus groups with SLT ($n = 5$) and PWA ($n = 11$). The third Phase, 'Materialize', encompassed the technical development of a proof-of-concept (POC) version of the iVR-based assessment instrument. Building on our previous research on the ANELT [7, 8], we used, with permission, the Child scenario from the original pen-and-paper Dutch version of the ANELT [2] and the Train Station scenario, taken from the Dutch and adapted version of this test [7].Three conditions were created for each scenario: a) a standard condition, in which the virtual interlocutor was patient and understanding and in no background noise was present, b) a 'noise' condition, in which there was background noise was added, while the virtual interlocutor remained patient, c) an 'impatient' condition, in which there was no background noise, but the interlocutor responded impatiently. The final Phase, 'Testing', involved evaluating the beta version of the POC iVR-based assessment instrument through individual user-centred test sessions with both PWA ($n = 7$) and SLT ($n = 6$). During this phase, the iVR-based assessment was additionally administered to neurologically healthy control participants with no professional background in speech and language therapy (NHP, $n = 40$).

Results

Both SLT and PWA generally perceived the environment as realistic, understood what was expected of them, and indicated that they would respond in a manner similar to how they do in everyday communicative situations. In line with our hypothesis regarding immersion, PWA, SLT, and NHP produced only direct speech when responding appropriately to the scenarios. Despite suboptimal simulations of impatience in the virtual conversation partner across both scenarios - participants generally reported that the impatient condition of the Train Station scenario and the noise condition of the Child scenario were more difficult to respond to than the standard conditions. This suggests that, although further refinement of the virtual interlocutor's impatient behaviour is required, manipulating communicative complexity is technically feasible.

Discussion

The results obtained suggest that it is technically feasible to create iVR environments for the assessment of functional communication that are acceptable to (end) users and that enhance the construct validity of these tests. Despite remaining areas for improvement—such as enhancing the sense of time pressure through more effective non-verbal representations of the virtual interlocutor's impatience—iVR was found to have a clear added value for diagnostic

purposes. In particular, the increased immersion in the scenarios, and the resulting consistent use of direct speech, appear to contribute to a more accurate measurement of functional communication skills.

Funding
This work was supported by the NeuroLabNL: Small Projects for NWA routes 21/22, Dutch Research Counsil (NWO) under Grant NWA.1418.22.025.

References

1. van der Meulen, I., et al., *Scenario Test [test manual]*. 2008, Bohn Stafleu van Loghum: Houten.
2. Blomert, L., C. Koster, and M.-L. Kean, *Amsterdam-Nijmegen Test voor Alledaagse Taalvaardigheden [ANELT] [test manual]*. 1995, Swets & Zeitlinger: Lisse.
3. Olsson, C., P. Arvidsson, and M. Blom Johansson, *Relations between executive function, language, and functional communication in severe aphasia.* Aphasiology, 2019. **33**(7): p. 821–845.
4. Akkad, H., et al., *Mapping spoken language and cognitive deficits in post-stroke aphasia.* Neuroimage Clin, 2023. **39**: p. 103452.
5. Harmon, T.G., E. McDonald, and K. Steele, *Effects of cognitive and social demands on linguistic production for people with moderate, mild, or no aphasia.* Aphasiology, 2024. **38**(2): p. 281–304.
6. Dalemans, R.J.P., et al., *Social participation through the eyes of people with aphasia.* International Journal of Language & Communication Disorders, 2010. **45**(5): p. 537–550.
7. Ruiter, M.B., et al., *A transcription-less quantitative analysis of aphasic discourse elicited with an adapted version of the Amsterdam-Nijmegen Everyday Language Test (ANELT).* Aphasiology, 2023. **37**(10): p. 1556–1575.
8. Ruiter, M.B., et al., *Towards a quantitative measure of verbal effectiveness and efficiency in the Amsterdam-Nijmegen Everyday Language Test (ANELT).* Aphasiology, 2011. **25**(8): p. 961–975.

Longitudinal subjective accounts of co-occurring language and memory complaints in glioma patients undergoing brain surgery and comparison to neuropsychological assessment

Lena Rybka | Prof. Roel Jonkers | Milena Burzlaff | PD Dr. Julia Onken | Prof. Katharina Faust | Prof. Peter Vajkoczy | Prof. Thomas Picht | PD Dr. Tizian Rosenstock | Dr. Adrià Rofes

Introduction and aims

Language is commonly assessed in glioma patients (Rofes et al., 2017). However, language processing also relies on short-term or working memory (henceforth „memory"), such as in reading, conversation, writing, and sentence processing (Schwering & MacDonald, 2020). While the interaction of language and memory has been discussed in other clinical populations, research in glioma patients is scarce. First insights into the importance of those interactions in glioma patients has been reported by Rybka and colleagues (2024). Furthermore, analyses of subjective complaints and neuropsychological testing have shown discrepancies (Brownsett et al., 2019; Pertz et al., 2022), possibly caused by test standardization in other clinical populations. In the context of gliomas, their surgical and adjuvant treatment may also add to the symptom burden. Consequently, cognitive symptoms may change across the disease trajectory.

The aim of this study is hence twofold: to understand the extent of subjective co-occurring language and memory complaints in glioma patients longitudinally, including issues in specific language domains (i.e. conversation, sentence processing, reading, writing); and to gauge differences between subjective complaints and neuropsychological testing.

Methods

Forty-one native German speakers with a supratentorial glioma underwent a semi-structured interview and a neuropsychological test battery comprising the Aachen Aphasia Test (AAT), a digit span forward (DSF) and backward test (DSB). Assessments were conducted up to four times across the disease trajectory (preoperatively: T0; 2-5 months postoperatively: T1; 6-12 months postoperatively: T2; 12+ months postoperatively: T3).

The semi-structured interview included questions relating to language and memory during conversation, sentence processing, reading, and writing. Co-occurring subjective complaints were noted by LR. A clear distinction between short-term and working memory in the interview responses, was not always feasible. Consequently, these were accumulated as memory complaints.

The following tests from the AAT were chosen to match the domains of the interview: reading, writing, and the Token Test (for conversation). Sentence repetition, sentence comprehension, and sentence production were compiled into one score for sentence processing. A test performance below 90% was considered below norm. Patients were classified as impaired in memory if their performance differed significantly from healthy controls in the DSF or DSB tests on single case statistics.

Four complaint profiles were identified for each timepoint given subjective complaints and the neuropsychological test: memory complaints, language complaints, co-occurring complaints in memory and language, and no complaints.

Descriptive statistics were reported and differences in co-occurring subjective complaints and neuropsychological deficits were checked with Fisher's exact tests (Bonferroni corrected p-values).

Results

Across assessment times, 61% of patients dropped out of the study (e.g. due to recurrence). Therefore, 41 patients were assessed at T0, 27 patients at T1, 23 patients at T2, and 16 patients at T3.

Extent of Subjective Co-occurring Complaints of Memory and Language Longitudinally

Across the assessment period, four trajectory types were identified: 37% of patients reported persistent co-occurring complaints, while 20% of patients indicated new ones. 13% of patients indicated recovery from co-occurring complaints, while the remaining 30% reported no deficits across the trajectory. Relative to specific language domains, at T0, co-occurring complaints in conversation were most frequent (69%), followed by sentence processing and reading (54%). Longitudinally, co-occurrences in conversation were also reported, though less frequent (T1 and T2: 27%; T3: 17%). At T1 and T2, co-occurrences were most common in sentence processing (73% and 64% respectively), and reading complaints in 27% of patients at these timepoints. At T3, most co-occurrences were reported in reading (50%), and complaints in sentence processing declined (33%).

Subjective Complaints vs. Neuropsychological Deficits

A discrepancy in the groups was found: while co-occurring complaints was the largest group for subjective complaints at T0 (42%), T1 (63%), and T2 (57%), co-occurring deficits in neuropsychological assessments were less frequent (discrepancy for T0: 21,95%; T1: 33,33%; T2: 30,43%). Although both assessment types indicate the "no deficit" group as the largest group at T3, a discrepancy of 10% remains. Fisher's exact tests did not reveal statistically significant differences at any timepoint. Although not statistically significant, a trend towards a higher match rate across the trajectory can be observed (T0 and T1: 59%, T2: 74%, T3: 75%). To determine what may contribute to that improved match rate, further analysis on the different combinations of subjective complaints and neuropsychological deficits showed that discordant categories (e.g. both complaints but only language deficit) were most prevalent at T0 (82%) and decreased at later timepoints (T1 and T2: 71%, T3: 60%). Overall, there were more discordant categories with present subjective complaints but absent neuropsychological deficits.

Discussion

Regarding our first aim, co-occurring language and memory complaints were frequent in glioma patients and occurred longitudinally, with new co-occurrences arising throughout the trajectory. These results resonate with the literature, whereby subjective complaints in language and memory have been reported (Brownsett et al., 2019; Pertz et al., 2022). Regarding specific language domains, we found subjective complaints of conversation,

reading, and writing, similar to a previous study of our group (Rybka et al., 2024). In addition, this study addressed subjective co-occurring language and memory complaints in sentence processing.

As for the second aim, subjective complaints were more often reported than neuropsychological deficits detected. This has been partly reported in the previous literature for language (Brownsett et al., 2019) and memory (Pertz et al., 2022), and linked to an insensitivity of many standardized tests. However, these differences were not statistically significant in this study which may be driven by a small sample size or high attrition rate, but also within-group variability and rare events. Additional data from larger cohorts is therefore needed to confirm these findings. Nevertheless, subjective accounts may provide additional insights into the impact on patients' daily functioning. This may allow identifying functions that are most critical to preserve for the individual patient, and may complement clinical decision making, for example to adapt intraoperative test protocols or rehabilitative measures.

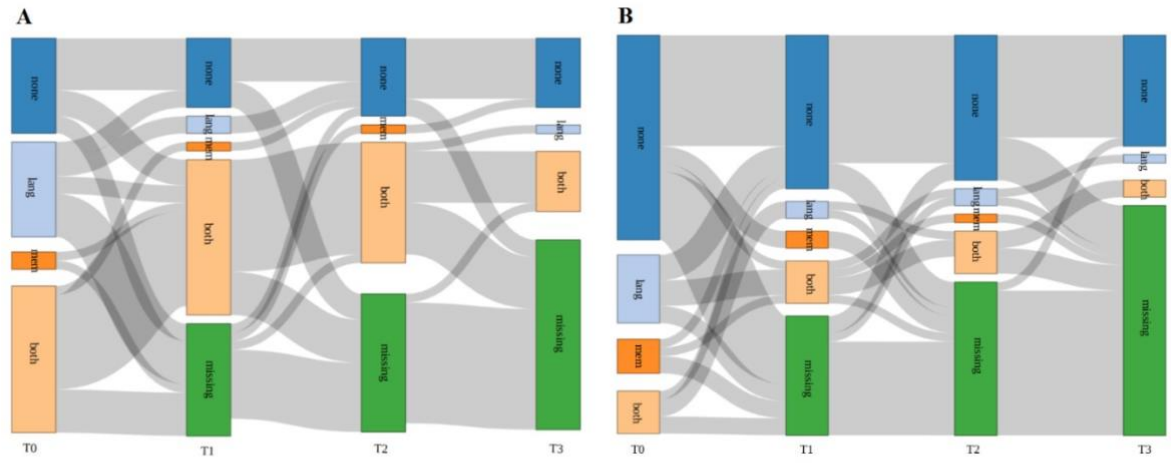
In sum, co-occurring language and memory deficits and complaints were present in glioma patients across all timepoints. Despite lack of statistical significance, subjective accounts may complement neuropsychological testing to reflect the relevance for patients' lived experience.

References

Brownsett, S. L. E., Ramajoo, K., Copland, D., McMahon, K. L., Robinson, G., Drummond, K., … De Zubicaray, G. (2019). Language deficits following dominant hemisphere tumour resection are significantly underestimated by syndrome-based aphasia assessments. *Aphasiology, 33*(10), 1163–1181.
Pertz, M., Kowalski, T., Jetschke, K., Schmieder, K., Schlegel, U., & Miller, D. (2022). Pre- and postoperative self-reported and objectively assessed neurocognitive functioning in lower grade glioma patients. *Journal of clinical neuroscience, 106*, 185–193.
Rybka, L., Jonkers, R., Burzlaff, M., Rosenstock, T., Vajkoczy. P., Picht, T., Faust, K., & Rofes, A. (2024). Preoperative subjective impairments in language and memory in brain tumor patients. *Frontiers in Oncology, 14*, 1475860.
Schwering, S. C., & MacDonald, M. C. (2020). Verbal Working Memory as Emergent from Language Comprehension and Production. *Frontiers in human neuroscience, 14*, 68.
Rofes, A., Mandonnet, E., Godden, J., Baron, M. H., Colle, H., Darlix, A., … Wager, M. (2017). Survey on current cognitive practices within the European Low-Grade Glioma Network: towards a European assessment protocol. *Acta neurochirurgica, 159*(7), 1167–1178.

Figure 1

Sankey Plots on Changes in Subjective Complaints (A) and Neuropsychological Deficits (B) Across Assessment Timepoints



Abbreviations: none = no complaints/deficits, lang = language complaints/deficits, mem = memory complaints/deficits, both = both language and memory complaints/deficits

Colors: dark blue = none, light blue = lang, dark orange = mem, light orange = both

Main Concept Analysis in Catalan and Spanish: performance of unimpaired bilingual speakers

Io Salmons | Eduard Artés | Helena Muntané-Sánchez | Àgueda Munar

Introduction and aims

The Main Concept Analysis (MCA; Kong, 2009) measures oral discourse efficiency by considering the number, completeness, and accuracy of main concepts produced. Main concepts are defined as those statements that provide essential information and contain at least one main action (Nicholas & Brookshire, 1995). The MCA has been useful for identifying discourse deficits in individuals with acquired language disorders in different languages, such as Mandarin or English (e.g., Kong et al., 2012; 2016; Gao et al., 2016).

This study is part of the ongoing adaptation of the MCA into Catalan, which is a minoritized language whose speakers are mostly bilingual. Despite increasing efforts to expand the multilingual resources in clinical settings, individuals with language disorders are often assessed and treated in only one language partly due to a lack of assessment tools in their L1, especially in minority languages (Matić et al., 2026). This study examined the production of main concepts in Catalan (L1) and Spanish (L2) by unimpaired bilingual speakers to compare performance across languages and to investigate whether an L1 advantage exists in oral discourse. Given the frequent language-mixing behavior in oral discourse by Catalan-Spanish bilinguals, establishing baseline performance in both languages is a necessary step before the MCA can be used clinically with Catalan-speaking individuals with language disorders. To our knowledge, this is the first examination of MCA performance in bilingual speakers.

Methods

Participants

A total of 31 bilingual speakers participated in the study (mean age = 46.06 years, range: 22–74 years; 18 females and 13 males; 2 left-handed). They were all native speakers of Catalan (L1) and Spanish (L2) and had been exposed to them before the age of 5. The participants came from different territories in Spain, specifically the Balearic Islands, Catalonia, and Valencia. They were all proficient in both languages, although some reported Catalan as their dominant language, whereas others described themselves as balanced bilinguals.

Materials and procedure

The participants completed the MCA in both Catalan and Spanish, in which they were asked to describe four sets of sequential pictures depicting different characters and events. Twenty-one participants started with Catalan, and 10 started with Spanish. Descriptions were recorded, and each main concept was evaluated and scored according to the measures of the MCA, which are based on Nicholas & Brookshire's (1995) proposal: accurate and complete (AC; 3 points); accurate but incomplete (AI; 2 points) if relevant information was omitted or not specific; inaccurate (IN; 1 point) if the utterance contained errors or irrelevant information; and absent (AB; 0 points). The main concept (MC) score for each language was calculated by summing the points across items.

Results

Despite participants showing near-ceiling performance (> 87%) in both languages, their results in Catalan (L1) were consistently, though slightly, higher than Spanish (L2) across all sets (Table 1: 1.3–3.9% difference). The difference between the two languages was significant in sets 1 and 4, and in the total MC score ($p < 0.05$), while no differences were found in sets 2 and 3. Moreover, greater intersubject variability was observed in L2 (SD = 7.6) than in L1 (SD = 4.7) MC scores.

Table 1. Mean scores (and standard deviations) by set

Table 1. Mean scores (and standard deviations) by set

Set		Catalan	Spanish	Wilcoxon
1	raw	11.7/12 (0.9)	11.6/12 (1.1)	V = 15, p = 0.048
	%	97.8 (7.7)	96.5 (9.6)	
2	raw	16.1/18 (2.1)	15.7/18 (2.1)	V = 129.5, p = 0.155
	%	89.4 (11.9)	87.3 (11.7)	
3	raw	13.8/15 (1.7)	13.3/15 (2.3)	V = 66, p = 0.157
	%	92.3 (11.2)	88.4 (15.6)	
4	raw	17/18 (1.4)	16.4/18 (1.9)	V = 97.5, p = 0.034
	%	94.4 (7.9)	91.0 (10.4)	
MC score	raw	58.7/63 (2.9)	56.9/63 (4.8)	V = 271, p = 0.016
	%	93.1 (4.7)	90.4 (7.6)	

Wilcoxon rank sum tests with continuity correction revealed that the order of language administration did not significantly affect the L1 advantage (W = 73.5, p = .188). In fact, there were no significant differences between the two order groups in Catalan (W = 61, p = 0.064) or Spanish (W = 93, p = 0.625) scores. Descriptively, greater intersubject variability was observed in Spanish scores, especially when administered first (SD = 7.0) compared to when it was administered second (SD = 3.5), while the variability in Catalan remained similar when administered first (SD = 2.7) or second (SD = 3.1).

Discussion

These findings suggest that, even in proficient bilingual unimpaired speakers of minoritized languages, subtle L1 advantages persist in oral discourse. The fact that greater intersubject variability was observed in L2 than in L1 MC scores, particularly when L2 was administered first, also suggests less consistent performance in L2, which provides additional evidence for the L1 advantage. Notably, the observed L1 advantage emerged despite some participants considering themselves balanced bilinguals, suggesting that self-reported balance does not eliminate L1 advantages in discourse tasks.

This study demonstrated, in line with previous research (e.g., Criel et al., 2021; Kong, 2021), that oral discourse assessment needs to consider cultural, geographic, and linguistic factors. The findings also highlight the importance of assessing Catalan-speaking individuals in their first language, regardless of their high proficiency in L2, as L1 performance was more consistent. One limitation of the study is that the near-ceiling performance may have reduced sensitivity to detect larger or more variable differences among the two languages. Future

research should expand the sample and explore other aspects of bilingualism, such as language-mixing patterns in bilinguals with and without aphasia and their implications for clinical assessment practices.

References

Gao, G., Kong, A., & Lau, K. (2016). Production of main concepts by Mandarin-speakers with traumatic brain injury in China: A pilot study. *Frontiers in Psychology*.
Kong, A. (2009). The use of main concept analysis to measure discourse production in Catonese-speaking pearsons with aphasia: a preliminary report. *Journal of Communication Disorders*, 42, 442-464.
Kong, A., Ross, A., & Pettigrew, C. (2012). A main-concept analysis for aphasic discourse in Irish-English speakers: Adaptation and preliminary report. *Journal of Clinical Speech and Language Studies*, 19, 19–43.
Kong, A. P.H., Whiteside, J. & Bargmann, P. (2016). The Main Concept Analysis: Validation and sensitivity in differentiating discourse produced by unimpaired English speakers from individuals with aphasia and dementia of Alzheimer type. *Logopedics Phoniatrics Vocology*, 41(3), 129-141.
Matić, A., Pourquié, M., Norvik, M., Kuvač Kraljević, J., Gram Simonsen, H., Martínez-Ferreiro, S., Fyndanis, V., Munarriz-Ibarrola, A., Soroli, E., Kong, A., ... Arslan, S. (2026). Setting a research agenda for the assessment and treatment of aphasia in minority languages. *Cortex*, 198, 13–26.
Nicholas, L.E., & Brookshire, R.H. (1995). Presence, completeness, and accuracy of main concepts in the connected speech of non-brain-damaged adults and adults with aphasia. *Journal of Speech and Hearing Research*, *38(1)*, 145-156.

Acknowledgments

We thank all participants and Maria Macià for assistance with data collection.

Effects of psycholinguistic variables on repeated written naming attempts across aphasic deficits.

Salpéteur Laure | Hameau Solène

Introduction and aims

In repeated spoken picture naming, each successful naming attempt is hypothesized to strengthen the connections between an item's semantic and phonological representations (Nickels, 2002). If, as cognitive models suggest, analogous processes operate across spoken and written modalities, a similar reinforcement may occur in written naming—both between semantic and lemma representations and between lemma and orthographic (lexeme) representations. This would predict a reduction in latencies (i.e., the interval between picture offset and written naming onset) for correctly produced items across trials. At the orthographic (lexeme) level, a comparable facilitation effect on word production time (i.e., the interval between the start and the end of word writing) might also emerge.

Investigating the effects of psycholinguistic variables can provide critical insights into the underlying cognitive mechanisms of language production (e.g., Hameau et al., 2020; Nickels et al., 2022; Shallice, 1988). Exploring these variables in written picture naming—a modality less studied than spoken naming—could enhance our understanding of the processes involved in repeated naming tasks. Furthermore, examining how these variables interact with the cognitive deficits of people with aphasia (PWA) through a single-case study design may help clarify how access deficits or damage to orthographic representations manifest and how improvements to these processes likely operate.

This study aims to analyze the psycholinguistic variables influencing latencies and production times in repeated correct written naming. Additionally, it will explore how these variables interact with the distinct cognitive profiles of three PWA, each presenting different deficits, either at the interface between semantic and lexical representations, or at the level of orthographic representations.

This approach will enable us to a) Identify which word-related variables seem to determine the potential of these words to benefit from multiple naming attempts, b) further characterize the distinction between access deficits and orthographic lexicon impairments through the examination of psycholinguistic variables, and c) gain insight into the mechanisms underlying improvement in the affected processes.

Methods

Participants

This study investigates written word production deficits in three individuals with chronic aphasia (VW, SC, and JC), all of whom had normal spelling abilities before the lesion. Comprehensive background assessments were conducted to characterize their impairments. VW demonstrates a mild lexical retrieval deficit, with 18% errors in the written naming of 720 pictures, primarily at the lemma level (84% of errors), including semantic substitutions (e.g., writing *"ARTISAN"* for *"SCULPTOR"*) or omissions. SC, in contrast, exhibits a stronger post-lemma deficit, producing 44% errors in a 720-picture naming task. Of these errors, 73% occur at the post-lemma stage, with 25% of all errors (35% of the post-lemma errors) being

phonologically plausible misspellings (e.g., “SCULPTER" for "SCULPTOR"). JC presents a mixed profile, with 16% errors in naming 720 pictures, split between lemma-level errors (43%, including semantic errors or omissions) and phonologically plausible misspellings (43%).

Experimental design

This case-series study employs a repeated written picture-naming task to investigate the word-specific characteristics that modulate the effect of repeated written naming across a set of 100 to 240 picturable items.

At this stage, data collection is complete for the three PWA but analysis was not performed yet.

The three PWA were presented with pictures representing individual items and asked to write down the corresponding word in a single attempt, without feedback. Performances were recorded using a Wacom graphic tablet paired with the Eye and Pen V3.0 software, allowing precise time measurements (Alamargot et al., 2006).

Latencies and production times for each correct naming response were retrieved for each PWA at five time points, within a 24-week period, with varying intervals between measurements.

To isolate the effects of variables related to participants’ language deficits from the expected effects of these variables, matched control subjects will also perform the task at the same intervals as the participants.

The effects of psycholinguistic factors—affecting semantic, lexical, and post-lexical processing—on response latencies and production times, along with their interaction with measurement timing, will be examined in relation to each participant’s cognitive impairment profile.

Data analysis

(Generalized) Linear mixed effects models will be used to assess the interaction between psycholinguistic factors and changes in latency and production times in written picture naming across successive measurement points for each participant.

The following psycholinguistic variables will be examined :

- Semantic variables: number of semantic neighbors, number of semantic features.
- Lexical variables: frequency, number of orthographic neighbors, number of homophones.
- Orthographic (sublexical) variables: length, number of morphemes, orthographic consistency (phoneme-to-grapheme correspondence).

Results

Results will be presented at the conference in October 2026.

Discussion

We expect that variables affecting impaired processes in PWA will have a greater effect on measured latencies and production times compared to control subjects.

It is hypothesized that the degree and nature of a participant's impairment will modulate the impact of psycholinguistic variables. For instance, individuals with an impairment of the orthographic lexicon may exhibit greater sensitivity to orthographic variables, such as written word consistency or length, while those with a lexical label retrieval difficulty may show stronger effects of semantic or lexical variables, including the number of semantic features or lexical frequency.

Furthermore, we expect effects of each variable to manifest exclusively at their predicted processing levels: semantic and lexical variables should influence latency times, while sublexical/orthographic variables should affect production times.

Finally, improvement through repeated written naming trials should manifest as a gradual attenuation of critical variable effects according to each participant's deficit, modulated by both the number of repetitions and the time interval between them.

The results of this study will inform current theories relating to the cognitive underpinnings of aphasia recovery. Clinically, its outcomes are expected to help with the prediction of outcomes for written single-word treatment based on individual deficit profiles.

References

Hameau, S., Biedermann, B., Fieder, N., & Nickels, L. (2020). Investigation of the effects of semantic neighbours in aphasia: A facilitated naming study. *Aphasiology*, *34*(7), 840-864.
Nickels, L. (2002). Improving word finding: Practice makes (closer to) perfect?. Aphasiology, 16(10-11), 1047-1060.
Nickels, L., Lampe, L. F., Mason, C., & Hameau, S. (2022). Investigating the influence of semantic factors on word retrieval: Reservations, results and recommendations. *Cognitive Neuropsychology*, *39*(3-4), 113-154.
Shallice, T. (1988). From neuropsychology to mental structure. *Cambridge University Press*.
Alamargot, D., Chesnet, D., Dansac, C. (2006) Eye and Pen: A new device for studying reading during writing. *Behavior Research Methods 38*, 287–299.

Mild aphasia in patients with brain tumors and after stroke: objective and subjective insights

Djaina Satoer, Michelle van Steijn, Joost Schouten, Eelke Bos, Arnaud Vincent, Elke De Witte

Introduction

Currently, there is a lack of sensitive objective and subjective measurements to detect mild language disturbances (Mozeiko et al. 2020). A mild language disorder in the context of neurological diseases can have negative effects on daily communication as it is known in patients with moderate or severe aphasia (Cruice et al. 2006). Therefore, the digital Diagnostic Instrument for Mild Aphasia (DIMA) was developed. DIMA examines the most important linguistic levels: phonology, semantics and (morpho)-syntaxis at production and comprehension (Satoer et al., 2022). We aim to validate DIMA in a large group of patients with brain tumors (gliomas and meningiomas) and patients after stroke. In addition, we developed a short questionnaire for mild aphasia (QMA). The QMA consists of 24 questions for subtle language complaints at the levels of production, comprehension, reading and writing (responses via likert scale 1-5).

Methods

Patients with gliomas and meningiomas (N=82) were tested with DIMA as part of standard care preoperatively before (awake) surgery (mean age 47.37, sd 14.86) or after stroke (N=8) (mean age 54, sd 12.57). Accuracy scores were taken into account. All patients were compared to healthy participants (N=45) (mean age 45.51, sd 17.33), using an independent samples t-test. Data collection of the QMA in healthy population and patients is ongoing.

Results

Brain tumors (gliomas and meningioma)

DIMA has demonstrated sensitivity in a group of 82 patients with brain tumors (gliomas and meningiomas) compared with a group of 45 healthy Dutch-speaking individuals. Significant differences were observed in particular for the subtests repetition of compound words, sentence repetition, semantic odd-picture-out and sentence completion, with the patient group performing significantly worse than the healthy control group.

Stroke

Similarly, DIMA proved sensitive in a group of eight patients after stroke compared with the same group of 45 healthy Dutch-speaking individuals. Significant group differences were found for the subtests of sentence repetition, semantic odd-picture-out, and sentence completion, again showing poorer performance in the patient group relative to the healthy controls.

Table 1 presents the corresponding means, standard deviations, and results of the statistical *t*-tests ($P \leq 0.05$*, $P \leq 0.01$**).

Discussion and conclusion

DIMA appeared to be sensitive in a large patient group with brain tumors (gliomas and meningiomas) and a small group of patients after stroke. Language impairments were found in both patient groups at the linguistic levels phonology (sentence repetition), semantics (odd-picture-out) and syntax (sentence completion), but were more prominently present in patients with brain tumors as repetition compounds, syntactic sentence judgement were also impaired. This is in accordance with results from the pen-and-paper DIMA clinical validation and shows that mild aphasia can occur after lesions in widespread areas, i.e. both left and right hemisphere at intra- (glioma) and extra-axial (meningioma) level in addition to the cerebellum (Satoer et al. 2024, Vogrincic et al. 2025). The deviations at the main linguistic levels in both patient groups may also reflect earlier spontaneous speech analyses regarding the occurrence of repetitions, self-corrections and incomplete sentences in patients with glioma (Collee et al. 2026). We will further examine the long-term language postoperative outcome in patients with brain tumors (glioma and meningioma). In addition, the current sample size of patients after stroke remains limited and needs more volume as well as data from other neurological patient groups (.e.g primary progressive aphasia). Reaction times need to be analyzed in addition to data from the newly developed subjective measurement QMA to investigate the ecological validity of DIMA. These data will be analyzed in summer 2026. DIMA is a promising measurement for standard language evaluation in patients with suspected mild aphasia after neurological damage at various brain locations. It allows for more detailed baseline investigation before (awake) surgery and counseling about language recovery at different linguistic levels with indications for rehabilitation.

References

Collée, E., Vincent, A. J. P. E., Jiskoot, L. C., Bos, E. M., Schouten, J. W., Dirven, C. M. F., & Satoer, D. (2026). Spontaneous speech: A robust measurement before, during, and after awake brain surgery in patients with glioma. *British Journal of Neurosurgery, 40*(1), 85–93.

Cruice, M., Worrall, L., & Hickson, L. (2006). Perspectives of quality of life by people with aphasia and their family: Suggestions for successful living. *Topics in Stroke Rehabilitation, 13*(1), 14–24.

Mozeiko, J., & Pascariello, A. (2020). How are SLPs managing services for people with mild aphasia? *Journal of Communication Disorders, 84*.

Satoer, D., De Witte, E., Bulté, B., Bastiaanse, R., Smits, M., Vincent, A., Mariën, P., & Visch-Brink, E. (2022). Dutch Diagnostic Instrument for Mild Aphasia (DIMA): Standardization and a first clinical application in two brain tumour patients. *Clinical Linguistics & Phonetics, 36*(11), 929–953.

Satoer, D., Koudstaal, P. J., Visch-Brink, E., & van der Giessen, R. S. (2024). Cerebellar-induced aphasia after stroke: Evidence for the "linguistic cerebellum." *The Cerebellum, 23*(4), 1457–1465.

Vogrincic, B., Mooijman, S., Vincent, A., Bulté, B., De Witte, E., Visch-Brink, E., & Satoer, D. (2025). Precision in long-term language evaluation after awake brain tumor surgery. *Neuro-Oncology Practice*.

Table 1. Performance DIMA subtests patient groups (brain tumor, stroke) versus healthy controls. Yellow = brain tumor group deviant versus controls, Blue = both patient groups deviant versus controls ($P \leq 0.05$*, $P \leq 0.01$**).

DIMA Subtest	Group	Mean	SD
Repetition 3 syllabic words	Healthy Brain tumor Stroke	9.91 9.89 10	0.29 0.49 0
Repetition compound words	Healthy Brain tumor* Stroke	9.87 9.62 9.38	0.34 0.93 0.74
Repetition non-words	Healthy Brain tumor Stroke	9.20 9.25 9.50	1.24 1.46 0.76
Repetition sentences	Healthy Brain tumor** Stroke**	9.24 7.91 7.00	0.83 2.4 2.07
Semantic odd-picture-out	Healthy Brain tumor* Stroke*	9.22 8.62 5.63	1.09 1.80 3.74
Sentence completion	Healthy Brain tumor* Stroke*	9.13 8.28 6.63	1.16 1.93 2.33
Sentence judgment Phonology	Healthy Brain tumor Stroke	4.98 4,92 4.88	0.15 0.33 0.35
Sentence judgment Semantics	Healthy Brain tumor Stroke	4,84 4.81 4.63	0.37 0.44 0.52
Sentence judgment Syntax	Healthy Brain tumor** Stroke	4.60 4.11 3.88	0.65 1.10 1.55

Sentence completion for pre- and postoperative language testing in neurosurgical patients

Johanna Scholtz | Sabine Weiss | Horst M. Müller | Kirsten Heckmann | Matthias Simon

Introduction and Aims

Tumor-related aphasia can affect all language modalities and has a wide range of severity. In neurosurgical patients, diagnosing postoperative aphasia shortly after surgery is difficult due to the patients' impaired cognitive function and the potential effects of anesthesia [1]. Hence, a short but distinct method for evaluating different aspects of language pre- and postoperatively is needed. Additionally, it is crucial to incorporate highly functional language tasks, as standard tasks might not be able to detect mild symptoms in patients with tumor-related aphasia [2]. Sentence completion tasks have been described to detect mild aphasic deficits beyond the sensitivity of standard tests (e.g., object naming) [2]. This type of task involves reading comprehension, oral language production and morphosyntactic processing. Here, patients read sentences in which the final word is missing and then add a word that fits both the syntactic structure and the intended meaning. In this study, we examined pre- and postoperative language performance in neurosurgical patients. Specifically, we hypothesized differences from healthy controls already before surgery and aimed to follow postoperative changes over time. Accordingly, we conducted both sentence completion and object naming tasks, hypothesizing that sentence completion would be more sensitive in detecting subtle language deficits.

Methods

For sentence completion, 50 items from a previous study [3] were visually presented. The sentences contain different semantic relations (e.g., taxometric, instrumental) with varying word types to be added (e.g., nouns, verbs). Additionally, quantitative psycholinguistic criteria (e.g. cloze probability, syllable number, word frequency) were controlled for [3]. Object naming was performed with the Boston Naming Test [BNT; 4]. For both tasks, the patients' oral answers were recorded with a digital recorder and later analyzed manually using a sound editor. Here, we assessed the number of errors and the reaction time needed for word retrieval, which was measured from the onset of visual stimulus presentation to the start of the participants' verbal response. Prior to applying it to our patients, we piloted the visual sentence completion task in a group of 15 healthy adults (6 male; ø age 39.6 ± 13.5 years) to gain data on reaction time in healthy subjects. Later, 43 patients with intracranial lesions (24 male; ø age 55.1 ± 16.2 years; 32 right-handed) performed both the BNT and the SC task prior to surgery ("PRE"). Postoperative data (approximately 3 days after surgery) is available for 38 patients ("POST"). 24 patients performed a follow-up approximately 3 months after surgery ("FU"). The Berlin Aphasia Score [BAS; 5] was used to assess overall language performance before and 3 months after surgery. In a subgroup of patients (N = 23), we used the Trail-Making-Test [TMT, version B; 6] to assess cognitive flexibility prior to surgery.

Results

Both preoperatively (PRE) and 3 months after surgery (FU), 4 patients each were classified as having aphasia based on the BAS [6]. The total score of all patients for PRE (N = 43; ø 411.3 ± 18.5) and FU (N = 24; ø 405.7 ± 29.4) did not differ significantly (t = .929, p = .357). For the healthy adults, we calculated a mean sentence completion time of 2268.6 (± 1139.9) ms.

Preoperatively, our patients had a mean sentence completion time of 3725 ± 2832 ms. Patients performed significantly slower than the healthy group (β = −0.179, SE = 0.047, t = −3.84, p < .001). Age did not significantly predict the outcome (p = .431). There was a significant correlation between the preoperative reaction times for sentence completion and the preoperative TMT (: ø 83.9 ± 31.3 ms; ρ = .423, p = .044), but not for object naming and the TMT (ρ = .358, p = .093). Postoperatively, patients showed significantly slower reaction times in the sentence completion task (ø 4290 ± 2774 ms) compared to the preoperative session (β = -0.075, SE = 0.022, t = -3.45, p = .004). In addition, the number of errors increased significantly from PRE- to POST-test (PRE ø 5.0 ± 4.9 errors; POST ø 8.4 ± 6.5*;* β = 0.770, SE = 0.170, z = 4.529, p < . 001). For object naming, neither reaction times (PRE ø 1290 ± 988 ms; POST ø 1266 ± 999 ms) nor error numbers (PRE ø 11.6 ± 6.4 errors; POST ø 12.6± 11.2) changed between PRE and POST. At the 3-month follow-up (FU), reaction times in the sentence completion task remained stable (ø 4327.3 ± 2837.9 ms). In contrast, the number of errors (ø 4.8 ± 6.7 errors) decreased significantly compared to the POST session (β = 0.075, SE = 0.516, z = 4.160, p < .001). For object naming, reaction times (ø 1350 ± 1006 ms) and error numbers (ø 12.9 ± 11.5 errors) remained unchanged at follow-up.

Discussion

Overall, the assessment of reaction time for the BNT and sentence completion was time-efficient for patients and feasible already a few days after surgery. The comparison with healthy controls suggests that patients already performed significantly slower in sentence completion prior to surgery, indicating the presence of preoperative aphasia. Notably, aphasia was classified in only 4 patients according to the BAS [5], highlighting the sensitivity of the sentence completion task in detecting subtle preoperative language deficits. The correlation with the TMT [6] suggests that sentence completion might also reflect nonverbal cognitive functions, which may be beneficial for assessing more general cognitive performance. Postoperatively, patients performed significantly slower and made significantly more errors in the sentence completion task. Although reaction times for sentence completion did not change at follow-up, patients made significantly fewer errors compared to the immediate postoperative assessment. In contrast, neither reaction times nor error numbers in object naming showed significant changes across the PRE, POST, and FU tests. This suggests that sentence completion is more sensitive to subtle language impairments. Overall, the findings support its use as a sensitive tool for detecting preoperative and postoperative deficits and for monitoring language recovery over time.

References

[1] Sun, J., Du, X., & Chen, Y. (2024). Current progress on postoperative cognitive dysfunction: An update. *Journal of Integrative Neuroscience*, 23(12), 24.
[2] Satoer, D., De Witte, E., Bulté, B., Bastiaanse, R., Smits, M., Vincent, A., Mariën, P., & Visch Brink, E. (2022). Dutch Diagnostic Instrument for Mild Aphasia (DIMA): Standardisation and a first clinical application in two brain tumour patients. *Clinical Linguistics & Phonetics*, 36(11), 929–953.
[3] Scholtz, J., Weiss, S., Redecker, C., & Müller, H. M. (2023). Sentence completion in progressive supranuclear palsy following transcranial direct current stimulation. *NPJ Parkinson's Disease*, 9(1), 162.
[4] Kaplan, E., Goodglass, H. & Weintraub, S. (2001). *Boston Naming Test*. 2nd Ed. Austin: Pro-Ed.

[5] Tuncer, M. S., Salvati, L. F., Grittner, U., Hardt, J., Schilling, R., Bährend, I., Silva, L. L., Fekonja, L. S., Faust, K., Vajkoczy, P., Rosenstock, T., & Picht, T. (2021). Towards a tractography-based risk stratification model for language area associated gliomas. *NeuroImage: Clinical*, 29, Article 102541.
[6] Delis, D.C., Kaplan, E. & Kramer, J.H. (2021). *Delis-Kaplan Executive Function System.* (German Version). München: Pearson.

Overall Language Impairment and Cross-Language Effects Shape Reading in Bilinguals with Aphasia

Michael Scimeca | Isabel Yu | Amelia Andre | Maria Varkanitsa | Swathi Kiran

Introduction and aims

Reading impairment, or alexia, is common after left-hemisphere stroke in monolinguals with aphasia (Brookshire et al., 2014), yet the mechanisms underlying alexia in bilinguals with aphasia (BWA) remain poorly understood. Current theoretical accounts suggest that reading depends on modality-independent semantic and phonological systems that also support spoken language, such that damage to these systems should affect tasks like oral reading and oral repetition in similar ways (Patterson & Lambon Ralph, 1999; Lambon Ralph & Patterson, 2007; Woollams, 2014). In bilingual readers, however, reading is also shaped by cross-language activation that creates competition between shared phonological and orthographic representations (Dijkstra & van Heuven, 1998, 2002; Dijkstra et al., 2019). The present study examined the extent to which alexia in BWA reflects both central language impairment and cross-language effects. The study had three aims: (1) to quantify reading performance in BWA relative to bilingual neurological controls (BNC); (2) to examine whether non-orthographic semantic and phonological abilities predicted reading outcomes; and (3) to evaluate whether cross-language phonological and orthographic effects modulated single-word reading accuracy.

Methods

Participants were 14 Spanish-English BWA following left-hemisphere stroke and 14 age- and education-matched BNC. The BWA had a mean age of 51.2 years (SD=14.9, range=27–75), had 16.4 mean years of education, (SD=3.6, range=12–22), and were 112.9 mean months post-onset (SD=65.3, range=41–237). Most reported Spanish as their first language and mean age of second-language acquisition was 11.3 years (SD=7.7, range=1–21). Western Aphasia Battery–Revised Aphasia Quotient scores indicated moderate aphasia severity in English (M=79.9, SD=20.0, range=42.9–98.0) and Spanish (M=74.2, SD=17.9, range=40.1–95.4). BNC had a mean age of 38.2 years (SD=16.1, range=24–72) and 17.6 years of education (SD=3.37, range=12–22).

Participants completed English and Spanish language assessments. Reading outcomes included oral reading of real words and nonwords and silent reading comprehension at the word and sentence level. BWA also completed non-orthographic semantic tasks, consisting of auditory word- and sentence-picture matching, and non-orthographic phonological tasks, including auditory lexical decision, word/nonword repetition, sound/syllable deletion, and rhyme judgment. Task accuracy was scored at the item level, and language-specific semantic and phonological composite scores were created by averaging standardized performance across semantic or phonological subtests within participant.

For cross-language effects, phonological and orthographic similarity were computed for translation equivalents across English and Spanish. Phonological forms were transcribed into IPA and aligned using a weighted Levenshtein procedure that allowed partial similarity

among non-identical phonemes based on shared phonological features. Orthographic similarity was computed using normalized Levenshtein distance over character strings. Similarity values were scaled such that higher values indicated greater overlap.

Generalized linear mixed-effects models were fit to trial-level accuracy to estimate the probability of a correct response across reading tasks. Aim 1 tested group differences in oral reading and reading comprehension with fixed effects of group (BWA vs. BNC), test language (Eng vs. Spn), domain (oral reading vs. reading comprehension), self-reported premorbid reading ability, and relevant two- and three-way interactions. Subsequent models for Aim 2 tested whether semantic and phonological composites predicted reading performance in BWA only across three oral reading outcomes, i) real words, ii) nonwords, and iii) all words and three reading comprehension outcomes, i) words, ii) sentences, and iii) all items. For Aim 3, cross-language single-word reading accuracy was predicted from phonological similarity, orthographic similarity, their interaction, test language, and reading domain, adjusting for cross-language phonological and orthographic neighborhood density. Random intercepts for participants and items were included across models.

Results

Across languages and tasks in Aim 1, BWA showed significantly lower reading accuracy than BNC ($b=-2.19$, $z=-3.32$, $p<.001$). Oral reading was disproportionately impaired relative to reading comprehension, demonstrated by a significant Group × Domain interaction ($b=-1.10$, $z=-15.36$, $p<.001$). There was no reliable overall English-Spanish difference, nor a significant Group × Test Language × Domain interaction. Premorbid reading proficiency also positively predicted accuracy ($b=0.64$, $z=2.88$, $p<.01$), linking stronger pre-stroke reading ability to better post-stroke performance.

Non-orthographic semantic and phonological predictors showed task-specific effects in Aim 2. For oral reading, stronger phonological abilities predicted higher accuracy for real words ($b=2.76$, $z=7.31$, $p<.001$), nonwords ($b=1.47$, $z=6.07$, $p<.001$), and overall oral reading ($b=1.93$, $z=10.15$, $p<.001$). Semantic abilities additionally predicted real-word oral reading ($b=0.95$, $z=2.03$, $p=.04$), but not nonword or overall oral reading. For reading comprehension, semantic abilities predicted word-level comprehension ($b=1.43$, $z= 5.07$, $p<.001$). Both semantic and phonological abilities predicted sentence comprehension (semantic: $b=0.60$, $z=3.34$, $p<.001$; phonology: $b=0.91$, $z=4.08$, $p<.001$) and overall comprehension accuracy (semantic: $b=0.80$, $z=6.02$, $p<.001$; phonology: $b=0.63$, $z=4.33$, $p<.001$).

Aim 3 results showed that cross-language structure also influenced single-word reading accuracy. As shown in Figure 1, greater phonological similarity between translation equivalents predicted higher accuracy ($b=1.38$, $z=2.08$, $p=.03$), whereas greater orthographic similarity predicted lower accuracy ($b=-1.10$, $z=-2.54$, $p=.01$). Cross-language orthographic neighborhood density also facilitated accuracy ($b=0.08$, $z=2.51$, $p=.01$). The interaction between phonological and orthographic similarity was not significant.

Discussion

Across oral reading and reading comprehension, BWA performed significantly worse than BNC which suggests that alexia is a pervasive consequence of left-hemisphere stroke in this population also (Brookshire et al., 2014). Additionally, the relationship between premorbid reading proficiency and post-stroke accuracy highlights bilingual literacy experience as an

underexamined source of variability in clinical reading outcomes. Non-orthographic semantic and phonological abilities further predicted reading across domains, providing initial support for extending modality-independent accounts of alexia to BWA. Under this account, reading impairment reflects disruption to central semantic and phonological systems rather than damage to orthographic representations alone (Patterson & Lambon Ralph, 1999; Lambon Ralph & Patterson, 2007). Beyond these central language predictors, reading accuracy was also shaped by cross-language overlap. Phonological similarity facilitated accuracy, whereas orthographic similarity interfered with accuracy, which suggests that bilingual word recognition after stroke is marked by both cross-language activation and competition—typical of BNC. The positive effect of orthographic neighborhood density, despite negative effects of direct orthographic similarity, further suggests that cross-language structure may operate differently across levels of bilingual lexical organization. These findings indicate that reading impairment in BWA reflects aphasia-related semantic and phonological impairment operating within an interactive bilingual lexical system (Dijkstra & van Heuven, 2002).

References

Brookshire, C. E., Wilson, J. P., Nadeau, S. E., Gonzalez Rothi, L. J., & Kendall, D. L. (2014). Frequency, nature, and predictors of alexia in a convenience sample of individuals with chronic aphasia. *Aphasiology*, *28*(12), 1464–1480. https://doi.org/10.1080/02687038.2014.945389

Dijkstra, T., & van Heuven, W. J. B. (1998). The BIA model and bilingual word recognition. In *Localist connectionist approaches to human cognition* (pp. 189–225). Lawrence Erlbaum Associates Publishers.

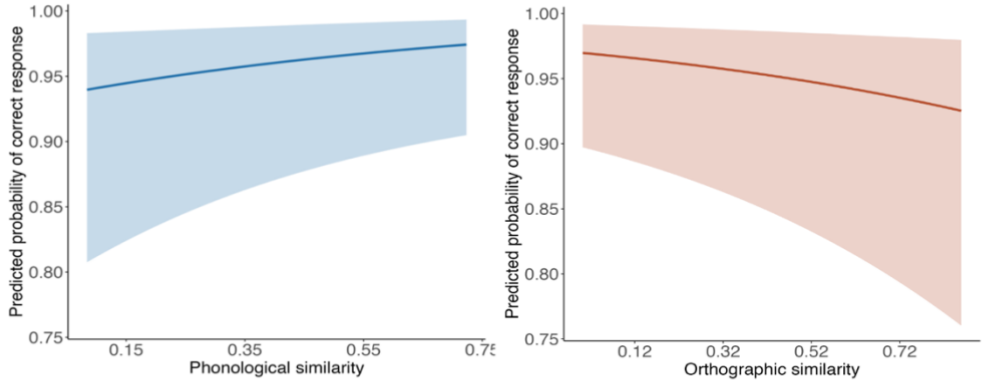
Dijkstra, T., & Heuven, W. J. B. van. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*(3), 175–197. https://doi.org/10.1017/S1366728902003012

Dijkstra, T., Wahl, A., Buytenhuijs, F., Halem, N. V., Al-Jibouri, Z., Korte, M. D., & Rekké, S. (2019). Multilink: A computational model for bilingual word recognition and word translation. *Bilingualism: Language and Cognition*, *22*(4), 657–679. https://doi.org/10.1017/S1366728918000287

Lambon Ralph, M. A., & Patterson, K. (2007). Acquired disorders of reading. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 413–430). Blackwell.

Patterson, K., & Lambon Ralph, M. A. (1999). Selective disorders of reading? *Current Opinion in Neurobiology*, *9*(2), 235–239. https://doi.org/10.1016/S0959-4388(99)80033-6

Figure 1. Effect of Phonological and Orthographic Similarity on Word Reading in BWA.



Caption. Predicted probability of a correct response as a function of phonological similarity (left) and orthographic similarity (right) for single-word reading. Each panel shows the marginal effect of one similarity metric derived from the mixed-effects model while the other similarity metric and all covariates were held at their centered means. Shaded regions indicate 95% confidence intervals around model-predicted probabilities.

Understanding language change in cognitive aging from a usage-based perspective: Implications for Aphasia

Ioana Sederias | Sophie Froud | Jessica Jiang | Lucy B. Core | Benjamin A. Levett | Jason D. Warren | Chris Hardy | Vitor Zimmerer

Introduction and aims

Accurate assessment of language production in aphasia requires an understanding of language in healthy populations across the lifespan, and particularly the effects of different demographic variables. However, existing normative language frameworks often fail to account for demographic factors, leading to highly heterogeneous samples in which substantial variance in language outcomes remains unexplained (Folia et al., 2022). Moreover, normative studies primarily focus on isolated aspects of grammatical (e.g., syntactical complexity) or lexical abilities (e.g., naming), rather than providing a comprehensive account of patterns of linguistic changes (Benitez-Buraco and Ivanova, 2023). This approach is also reflected in aphasia research and clinical practice, underpinning characterisations, assessments and treatments of the disorder. For example, definitions of some aphasia subtypes rely on a strict division between lexis and grammar (i.e., agrammatic vs anomic (word-finding) aphasia), despite substantial variability in the nature of impairments observed within each subtype (Marshall, 2010). Overall, this broadly aligns with generativist assumptions which separate lexical and grammatical components of language.

Alternative approaches, in particular usage-based theories, do not make a categorical distinction between words and phrases and place high emphasis on frequency and familiarity effects, extending them to phrase level (Ambridge and Lieven, 2011). This approach offers a novel way to detect changes in language at the sentence level. For instance, frequently co-occurring word sequences may be processed more easily. Applications of usage-based approaches highlight a pattern of preservation of familiar phrases in aphasia language production (Zimmerer et al., 2018). Overreliance on frequent structures, while often not noticeable in everyday communication, could indicate underlying cognitive changes. Further consideration of usage-based measures could therefore aid the detection and tracking of such changes.

Clinical applications of usage-based frameworks are limited by several gaps. Firstly, there are currently no normative benchmarks in healthy populations of such patterns. Moreover, it remains unclear which usage-based phenomena are indicative of cognitive decline, how they relate to formal grammatical measures (e.g., lexical diversity, syntactic complexity) and lastly, how they might be affected by demographics. This study aims to address these gaps by conducting a comprehensive exploration of language production in a healthy ageing population, specifically exploring demographic effects and implementing usage-based measures.

Methods

A secondary analysis was conducted on Cookie Theft picture-description transcripts from 43 healthy older adults (mean age = 66.56, SD = 8.38, 26 F, 17 M, mean education = 15.62, SD = 2.39). Language variables were extracted partly manually and partly using automated tools, including the Frequency In Language Analysis Tool (FLAT, Zimmerer et al., 2018), the Tool for the Automatic Analysis of Syntactic Sophistication and Complexity & Tool for the

Automatic Analysis of Lexical Diversity (TAASSC, & TAALD, Kyle, 2016) and Computerised Propositional Idea Density Rater (CPIDR, Brown et al. 2008).

In total, 15 language variables characterising four different levels of language processing, namely lexical, semantic, fluency and syntactic were considered. Formal measures of syntactic complexity (e.g., clauses per sentence) and lexical diversity (e.g., moving type token ratio) were complemented with usage-based metrics capturing predictability and conventionality aspects of word combinations (e.g., collocation strength). Language outcomes were analysed in relation to three demographic factors: age, sex, and years of education, as well as cognitive ability, measured by the Wechsler Abbreviated Scale of Intelligence Matrix Reasoning and Vocabulary subtest (Wechsler, 2011), and Digit Span Test (Wechsler, 1987).

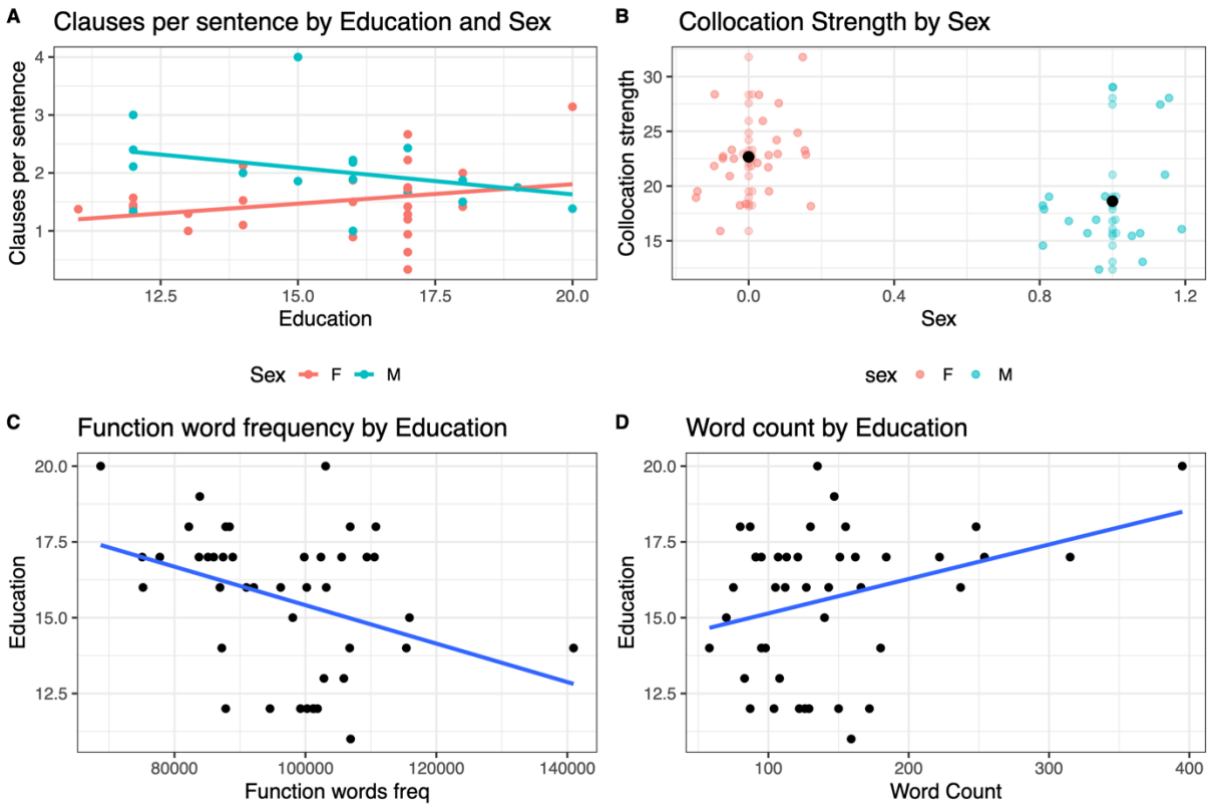
Results

A series of linear regression models revealed that sex and education were predictors of several linguistic production measures across the lexical and syntactic domains (Figure 1). Males were more likely to produce longer sentences ($\beta = 1.56$, $p = .018$) and less familiar word combinations ($\beta = -0.23$, $p = .013$). More years of formal education were associated with a small but significant increase in the number of words produced ($\beta = .013$, $p = .028$) and a decrease in function word frequency ($\beta = -.000063$, $p = .023$), indicating greater grammatical complexity. Sex effects on sentence length were moderated by education, albeit with marginal significance ($\beta = -.16$, $t = -1.95$, $p = .058$), with males showing higher scores than females at lower levels of education, a difference which decreased as education years increased. Addition of usage-based measures at the syntactic level significantly improved the model fit for syntactic ability differences in sex ($\chi^2 = 9.72$, $p = .021$). We found no significant effects of age and cognitive abilities.

Discussion

Demographic factors impacted language performance. Based on our data, education and sex are more important for linguistic profiling than age, suggesting that the former variables need greater consideration when modelling language and cognitive decline in general in neurological populations. Usage-based variables were sensitive to demographic differences, and offered additional explanatory power alongside formal measures. Our findings suggest that these demographic and linguistic factors, which are often not prioritized in models of language and cognitive decline, should be given greater consideration. By establishing normative patterns of change as well as the relationships between language measures and demographics, this work provides a foundation for more sensitive and comprehensive models of language variation, with direct implications for clinical assessment.

Figure 1. *Significant relationships between linguistic measures and demographic variables: (A) Sentence length as a function of education and sex (interaction), (B) differences in collocation strength between males and females, (C) function word frequency as a function of education, and (D) word count as a function of education.*



References

Ambridge, B., & Lieven, E. (2015). A constructivist account of child language acquisition. *The handbook of language emergence*, 478-510.

Benítez-Burraco, A., & Ivanova, O. (2024). Language in healthy and pathological ageing: Methodological milestones and challenges. *International Journal of Language & Communication Disorders*, *59*(1), 4-12.

Brown, C., Snodgrass, T., Kemper, S. J., Herman, R., & Covington, M. A. (2008). Automatic measurement of propositional idea density from part-of-speech tagging. *Behavior Research Methods*, 40(2), 540-545

Folia, V., Liampas, I., Ntanasi, E., Yannakoulia, M., Sakka, P., Hadjigeorgiou, G., ... & Kosmidis, M. H. (2022). Longitudinal trajectories and normative language standards in older adults with normal cognitive status. *Neuropsychology*, *36*(7), 626.

Kyle, K. (2016). Measuring syntactic development in L2 writing: Fine grained indices of syntactic complexity and usage-based indices of syntactic sophistication (Doctoral Dissertation).

Marshall, J. (2010). Classification of aphasia: Are there benefits for practice?. *Aphasiology*, *24*(3), 408-412.

Wechsler, D. (1987). Wechsler memory scale-revised. *Psychological Corporation*.

Wechsler, D. (2011). *Wechsler Abbreviated Scale of Intelligence--Second Edition (WASI-II)*. APA PsycTests.

Zimmerer, V. C., Newman, L., Thomson, R., Coleman, M., & Varley, R. A. (2018). Automated analysis of language production in aphasia and right-hemisphere damage: frequency and collocation strength. *Aphasiology*, *32*(11), 1267-1283

Morphological Relations and Word Networks in People with Post-Stroke Aphasia

Jo-Anne H. van der Sluijs | Roel Jonkers | Elisabeth Beyersmann | Judith Feiken | Adrià Rofes

Introduction and aims

People with post-stroke aphasia can have difficulties with morphology. Inflection, derivation, and compounding can be differentially affected in production, as seen in spontaneous speech, reading and repetition tasks (Badecker & Caramazza, 2017). Newer tasks are now proposed to assess morphology, such as free association (Rofes et al., 2026), verbal fluency and relatedness judgement tasks. In verbal fluency, people say as many words as possible in one minute. In relatedness judgement tasks, people rate the relatedness of two words at a time. These tasks allow assessment of morphological processing in a more explicit manner than spontaneous speech, reading or repetition tasks, while obtaining information regarding the use of morphology that goes beyond accuracy. However, so far, these tasks were mainly used to study semantic word relations (Siew et al., 2019).

One of the ways in which this in-depth analysis can be accomplished is through studying word networks. Word networks reflect the structure of the mental lexicon (Siew et al., 2019). Semantic networks based on verbal fluency are less efficient and more partitioned for people with fluent and nonfluent aphasia than for healthy controls (Pham et al., 2025). Nevertheless, networks including more than one type of word relation (i.e., semantics, phonology) better predict picture naming performance by individuals with aphasia than single-layer networks (Castro & Stella, 2019). For brain tumor patients after surgery, modularity (i.e., how well a network divides into smaller groups of nodes) and clustering coefficient (i.e., whether neighbors of a node are also connected to each other) in semantic networks based on relatedness judgement correlate with measures from category and/or letter fluency (Gaudet, 2022).

In this study, we investigate morphological impairments in people after stroke, applying in-depth analyses novel to the domain of morphology. More specifically, we study the production of morphologically related words in the setting of a fluency task, and relate participants' performance in this task to their individual morphological networks constructed from relatedness ratings.

Methods

At the time of abstract submission we recruited four Dutch individuals with post-stroke aphasia in the chronic phase as well as multiple age-matched controls ($N = 5$). By the time of the conference, we aim to have five people after stroke. An inclusion criterion for individuals with aphasia is that language comprehension should be sufficient to understand the task instructions.

Two main tasks are employed: a morphological fluency task and a relatedness judgement task. In the morphological fluency task, participants are presented with 10 Dutch words (five nouns and five verbs), one at the time, and they are asked to produce as many morphologically related words as possible in one minute (see Figure 1). In the relatedness judgement task, participants rate how related they think word pairs are on a scale from 0 (not related) to 100 (strongly related). The instructions do not specify any particular type of

information (e.g., semantics) to use in rating the word pairs. Word pairs are either semantically-related, orthographically-related, morphologically-related, or unrelated. All possible word pair combinations are presented among a set of words with a particular relation (e.g., semantically-related). For each word relation type, individual, relatedness-based networks are constructed, with words as nodes and word pair ratings as weighted edges (i.e., connections).

Using single-case statistics, we compare morphological fluency performance of the individuals with aphasia to the controls. We explore differences in the number of items produced, the use of morphological operations (i.e., inflection, derivation, compounding), and whether noun versus verb cues elicit different responses. For the relatedness judgement task, ratings for each word pair relation (i.e., semantically-related, orthographically-related, morphologically-related, unrelated) are compared to each other and between groups. Network metrics (e.g., modularity, clustering coefficient) are compared between groups for each of the four word relations. Metrics of each individual's morphological network are compared with their number of responses in morphological fluency.

Expected results

Data collection is still ongoing. We hypothesize that individuals with aphasia will produce fewer morphologically related words than controls in morphological fluency. Depending on the specific language impairments, fewer inflected, derived or compounded word forms may be produced. If inflection is impaired, fewer responses are expected for verb than noun cues. In relatedness judgement, word pairs could be rated as less related by people with aphasia than controls, indicating possible language deficits for that word relation (i.e., semantics, orthography, morphology). Comparisons of network metrics might reveal less efficient underlying networks for individuals with aphasia (e.g., lower clustering coefficient and higher modularity; Pham et al., 2025). If relatedness judgement networks reflect the organization of words in the mind for both comprehension and production, less efficient morphological networks could be related to producing fewer morphologically-related responses in the fluency task.

Discussion

Morphological fluency could detect mild morphological deficits in people with aphasia by testing the production of morphologically complex words in the same word family. Observing different patterns in the use of morphological operations for individuals with distinct (degrees of) language impairments might provide further insight into the representation of morphology in the mental lexicon. Constructing word networks from relatedness judgement allows for the quantification of the structure with which words are stored in relation to each other. Relating these networks to fluency performance provides an in-depth, structural perspective on how (morphological) word relations are stored and/or accessed, beyond current testing paradigms.

References

Badecker, W., & Caramazza, A. (2017). Morphology and Aphasia. In A. Spencer & A. M. Zwicky (Eds), *The Handbook of Morphology* (1st edn, pp. 390–405). Wiley. https://doi.org/10.1002/9781405166348.ch20

Castro, N., & Stella, M. (2019). The multiplex structure of the mental lexicon influences picture naming in people with aphasia. *Journal of Complex Networks*, *7*(6), 913–931. https://doi.org/10.1093/comnet/cnz012
Gaudet, L. A. (2022). *Semantic networks following awake brain surgery: Perspectives from an associative judgement task and category fluency* [Master's thesis, University of Groningen]. https://doi.org/10.13140/RG.2.2.21652.07049
Pham, C. T., Castro, N., & Lee, J. (2025). Lexical retrieval in fluent and nonfluent aphasia: A network analysis of verbal fluency data. *Frontiers in Human Neuroscience*, *19*, Article 1710907. https://doi.org/10.3389/fnhum.2025.1710907
Rofes, A., Beyersmann, E., Rossetto, A., & Castro, N. (2026). Limited evidence of morphology in semantic and phonological free association tasks. *Language and Cognition*, *18*, Article e15. https://doi.org/10.1017/langcog.2026.10065
Siew, C. S. Q., Wulff, D. U., Beckage, N. M., & Kenett, Y. N. (2019). Cognitive network science: A review of research on cognition through the lens of network representations, processes, and dynamics. *Complexity*, *2019*(1), Article 2108423. https://doi.org/10.1155/2019/2108423

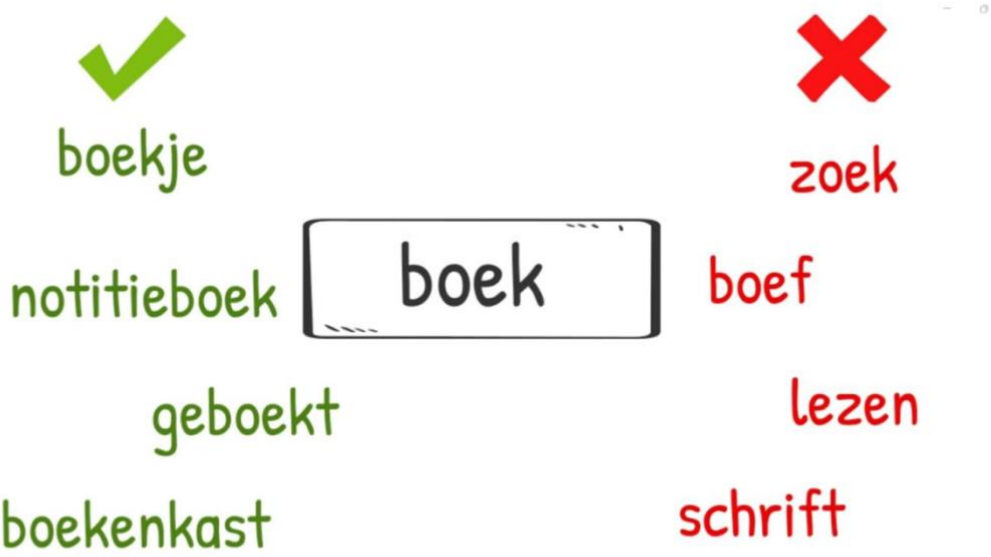


Figure 1. Examples of morphologically-related words on the left for the word *boek* (book), and examples of non-morphologically-related words (i.e., semantically-related or orthographically-related) on the right, as presented and explained to participants. Morphologically-related words, top to bottom: diminutive form of book, notebook, booked, bookcase. Non-morphologically-related words, top to bottom: search, crook, read, exercise book.

Syntactic abilities and short-term memory in the behavioral variant of frontotemporal dementia

Christiana Spiliou | Vasiliki Koukoulioti | Panagiotis Ioannidis | Michaela Nerantzini

Introduction and aims

The behavioral variant of the frontotemporal dementia (hereafter bvFTD) is characterized by executive dysfunction and behavior and social cognition deficits (Boeve et al., 2022). Whereas lexicosemantic deficits have been often found in bvFTD (Geraudie et al., 2021), findings on sentence processing are equivocal. For example, subtle decline in sentence complexity at the initial stages as well as fewer grammatical sentences as the disease progresses, have been documented in connected speech (Ash et al., 2019). Charles et al. (2014) found impaired performance in the comprehension of long and complex sentences and correlation between sentence comprehension and working memory. However, other studies suggest intact sentence comprehension and repetition even after 8 years of disease onset (Hardy et al., 2016). Interestingly, sentence production in bvFTD has been studied mainly by means of connected speech, in which language abilities are interwoven with general communication abilities. Therefore, it remains unclear whether impaired performance reflects a primary language disorder or can be attributed to the behavioral deficit. Moreover, the relation of sentence production with short-term memory in bvFTD has not been examined yet, to the best of our knowledge.

The aim of the present study is to investigate sentence production in patients with bvFTD by means of a sentence repetition task and explore the relation of sentence production to short-term memory. Specifically, the study addresses the following questions: 1) is sentence production impaired in bvFTD and 2) is there a relation between the performance in sentence repetition and verbal short-term memory?

Methods

Seven patients with the behavioral variant of frontotemporal dementia (bvFTD) participated (4 females) (mean age = 71.9 years, SD = 8.4, mean years of education = 11.4, SD = 4.1). Participants were recruited from the Outpatient Clinic for Dementia and Related Disorders, Second Department of Neurology, AHEPA University Hospital, where they underwent clinical and neuropsychological assessment, as well as brain MRI.

Syntactic repetition abilities were assessed using the LITMUS test (Chondrogianni et al., 2013), which includes a total of 32 sentences with a relatively constant length, each corresponding to one of the following target structure (Subject Verb Object, Complement Clauses, Clitic Left Dislocation structures, Negations, Relative Clauses, Adverbials, Wh-questions, and Coordination). Apart from verbatim repetitions, we also calculated: 1) the number of grammatical responses, independently of successful repetition, in order to evaluate general sentence production abilities, and 2) the number of successful productions of the target structure, independently of overall grammaticality, in order to identify challenging structures. Phonological short-term memory was assessed using a pseudoword repetition task (Talli et al., 2023), consisting of 24 pseudowords whose length increases progressively (3, 4, 5, and 6 syllables), with syllabic structure alternating between CV and CVC.

Results

Table 1 presents the mean percentage of verbatim repetitions, grammatical responses and responses with successful target structure production. Patients had low performance in producing verbatim repetitions. Indeed, six out of seven patients scored two standard deviations below the healthy controls' mean of Kaltsa et al. (2024). Interestingly, structural complexity did not affect performance, as complex sentences, such as wh-questions were less impaired than simple ones (e.g. subject-verb-object), which will be discussed later. There was a marginally significant difference among the examined structures *(Friedman's* $x^2(7)=14.3$, $p=0.047$), but no pair-wise comparison was significant. The majority of the patients' responses were grammatical (86%); however, only two patients performed within the normal range, which hints at a syntactic disorder. The difference among the structures concerning grammatical responses failed to reach significance *(Friedman's* $x^2(7)=13.8$, $p=0.055$). With regard to the successful production of the target structure, the most impaired structure was the clitic left dislocation, followed by complement clauses and negation. Although the difference among the structures was significant (*Friedman's* $x^2(7)=17.2$, $p=0.02$), no pair-wise comparison reached significance. Finally, phonological short-term memory did not correlate with any of the scores of the repetition task (*Spearman's* $\rho=0.17$, $p=.7$ for verbatim repetition, *Spearman's* $\rho=0.25$, $p=.6$ for grammatical responses, *Spearman's* $\rho=0.22$, $p=.6$ for successful target structure responses).

Error analysis revealed various error types, often multiple errors within one response. Word substitutions constituted 49% of the errors, followed by omissions (34%). Patients produced also additions of words or constituents, changed the word order or failed to respond, 9%, 6% and 2% of the errors, respectively. Substitutions primarily affected content words rather than function words (71% and 29% of the substitutions, respectively). Among content-word errors, semantic paraphasias were the most frequent (35%), followed by non-classifiable (24%) and phonological paraphasias (17%), light verb substitutions (13%), and perseverations (11%). Substitutions of functional words involved morphological features such as number, person, and tense.

Omissions typically involved core constituents, resulting in semantic ambiguity and/or ungrammaticality. For instance, the participants omitted components of multiword units or obligatory objects (e.g. "*The drivers left at the next stop", instead "The drivers left the passengers at the next stop"), violating argument structure. In many responses multiple omissions occurred.

Discussion

Taken together, the findings suggest a core impairment in sentence production. The failure to repeat verbatim complex sentences might be related to memory capacity restrictions or to impairment in lexical access. Nevertheless, most patients were impaired in producing grammatical sentences, violating core aspects of sentence production, such as verb subcategorization and thematic role assignment, which suggests a pure syntactic deficit. Crucially, comparison with a group of healthy participants (Kaltsa et al., 2024) revealed individual variability with some patients performing within the normal range and others being impaired. Concerning the second research question, although the failure to repeat verbatim sentences might imply the contribution of a verbal short-term memory impairment, the performance in the repetition task was not found to correlate with the short-term memory task. Finally, low performance in Subject Verb Object Structure seems to be related to

specific items, which deserves further investigation. Concluding, despite the limited number of participants and the small number of items, the present study provides evidence for a syntactic impairment beyond short-term memory and lexical deficits in bvFTD

Selected References

Ash, S., Nevler, N., Philips, J., Irwin, D. J., McMilan, C. T., Rascovsky, K., & Grossman, M. (2019). A longitudinal study of speech production in primary progressive aphasia and behavioral variant frontotemporal dementia. *Brain and Language*, *194*, 46 -57. https://doi.org/10.1016/j.bandl.2019.04.006

Chondrogianni, V., Andreou, M., Nerantzini, M., Varlokosta, S., & Tsimpli, I. M. (2013). The Greek Sentence Repetition Task. COST Action IS0804.

Geraudie, A., Battista, P., Garcia, A. M., Allen, I. E., Miller Z. A., Gorno - Tempini, M. L., Montembeault, M. (2021). Speech and language impairments in behavioral variant frontotemporal dementia: A systematicreview. *Neuroscience & Biobehavioral Reviews*, *131*, 1076-1095. https://doi.org/10.1016/j.neubiorev.2021.10.015

Kaltsa, M., Tsolaki, A., Lazarou, I., Mittas, I., Papageorgiou, M., Papadopoulou, D., Tsimpli, I. M., & Tsolaki, M. (2024). Language Markers of Dementia and Their Role in Early Diagnosis of Alzheimer's Disease: Exploring Grammatical and Syntactic Competence via Sentence Repetition. *Journal of Alzheimer's Disease Reports*, *8*(1), 1115-1132. https://doi.org/10.3233/ADR-230204

Talli, I., Kotsoni, P., Stavrakaki, S., & Sprenger-Charolles, L. (2023). Assessing phonological short-term memory in Greek: Reliability and validity of a non-word repetition test. *Frontiers in Psychology*, *13*. https://doi.org/10.3389/fpsyg.2022.904268

Table 1

Structure	Verbatim Repetition	Grammatically Correct	Target Structure
SVO	36%	86%	86%
Complement Clauses	57%	79%	71%
Negation	43%	71%	75%
Relative Clauses	46%	86%	89%
Adverbials	29%	96%	82%
CLD	36%	93%	61%
Coordination	54%	89%	96%
Wh-Questions	57%	86%	89%

Percentage correct verbatim repetitions, grammatical responses and successful target structure responses for each tested structure. SVO: Subject Verb Object, CLD: Clitic Left Dislocation

‘Striped shirt in blue and orange’ – Verbose discourse production in people with focal epilepsy

Strinzel, A-M., De Beer, C., Osthoff, S., Cloppenborg, T., Grewe, P. & Lorenz, A.

Introduction

A variety of language impairments have been described in people with focal epilepsy (PWE), such as frontal and temporal lobe epilepsy. Epilepsy causes morphological and functional changes in the brain that manifest as cognitive impairments, including lexical-semantic and pragmatic deficits (Bell et al., 2001). While word finding difficulties in PWE are not consistently detected, abnormalities in discourse behavior (macrolinguistic level) seem to be more apparent (Dutta et al., 2020). Previous studies used picture descriptions or story retelling tasks and documented a higher number of pauses, a more verbose narrative style, difficulties in coherent narratives, and an increased amount of detail in the narrative discourse of PWE (D’Aprano et al., 2024; Dutta et al., 2020). The aim of the current project is to investigate the semantic-pragmatic performance of PWE in German using a picture description task. It was tested how picture descriptions were affected by the underlying lesion location or epileptic focus in the brain.

Methods

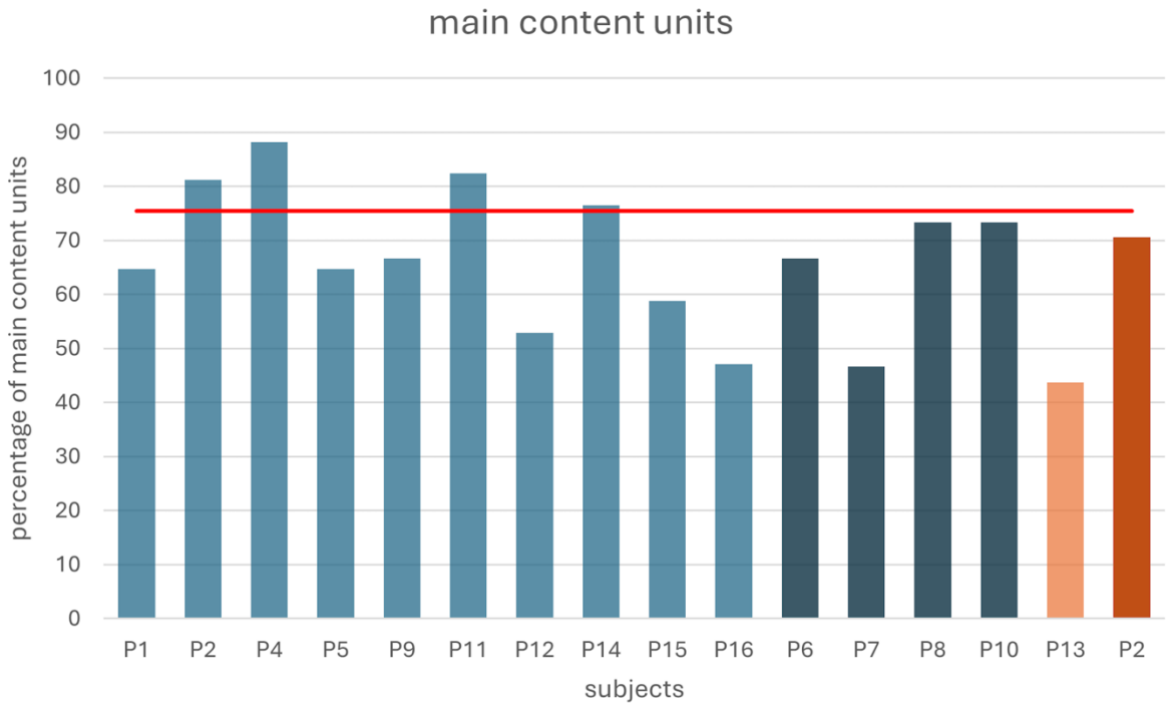
The new version of the ‘Cookie Theft’ picture by Berube et al. (2019) was used. The task involves describing the events depicted in the picture. This is an ongoing study, and data collection is not yet completed. Until today, 16 adults with focal epilepsy were tested. 14 patients with temporal lobe epilepsy (4 right-hemispheric, 9 left-hemispheric, 1 unclear) and 2 patients with frontal lobe epilepsy (one each localized in the left and right hemispheres) have been included. The participants (mean age: 40.12; SD=14.61) are recruited at Mara Hospital in Bielefeld. Additionally, data from a small group of matched neurotypical adults are currently collected (n=20). Furthermore, preliminary German norm data are available (Bayrhof et al., 2024; Wunderlich et al., 2024). All PWE completed the ‘Cookie Theft’ test in addition to various neuropsychological and linguistic pre-assessments. The participants were presented with the picture and received the following instruction: “Now you will see a picture depicting an everyday situation. Please describe everything that is happening in the picture. Please use complete sentences”. The responses were audio recorded and transcribed. The transcripts were analyzed and compared to the available German norm data regarding main content words (e.g. children, dog) and content units (e.g. The man is washing the dishes), produced by neurotypical, language unimpaired adults (healthy controls: HC) in different age ranges (Bayrhof et al., 2024; Wunderlich et al., 2024). Furthermore, an extended qualitative analysis of the data is currently conducted.

Results

The PWE as a grouip produced an average of 18.06 (SD=2.05) out of 25 main content words, which is in the normal range for both nouns (PWE: 11.63; HC: 10.6) and verbs (PWE: 6.31; HC: 6.5) (Bayrhof et al., 2024). While four PWE showed below-average scores of main content words (two of whom had right-hemispheric

temporal lobe epilepsy and two had left-hemispheric temporal lobe epilepsy), 12 PWE were unimpaired in this measure (content words). In contrast, most of the PWE produced fewer main content units than the age-matched HC. The proportions of main content units produced by PWE are depicted in Figure 1. One hundred percent reflects the calculated total number of main content units produced by HC (20–39 years: 17 content units; 40–59 years: 15 content units; 60–79 years: 16 content units). Twelve PWE produced fewer content units compared to the averaged norm data (PWE: 66.1% vs. HC: 75.44%), four PWE were in the normal range (see Figure 1). Notably, a high degree of interindividual variance is observed in the production of specific content units (see Figure 1).

Figure 1. Percentage of main content units for every participant. Light blue: temporal epilepsy left-hemispheric. Dark blue: temporal epilepsy right-hemispheric. Light orange: frontal epilepsy left-hemispheric. Dark orange: frontal epilepsy right-hemispheric. Red line: mean of main content units in norm data (HC). P=participant



A qualitative analysis is used to analyze expository discourse. Eight out of 16 (50%) participants exhibited unusual narrative behavior in relation to the task, as they mentioned irrelevant details such as the appearance of the people, their clothing, or a very detailed description of the picture background. In addition, 2 out of 16 (12.5%) participants exhibited abnormalities by providing incorrect information not depicted in the image. For example, participant P6 counted various objects in the image, such as windows and cookies, which was irrelevant for the task. Results of inferential statistical analysis will be presented.

Discussion

Our current analysis shows that the PWE were mostly unimpaired in the production of main content words when compared to age-matched controls. Only four out of 16 PWE were impaired in this measure. This is consistent with findings of Dutta and colleagues (2020), who also demonstrated that lexical-semantic processes,

including word finding, are relatively spared in PWE. This contrasts with the naming of main content units (concepts), showing lower rates of main content units in PWE compared to HC. Possibly, the main content units no longer received sufficient attention due to the numerous details in picture descriptions of PWE. An initial qualitative analysis revealed anomalies in the narrative structure in PWE. In half of the data, there were strikingly detailed descriptions of the people (hair color, clothing) and the picture background. This is striking, because PWE were instructed to focus on events in the picture, whereas object descriptions are not explicitly requested. This is in line with findings in the literature, which report verbose, circumstantial, and highly detailed narratives in PWE (D'Aprano et al., 2024; Dutta et al., 2020).

References:

Bayrhof, A., Mair-Ramm, V., Frieg, H., Wunderlich, A., & Aichert, I. (2024, November 16). Diagnostik der Textproduktion mit dem neuen Cookie-Theft-Bild – Teil B: Erhebung von Normdaten für den Kernwortschatz [Poster]. Jahrestagung der GAB, Leipzig.

Bell, B. D., Hermann, B. P., Woodard, A. R., Jones, J. E., Rutecki, P. A., Sheth, R., Dow, C. C., & Seidenberg, M. (2001). Object naming and semantic knowledge in temporal lobe epilepsy. Neuropsychology, 15(4), 434–443. https://doi.org/10.1037/0894-4105.15.4.434

Berube, S., Nonnemacher, J., Demsky, C., Glenn, S., Saxena, S., Wright, A., Tippett, D. C., & Hillis, A. E. (2019). Stealing Cookies in the Twenty-First Century: Measures of Spoken Narrative in Healthy Versus Speakers With Aphasia. American Journal of Speech-Language Pathology, 28(1S), 321–329. https://doi.org/10.1044/2018_AJSLP-17-0131

D'Aprano, F., Malpas, C. B., Roberts, S., & Saling, M. M. (2024). Macrolinguistic function in temporal lobe epilepsy: A reinterpretation of circumstantiality. Aphasiology, 38(2), 181–204. https://doi.org/10.1080/02687038.2023.2174371

Dutta, M., Murray, L. L., Miller, W., Innis, I., & Newman, S. (2020). Cognitive–Linguistic Functions in Adults With Epilepsy: Preliminary Electrophysiological and Behavioral Findings. Journal of Speech, Language, and Hearing Research, 63(7), 2403–2417. https://doi.org/10.1044/2020_JSLHR-19-00351

Wunderlich, A., Bayrhof, A., Mair-Ramm, V., & Aichert, I. (2024, November 16). Diagnostik der Textproduktion mit dem neuen Cookie-Theft-Bild Teil A: Erhebung von Normdaten zu Hauptkonzepten [Vortrag]. Jahrestagung der GAB, Leipzig

First insights on nonfluent aphasia in Georgian: A quantitative analysis of spontaneous speech with a focus on morphosyntax

Sopho Tsagareishvili | Tamar Giorgobiani | Manana Gabashvili | Mauro Viganò | Seçkin Arslan

Introduction and aims

People with nonfluent aphasia (PWA) typically produce short, telegraphic utterances marked by reduced syntactic complexity, frequent omissions and substitutions of function words, and impaired verbal morphology in spontaneous speech. This constellation of features is commonly referred to as agrammatism (see Menn & Obler, 1990, among others). However, theoretical accounts of agrammatic aphasia remain biased toward well-studied European languages such as English (Beveridge & Bak, 2011). Despite the presence of linguistic universals, substantial evidence shows that the specific characteristics of language impairments cannot be generalized across languages. In particular, agrammatism manifests differently across languages because they vary in their morphosyntactic properties. For example, in several morphologically rich languages, including Turkish, Finnish, Japanese, West Greenlandic, and Indonesian, agrammatism in spontaneous speech is primarily characterized by syntactic simplification, while verbal morphology remains relatively preserved (Arslan et al., 2016; Anjarningsih et al., 2012; Nedergaard et al., 2020; Menn & Obler, 1990).

Research on aphasia in Georgian remains very limited, and no detailed analyses of aphasic spontaneous speech are currently available. This gap is particularly important given the typological profile of Georgian, which combines relatively flexible word order with highly complex morphology. In Georgian, the verb in particular encodes a large amount of information, including tense, aspect, mood, voice, person, case agreement, polypersonality, and directionality; therefore, much of the grammatical structure is encoded in the verbal system.

This pilot study intends to characterize how nonfluent aphasia manifests in spontaneous speech in Georgian, using quantitative measures of fluency, lexical diversity, morphosyntactic correctness, and syntactic complexity.

Methods

The study included six native Georgian-speaking post-stroke nonfluent PWA and six matched healthy controls (PWA: 1 female; age M=48, SD=14; controls: 2 female; age M=50, SD=14). Spontaneous speech samples were elicited using four tasks: a semi-structured interview, two sequencing picture tasks (Herluf Bidstrup comics), and the Cookie Theft picture description task. Recordings were manually transcribed, and a 200-word sample per participant was extracted for analysis. Lexical classification and morphological annotation were verified using the automated parser of the Georgian National Corpus. Disagreements were resolved through discussion. Quantitative measures of language production included: mean length of utterance; proportion of well-formed utterances; proportion of utterances with clause embedding; number of verbs and nouns; mean verbs per utterance; type-token ratio; noun-to-verb ratio; number of semantic and phonological paraphasias and verb morphology errors.

At the group level, normality was assessed with the Shapiro-Wilk test. Between-group differences were evaluated using independent-samples *t*-tests with Cohen's *d* as the effect size; given the small sample, Wilcoxon rank-sum tests were also conducted to confirm significance. At the individual level, each PWA was compared to the control group using the Crawford-Howell modified *t*-test. All analyses were conducted in R.

Results

At the group level, compared to controls, PWA showed a reduced mean length of utterance (PWA: M=3.09 (SD=0.55); controls: 5.34 (0.61); $t(10)=-6.67$, $p<.001$, $d=-3.85$), a lower proportion of well-formed utterances (PWA: 0.70 (0.18); controls: 0.95 (0.06); $t(10)=-3.27$, $p=.017$, $d=-1.89$), fewer utterances with embedded clauses (PWA: 0.01 (0.01); controls: 0.16 (0.10); $t(10)=-3.49$, $p=.017$, $d=-2.01$), and fewer verbs per utterance (PWA: 0.72 (0.19); controls: 1.34 (0.22); $t(10)=-5.24$, $p<.001$, $d=-3.02$).

At the lexical level, PWA showed a lower type-token ratio (PWA: 0.48 (0.04); controls: 0.56 (0.05); $t(10)=-3.28$, $p<.001$, $d=-1.89$), whereas no significant differences were observed in the number of nouns (PWA: 47.17 (14.18); controls: 50.00 (4.43); $t(10)=-0.47$, $p=.657$, $d=-0.27$), the number of verbs (PWA: 53.17 (21.72); controls: 49.33 (9.58); $t(10)=0.40$, $p=.704$, $d=0.23$), or the noun-to-verb ratio (PWA: 1.13 (0.29); controls: 1.00 (0.23); $t(10)=0.90$, $p=.340$, $d=0.52$).

At the individual level, three PWA followed the overall PWA group pattern. One additional PWA showed a reduced number of verbs ($t(5)=-5.44$, $p=.003$). The remaining two PWA were otherwise comparable to controls but, like the PWA group, showed reduced mean length of utterance ($t(5)=-4.31$, $p=.008$, and $t(5)=-2.08$, $p=.046$) and produced the highest number of verb morphology errors (five and three, respectively). Only a few cases of phonological and semantic paraphasias were observed in PWA. The distribution of performance across measures is shown in *Figure 1*.

Discussion

PWA showed a consistent group pattern, producing shorter and syntactically simpler utterances and reduced lexical diversity compared to controls, while overall verb production and verbal morphology were relatively preserved. At the same time, variability in error profiles and quantitative measures highlights substantial heterogeneity within the sample, underscoring the need for larger datasets and suggesting the presence of distinct profiles within nonfluent aphasia.

Verb access and verb morphology appear relatively preserved, although accompanied by increased repetitions, likely reflecting fillers or perseverations. This result likely reflects the structural properties of Georgian verbs, whose rich morphology encodes grammatical and sentence-level information, allowing them to function as meaningful standalone units. Moreover, this morphological complexity may protect against omissions, which would often result in an ill-formed verb, unlike in languages such as English, where omissions can still yield valid forms. At the same time, patients' speech is characterized by shorter and syntactically simpler utterances, which may represent a core feature of agrammatism in Georgian. This pattern is consistent with findings from other morphologically rich languages discussed in the introduction.

This study provides the first quantitative characterization of spontaneous speech in Georgian-speaking individuals with nonfluent aphasia. Picture sequencing and description tasks proved sensitive to impairment, capturing both lexical and morphosyntactic deficits. Future work will expand PWA and control samples to establish normative data, providing an evidence-based foundation for theoretical interpretation, clinical assessment, and rehabilitation. This is a particularly critical goal for aphasia research in under-investigated languages (Ali et al., 2022).

References

Ali, M., Soroli, E., Jesus, L. M., Cruice, M., Isaksen, J., Visch-Brink, E., & Brady, M. C. (2022). An aphasia research agenda–a consensus statement from the collaboration of aphasia trialists. *Aphasiology*, *36*(4), 555-574.

Anjarningsih, H. Y., Haryadi-Soebadi, R. D., Gofir, A., & Bastiaanse, R. (2012). Characterising agrammatism in standard Indonesian. *Aphasiology*, *26*(6), 757-784.

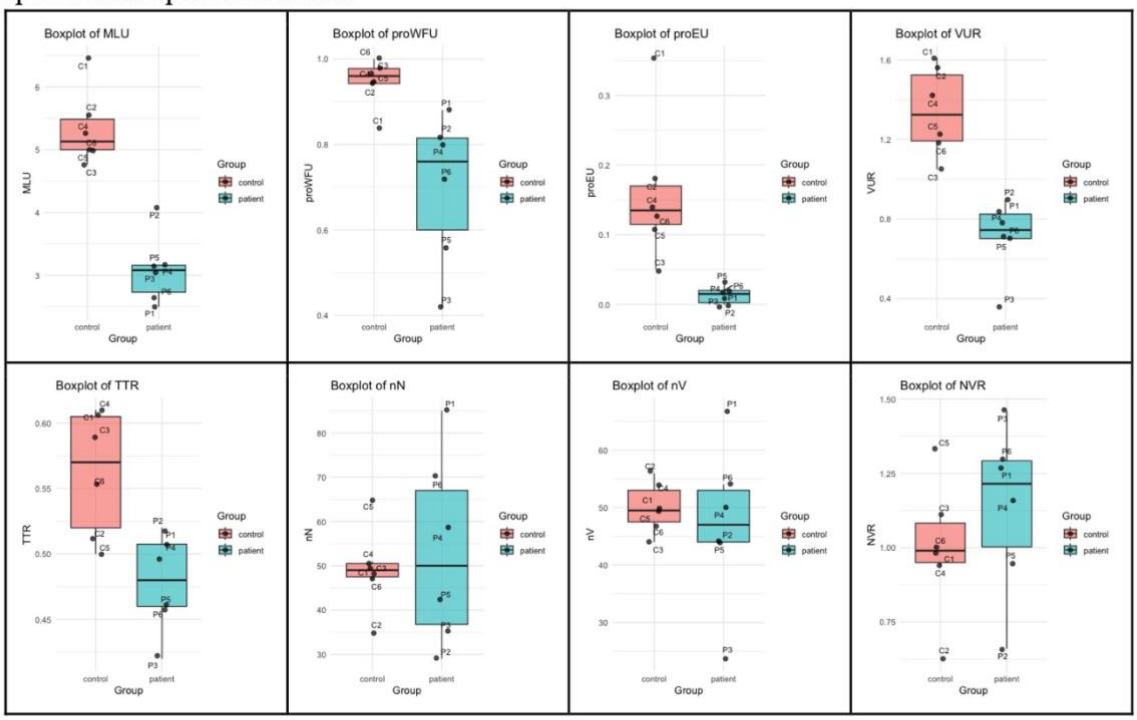
Arslan, S., Bamyacı, E., & Bastiaanse, R. (2016). A characterization of verb use in Turkish agrammatic narrative speech. *Clinical Linguistics & Phonetics*, *30*(6), 449-469.

Beveridge, M. E. L., & Bak, T. H. (2011). The languages of aphasia research: Bias and diversity. *Aphasiology*, *25*(12), 1451-1468.

Menn, L., & Obler, L.K. (Eds.). (1990). *Agrammatic aphasia: A cross-language narrative sourcebook*. John Benjamins Publishing.

Nedergaard, J. S., Martínez-Ferreiro, S., Fortescue, M. D., & Boye, K. (2020). Non-fluent aphasia in a polysynthetic language: Five case studies. *Aphasiology*, *34*(6), 675-694.

Figure 1. Group distributions for PWA and controls with individual performance across spontaneous speech measures.



Notes. MLU: Mean Length of Utterance; proWFU: proportion of well-formed utterances relative to the total n. of utterances; proEU: proportion of utterances with clause embedding relative to the total n. of well-formed utterances; VUR: n. of verbs per utterance ratio; TTR: type-token Ratio; nN: number of nouns; nV: number of verbs; NVR: noun-to-verb ratio.

Towards preserving errors: fine-tuning automatic speech recognition for literal transcription of aphasic speech

Frank Tsiwah | Rutger Luinge

Introduction and aims

Spontaneous speech is a rich source of diagnostic, language processing and language usage data in aphasia, and yet it remains underused in both clinical practice and research. This is because it is time consuming to transcribe and analyse. Time and resources are consistently the primary barriers to connected-speech work (Stark et al., 2021; Bryant et al., 2017), with estimates ranging from 1.5–3 hours of transcription per 10 minutes of speech (Boles, 1998) to 60 minutes of combined transcription and analysis per minute of conversation (Armstrong et al., 2007). This constrains sample sizes for studies using spontaneous speech and impedes a holistic assessment and/or treatment of language impairment in functional communication context.

Automatic speech recognition (ASR) offers a potential of automating spontaneous speech transcription. For example, *Whisper*, a state-of-the-art ASR model achieves near-human accuracy on healthy speech but was trained solely of neurotypical speech data and is engineered to "clean up" transcription outputs by auto-correcting all errors. That is, on aphasic speech, it normalises paraphasias to the nearest real word and has been found to hallucinate content on disfluent input (Koenecke et al., 2024), and thus, removing the exact information (speech and language errors) clinicians and researchers need for profiling people with aphasia (PWA).

A clinically useful ASR system must provide literal transcriptions that preserve errors (e.g., lexical and grammatical errors), neologisms, deviant sounds, and partial words – the type of errors central to diagnosis, aphasia subtype classification, and therapy planning. This study therefore addresses three questions: (1) Can an end-to-end ASR model be fine-tuned to transcribe aphasic speech literally? (2) Does it generalise across aphasia subtypes and different recording settings? (3) How does it compare with Whisper v3 as a general-purpose baseline? To our knowledge, this is the first ASR study explicitly designed and trained to enforce literal, error-preserving transcription of aphasic speech.

Method

We used five English language corpora from AphasiaBank (*NEURAL, NEURAL 2, UNH, Wright, Richardson*), containing picture descriptions (e.g., cookie theft, broken window, and cat rescue), narrative (e.g., Cinderella story) and open-ended question tasks from people with aphasia, with time-aligned transcripts and subtype labels.

Because literal transcription requires closely representing non-words and atypical pronunciation data in the transcripts, we first converted each word to an IPA-based phonemic representation using a grapheme to phoneme model. This step allowed us to treat the transcription problem as predicting phoneme strings rather than orthographic words. The audio and phoneme sequences were then aligned at the sentence level using forced alignment.

Three pre-trained versions of *wav2vec 2.0* (Baevski et al., 2020) ASR-model: base monolingual (English), multilingual XLSR, and large monolingual (English), were compared.

Each model was fine-tuned with identical hyperparameters and data. The best-performing model was then selected and subsequently trained on increasing amounts of data from all five corpora to examine the effect of training-set size and across-PWA variability. Performance was evaluated per aphasia subtype (Wernicke, Broca, Anomic, Conduction, Transcortical Motor) and across clinics (different data repositories on AphasiaBank), using phoneme error rate (PER) and word error rate (WER). Whisper v3 was evaluated on the same test sets under default settings.

Results

Wav2vec 2.0 large – English monolingual was the best-performing model, reaching PERs of 0.066 (controls), 0.170 (fluent aphasia), and 0.351 (non-fluent aphasia) on 8 hours of training data, outperforming the base and XLSR variants on all groups. Note that the numbers are in probabilities ranging from 0 – 1, where 0 and 1 are equivalent to 0% and100% error rates, respectively. At the aphasia subtype level, performance varied in systematic and clinically interpretable ways. For Wernicke's aphasia, the model achieved a PER of 0.076 and a WER of 0.122. For Anomic aphasia, PER was 0.125 and WER 0.212, while Conduction aphasia showed PER 0.150 and WER 0.255. Transcortical Motor aphasia was in between, with PER 0.105 and WER 0.169. As expected, Broca aphasia was the most challenging: PER was 0.267 and WER 0.373. These results are reported in Figure 1. Error analyses using phoneme level confusion matrices showed that many remaining errors involved predictable confusions between acoustically similar sounds, such as stops differing only in voicing (e.g., /ð/ ↔ /θ/, /s/ ↔ /z/).

Discussion

This study shows that literal, error preserving automatic transcription of aphasic speech is feasible with current end to end ASR technology, even when only modest amounts of carefully curated aphasic data are available. By training wav2vec 2.0 to output phoneme sequences instead of words, the system captures non words, sound substitutions and partial words that are usually lost when off the shelf systems "clean up" the signal. In direct comparison with Whisper v3, the fine-tuned model halves word error rates on aphasic speech while preserving clinically relevant error patterns, suggesting that aphasia adapted ASR can add value on top of much larger general-purpose models.

The subtype pattern is also informative. The model's performance on Wernicke and other fluent aphasias suggests that the model can follow highly fluent but error rich speech without "correcting away" neologisms and paraphasias, whereas higher error rates for Broca and other non-fluent profiles likely reflect shorter utterances, more challenging acoustics and greater inconsistency in the manual reference transcripts. Practically, a phoneme-level transcription that preserves errors should reduce the cost of human post-editing of spontaneous speech relative to transcribing from scratch.

There are nonetheless important limitations. The training data are restricted to English and to a small set of clinics and recording setups, so performance cannot be assumed to generalise to other languages or dialects. Phoneme level models depend on a fixed symbol inventory and may not fully capture gradient articulatory deviations or voice quality changes, and all outputs remain machine generated and require human post-editing. Nevertheless, the present work points to a practical path forward: with a reasonable amount of aphasic training material and a phoneme focused design, ASR can produce literal transcriptions that outperform a widely

used general purpose model while preserving the linguistic information that matters most in aphasia research and clinical documentation.

References

Armstrong, L., Brady, M., Mackenzie, C., & Norrie, J. (2007). Transcription-less analysis of aphasic discourse: A clinician's dream or a possibility?. *Aphasiology, 21*(3-4), 355-374.
Baevski, A., Zhou, Y., Mohamed, A., & Auli, M. (2020). wav2vec 2.0: A framework for self-supervised learning of speech representations. *Advances in neural information processing systems*, *33*, 12449-12460.
Boles, L. (1998) Conversational discourse analysis as a method for evaluating progress in aphasia: a case report. *Journal of Communication Disorders, 31*, 261–274.
Bryant, L., Spencer, E., & Ferguson, A. (2017). Clinical use of linguistic discourse analysis for the assessment of language in aphasia. Aphasiology, 31(10), 1105-1126.
Koenecke, A., Choi, A. S. G., Mei, K. X., Schellmann, H., & Sloane, M. (2024, June). Careless whisper: Speech-to-text hallucination harms. In *Proceedings of the 2024 ACM conference on fairness, accountability, and transparency* (pp. 1672-1681).
Stark, B. C., Dutta, M., Murray, L. L., Bryant, L., Fromm, D., MacWhinney, B., ... & Sharma, S. (2021). Standardizing assessment of spoken discourse in aphasia: A working group with deliverables. *American Journal of Speech-Language Pathology, 30*(1S), 491-502.

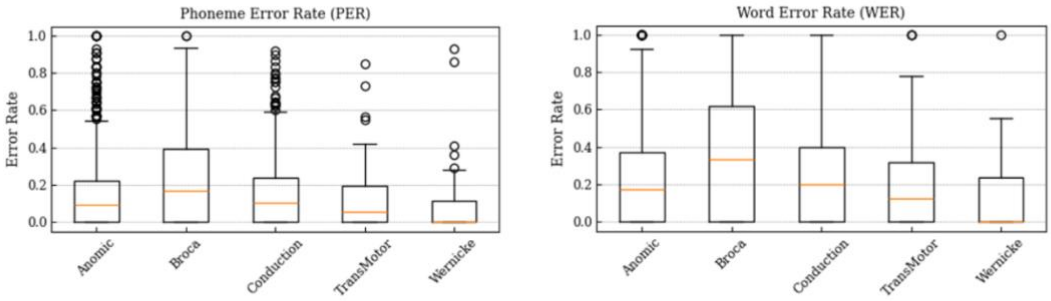


Figure 1. Phoneme error rate (PER, left) and word error rate (WER, right) of the fine-tuned wav2vec 2.0 model for each aphasia subtype (Anomic, Broca, Conduction, Transcortical Motor, Wernicke). The boxplots show how error rates vary across speakers within each subtype, with lower values indicating more accurate, more literal transcriptions.

Hungarian adaptation of the Apraxia of Speech Rating Scale–3.5: A pilot study of differential diagnosis in post-stroke AOS and aphasia

Vivien Vásári | Lilla Zakariás

Introduction and aims

Post-stroke aphasia and apraxia of speech (AOS) frequently co-occur. Differentiating their symptoms in cases of comorbidity is often challenging due to overlapping etiological and neuroanatomical characteristics (Basilakos et al., 2015; Code, 2021). Differentiating AOS from aphasic phonological production errors, particularly in Broca's and conduction aphasia, is especially challenging (Jonkers et al., 2017; Ogar et al., 2005). Currently, no adapted assessment tool is available in Hungarian clinical practice for assessing the presence and severity of AOS. The aim of this study was to introduce the Hungarian adaptation of the *Apraxia of Speech Rating Scale–3.5* (Duffy et al., 2023) and to present results from a pilot study examining how well the scale differentiates post-stroke individuals with and without AOS.

Methods

A total of 15 participants (6 women, age range: 42-76 years) participated in the pilot study. Inclusion criteria were: subacute or chronic ischemic or haemorrhagic stroke; aphasia (based on the Hungarian Aphasia Screening Test [HAST; Zakariás & Lukács, 2023]) and/or AOS (as determined by the clinical judgment of the treating speech-language pathologists); native Hungarian-speaking adults; and adequate hearing and vision, with or without correction. Exclusion criteria included dysarthria, comorbid neurological or psychiatric disorders, and moderate to severe hearing impairment. Participants were assigned to two groups: (1) AOS, with or without aphasia (AOS only: n = 2; AOS + aphasia: n = 6); and (2) no AOS, with aphasia only (n = 7). Participants with aphasia were classified as having Broca's, conduction, or anomic aphasia.

The original scale consists of 13 rated features, each scored on a 5-point scale (0–4), with higher scores indicating greater severity. Features are grouped into three subscales: Phonetic, Prosodic, and Other, and a Total score is derived by summing subscale scores. Ratings are based on performance during spontaneous speech, repetition, picture description, and reading tasks. During adaptation, English lexical items were matched to Hungarian equivalents based on lemma frequency, word class, vowel–consonant structure, and semantic properties. Sentence adaptation considered syllable count, word count, syntactic complexity, consonant cluster density. Due to structural differences between the two languages, the Hungarian stimuli differ from the original English version. Assessments were conducted across two sessions: the first included aphasia screening (HAST), administration of the Hungarian ASRS–3.5, and nonverbal oral apraxia assessment using ABA–2 Subtest 3 (Dabul, 2000); the second included the administration of the Western Aphasia Battery (Osmánné-Sági, 1991) when aphasia was present. Audio recordings were obtained during scale administration using a wireless, dual-channel, noise-reducing microphone (Boya Boyalink 2 Lavalier). The scale was evaluated by a speech-language pathologist with 7 years of clinical experience. Due to the small sample size and non-normal distribution of the data, analyses were conducted in JASP (version 0.96.0) using non-parametric Mann–Whitney U tests to compare groups on ASRS–3.5 subscale and total scores, while descriptive statistics were used to examine feature frequencies across groups.

Results

Comparisons revealed significant group differences in the Phonetic ($U = 5.5$; $p = 0.008$) and Prosodic ($U = 10.5$; $p = 0.046$) subscales. A significant difference was also observed in the Total score ($U = 8.5$; $p = 0.027$). No significant difference was found for the Other Features subscale ($U = 19$; $p = 0.314$). Results are presented in Table 1. Descriptive analyses indicated that phonetic distortions, distorted additions, increased errors with length or complexity, segmentation across words, slow speech rate, alternating motion rate abnormalities, and visible or silent groping occurred more frequently in the AOS group, suggesting that these may represent candidate diagnostically informative markers of AOS in this sample. Furthermore, all but one participant in the AOS group demonstrated at least five ASRS features rated above 0, whereas all but one participant in the aphasia-only group demonstrated four or fewer features.

Discussion

Significant differences between the AOS and no-AOS groups were observed in the Phonetic and Prosodic subscale scores, as well as in Total score. These findings suggest that the Hungarian ASRS–3.5 captures characteristic features of AOS and may help differentiate individuals with AOS from those with aphasia without AOS. The scale thus may support identification of AOS in Hungarian research and clinical contexts and aid differential diagnosis. These findings should be interpreted cautiously given the small pilot sample. Future studies should examine interrater reliability and validate the scale in larger samples, enabling investigation of its ability to differentiate AOS from specific aphasia subtypes (e.g., Broca's and conduction aphasia) and to further characterize hallmark diagnostic features.

References

Basilakos, A., Rorden, C., Bonilha, L., Moser, D., & Fridriksson, J. (2015). Patterns of poststroke brain damage that predict speech production errors in apraxia of speech and aphasia dissociate. *Stroke, 46*(6), 1561–1566.
Code, C. (2021). Contemporary issues in apraxia of speech. *Aphasiology, 35*(4), 391–396.
Dabul, B. (2000). ABA-2: Apraxia Battery for Adults (2dn edition). Pro-Ed, Austin, Texas
Duffy, J. R. et al. (2023). The Apraxia of Speech Rating Scale: Reliability, validity, and utility. *American Journal of Speech-Language Pathology, 32*(2), 469–491.
Jonkers, R., Feiken, J., & Stuive, I. (2017). Diagnosing apraxia of speech on the basis of eight distinctive signs. *Canadian Journal of Speech-Language Pathology and Audiology, 41*(3), 303–319.
Ogar, J., Slama, H., Dronkers, N., Amici, S., Gorno-Tempini, M. L. (2005). Apraxia of speech: An overview. *Neurocase, 11*(6), 427–432.
Osmánné Sági J. (1991). Az afázia klasszifikációja és diagnosztikája. *Ideggyógyászati Szemle, 44*, 339–362.
Zakariás, L., & Lukács, Á. (2023). The first aphasia screening test in Hungarian: A preliminary study on validity and diagnostic accuracy. *Plos one, 18*(8), e0290153.

Table 1. Group comparisons of ASRS–3.5 subscale and total scores between participants with and without AOS

		AOS, with or without aphasia	No AOS, aphasia only		Statistics
	Max	***M (SD)***	***M (SD)***	***U***	
Phonetic features	**16**	**3.13 (1.89)**	**0.29 (0.49)**	**5.5**	**(**
Prosodic features	**16**	**7.00 (5.78)**	**1.14 (1.22)**	**10.5**	
Other features	20	5.25 (3.01)	3.57 (1.13)	19.0	
Total score	**52**	**15.38 (9.49)**	**5.00 (1.83)**	**8.5**	

Note. AOS = apraxia of speech. $^{*}p < 0.05$; $^{**}p < 0.01$

Abstract morphosyntactic features versus overt morphophonological cues in the comprehension of French object relatives in Primary Progressive Aphasia

Mauro Viganò | Sophie Ferrieux | Lucie Grimont | Marie-Lea Muzerelle | Cécile Rebillard | Carlo Cecchetto | Marc Teichmann

Introduction and aims. Syntactic deficits are widely documented in aphasia, particularly in the comprehension of non-canonical sentence structures that require full syntactic parsing to assign thematic roles. A robust finding is the asymmetry between subject relatives (SR; 'Show me the magician that chases the cook') and object relatives (OR; 'Show me the magician that the cook chases'), with ORs typically harder to comprehend because the moved object must be interpreted across an intervening subject. Crucially, this difficulty is reduced when the head of the relative ('the magician,' the logical object) and the embedded subject ('the cook') differ in morphosyntactic features such as number, as in 'Show me the magician that the cooks chase' (Adelt et al., 2017 for German; Terzi & Nanousi, 2018 for Greek; Viganò et al., 2025 for Italian).

Within syntactic theory, this pattern is accounted for by *featural Relativized Minimality* (Rizzi, 2018; also termed *Generalized Minimality*: Garraffa & Grillo, 2008; Grillo, 2008), whereby OR difficulty depends on the structural relation between the moved object, its base-generated position, and the intervening subject, and is modulated by their similarity in abstract morphosyntactic features that are grammatically active in subject-verb agreement (Adelt et al., 2017; Viganò et al., 2025). An alternative explanation, grounded in *Cue-Based Retrieval* approaches (Engelmann et al., 2019), attributes difficulty to linear interference during incremental sentence processing and explains facilitation in terms of overt morphophonological cues (e.g., verb inflection) that guide retrieval even without full syntactic computation (Terzi & Nanousi, 2018). While the two accounts yield similar predictions, they crucially differ in positing either abstract, grammatically active features or surface morphophonological cues as the source of facilitation (cf. Viganò et al., 2026).

This study tests these competing accounts in the comprehension of French relative clauses across the three variants of PPA.

Methods. French provides an optimal testing ground for dissociating abstract morphosyntax from overt morphophonology, because the number feature participates in subject-verb agreement but is not always phonologically realised in verb inflection. For example, singular and plural forms differ in some verbs (e.g. *suivre*='to chase', 3P.SG: /syi/ vs. 3P.PL: /syiv/) but remain identical in others (e.g. *pousser*='to push', 3P.SG=3P.PL: /pus/). This property allows a direct test of whether facilitation arises from abstract feature computation or from overt phonological cues. We assessed auditory comprehension of SRs and ORs using a sentence-to-picture matching task in a 2×2×2 design manipulating (i) relative clause type (subject vs. object), (ii) number features (match vs. mismatch), and (iii) audibility of number marking (audible vs. non-audible). Participants included individuals with nonfluent/agrammatic PPA (nfavPPA, n=4), logopenic PPA (lvPPA, n=3), and semantic PPA (svPPA, n=8), together with matched healthy controls (n=23). Given the small sample size,

results are treated as preliminary and reported using descriptive statistics alongside Cohen's *h* effect sizes.

Results. Results revealed a robust SR-OR asymmetry in nfavPPA (SR=94%, OR=57%, diff=37%, *h*=0.9, strong effect) and lvPPA participants (SR=93%, OR=29%, diff=64%, *h*=1.4, strong effect), with markedly lower accuracy for ORs and almost ceiling performance for SRs. In contrast, svPPA participants and controls performed at ceiling across conditions (accuracy >95%). As expected, number mismatch facilitated OR comprehension in nfavPPA (mismatch=69%, match=46%, diff=23%, *h*=0.5) and lvPPA participants (mismatch=39%, match=19%, diff=20%, *h*=0.4). No such difference was observed in svPPA participants or controls (diff<4%), as expected given their ceiling performance. Crucially, phonological audibility of number marking yielded no meaningful differences across all groups (diff<4%). The distribution of performance across the three contrasts (SR vs. OR; number match vs. mismatch in ORs; audibility vs. non-audibility in mismatching ORs) is shown in *Figure 1*.

Discussion. These findings challenge *Cue-Based Retrieval* accounts in favour of *featural Relativized/Generalized Minimality*, as they show that number dissimilarity facilitates OR comprehension at an abstract level, independently of its morphophonemic realisation.

They also confirm an increased vulnerability to syntatic intervention configurations in nfavPPA and lvPPA, in line with the clinical profile of grammatical impairment in these variants (Teichmann et al., 2012; Patel et al., 2021). Importantly, this pattern does not point to degraded syntactic representations; rather, it suggests processing limitations that emerge specifically during the computation of specific configurations, as evidenced by the modulation of performance by the similarity between the elements involved (cf. Grillo, 2008).

More broadly, the results indicate that grammatical impairments do not affect syntax and morphology in parallel: while increased syntactic complexity is associated with greater difficulty, richer morphology and more morphological information can instead facilitate comprehension. This dissociation aligns with lesion-symptom mapping evidence in post-stroke aphasia showing distinct neural correlates for primarily morphological versus sentence-level deficits (Beber et al., 2025).

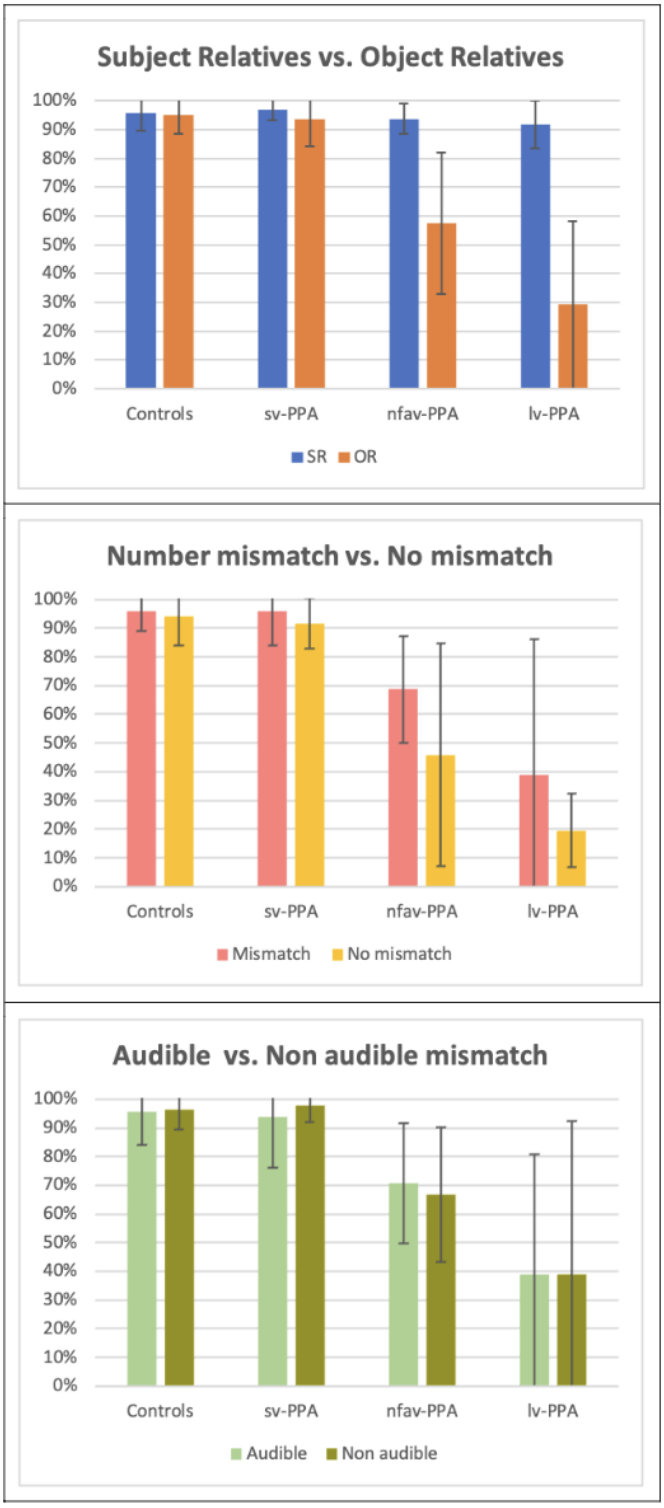
From a clinical perspective, these findings highlight the importance of assessing morphosyntactic and morphophonemic factors alongside syntax in aphasia, in order to disentangle their potentially independent contributions to sentence processing.

References

Adelt, A., Stadie, N.,…& Burchert, F. (2017). Feature dissimilarities in the processing of German relative clauses in aphasia. *Journal of neurolinguistics*, 44, 17-37.
Beber, S., Bontempi, G., Miceli, G., & Tettamanti, M. (2025). The neurofunctional correlates of morphosyntactic and thematic impairments in aphasia: A systematic review and meta-analysis. *Neuropsychology Review*, 35(3), 483-516.
Engelmann, F., Jäger, L. A., & Vasishth, S. (2019). The effect of prominence and cue association on retrieval processes: A computational account. *Cognitive Science*, 43(12), e12800.
Garraffa, M., & Grillo, N. (2008). Canonicity effects as grammatical phenomena. *Journal of neurolinguistics*, 21(2), 177-197

Grillo, N. (2008). *Generalized Minimality: Syntactic underspecification in Broca's aphasia.* PhD dissertation, Utrecht University.
Grodzinsky, Y. (1989). Agrammatic comprehension of relative clauses. *Brain and Language, 37*(3), 480-499.
Patel, N., Peterson, K.,...& Garrard, P. (2022). A 'Mini Linguistic State Examination' to classify primary progressive aphasia. *Brain Communications*, 4(2).
Rizzi, L. (2018). Intervention effects in grammar and language acquisition. *Probus*, 30(2), 339-367.
Teichmann, M., Kas, A.,...& Migliaccio, R. (2013). Deciphering logopenic primary progressive aphasia: a clinical, imaging and biomarker investigation. *Brain*, 136(11), 3474-3488.
Terzi, A., & Nanousi, V. (2018). Intervention effects in the relative clauses of agrammatics: The role of gender and case. *Glossa: a journal of general linguistics*, 3(1).
Viganò, M., Gilardone, G., Cassinelli, D., & Cecchetto, C. (2025). The challenge of variability for syntactic accounts of agrammatism: A study on feature dissimilarity in Italian relatives. *Studies in Language and Mind*, 6, 91-106.
Viganò, M., Toneatto, C., & Cecchetto, C. (2026). A quest for a unitary explanation of the object disadvantage in incremental processing: Evidence from a modified maze task with Italian relative clauses. *Applied Psycholinguistics*, 47.

Figure 1. Distribution of performance across groups (controls, svPPA, nfavPPA, lvPPA) as a function of the experimental factors. The three panels illustrate the key contrasts: SR vs. OR; number match vs. mismatch in ORs; and audibility vs. non-audibility in mismatching ORs.



A Usage-based Approach to Assessing Sentence Comprehension

Rinn Wang

Usage-based frameworks propose that linguistic knowledge is constantly shaped by experience, and that more familiar forms are typically more strongly entrenched and processed more efficiently. In research on production, this facilitative effect has often been indexed through frequency or collocation strength, and clinical evidence suggests that conditions such as aphasia and Alzheimer's disease can be characterized by increased use of highly familiar, strongly collocated utterances. By contrast, comprehension research has rarely examined how frequency-related information shapes processing.

We developed a novel usage-based sentence–picture verification paradigm that examines comprehension under the joint influence of structural and frequency-related factors. The paradigm includes two sentence types with distinct assessment aims: idiomatic sentences and reversible literal sentences. Idioms provide a useful test case of usage-based facilitation in conventionalised multiword expressions, and reversible literal sentences allow the examination of structurally driven thematic-role assignment under reduced semantic bias. Across both sentence types, items were balanced for active and passive voice and constructed around a shared set of target verbs. Two continuous usage-based predictors were examined alongside grammatical voice: sentence-level collocation strength, defined as the strength of association between adjacent word pairs within a sentence, and verb-level passive bias, defined as the frequency with which a verb occurs in passive constructions. Participants heard each sentence once and judged whether a presented picture matched its meaning. We report data from 45 neurotypical adults.

Results show high overall accuracy across sentence types, and we carried out reaction time (RT) analyses on correct responses only. Results revealed that processing cost was not explained by a simple facilitative effect of higher frequency or by a uniform disadvantage for passive constructions. Instead, voice and frequency-related information interacted in ways that varied across stimulus types. Specifically, higher verb passive bias ($b = -0.009$, $p < .001$) and higher collocation strength ($b = -0.006$, $p = .007$) were associated with shorter RTs for passive idioms, but longer RTs for active idioms. This pattern was not observed in reversible literal sentences.

These findings show, in a neurotypical baseline, that frequency-related information can modulate online sentence processing on a continuous scale. As such, this paradigm offers a promising route for detecting subtle comprehension costs in aphasia or other neurological conditions that may be missed by conventional domain-specific tests, while also portraying cognitive deficits more realistically in terms of how they affect everyday communication.

Measuring confrontation naming under increased cognitive demands in young stroke: the Picture Naming Cards Test

Rosemarije P.C. Weterings | Roy P.C. Kessels | Frank-Erik de Leeuw | Iris Wiegand | Mirjam C. Otters | Vitoria Piai

Introduction and aims

Individuals who suffered a stroke at young-adult age (<50 years, "young stroke" henceforth) often experience word-finding difficulties in daily conversation, which regularly remain unsubstantiated by standardized confrontation naming tests. We developed a novel speeded picture-naming test – the Picture Naming Cards Test – designed to enable more ecologically valid assessment of naming abilities, approximating everyday conversational challenges. The test does so by introducing internally imposed time pressure and semantic interference. The time it takes to complete a picture naming card (i.e., total naming duration) and total accuracy are quantified. We developed and validated the test materials. Additionally, we investigated whether the test's outcomes aligned with subjective word-finding complaints in young-stroke survivors, and if so, which outcome measure provided the best discriminative power in terms of sensitivity and specificity between young-stroke and matched controls.

Methods

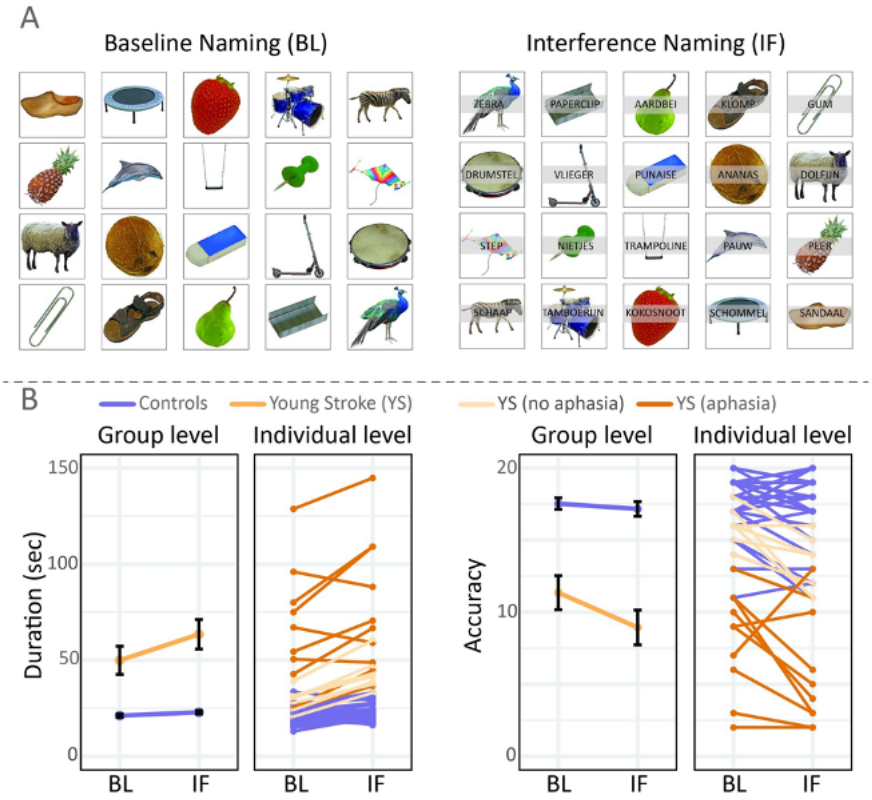
The Picture Naming Cards Test consists of two cards each containing the same 20 pictures (Figure 1A). In the baseline card condition, bare pictures are presented. In the interference card condition, each picture is overlaid with a written distractor word from the same semantic category (e.g., *pear* over the picture of a strawberry), increasing lexical competition. Experiments 1, 2, and 3 were conducted in a population of neurologically healthy adults. In Experiments 1 and 2, we validated the materials and investigated the effects of speed, repetition priming, and interference. Experiment 3 provided initial reference data based on a sample of 109 participants across a wide age range. In Experiment 4, 17 young-stroke survivors (with or without aphasia) reporting subjective word-finding difficulties and 30 individually matched, neurologically healthy controls completed the Picture Naming Cards Test.

Results

Across Experiments 1, 2, and 3, participants were significantly slower in the interference card condition than in the baseline card condition, demonstrating the expected semantic interference effect. Repetition priming and age-related slowing were also observed, next to concurrent validity with an established test of speeded picture naming under externally imposed time pressure. Importantly, total naming duration proved more sensitive than accuracy in capturing performance differences between the two naming card conditions. In Experiment 4, young-stroke survivors were disproportionately slower and made more errors than controls on both group and individual levels, especially in the interference card condition (Figure 1B). Their decreased performance could not be attributed to a lower general processing speed or aphasia diagnosis. Finally, total naming duration on the interference naming card had the best discriminative power between young-stroke and matched controls (Youden's index = .97, vs all other outcomes measures < .71), with the highest sensitivity (100%) and specificity (97%).

Discussion

Altogether, these findings demonstrate the initial validity of the Picture Naming Cards Test. By emphasizing speed, the test seems to capture a relevant aspect of word finding in everyday conversations. The reported reference data will facilitate establishing normative scores. We developed the Picture Naming Cards Test for Dutch, but we highlight its adaptable framework as it can be easily modified for use in other languages. Furthermore, we made sure the Picture Naming Cards Test is easily adaptable and practical for experimental or clinical use, either presented on a screen or as a straightforward paper-and-pencil task, timed with a stopwatch. Although technologies are increasingly being utilized in clinical settings, challenges remain at the implementational level. Finally, this initial demonstration of good discriminative power between a young-stroke population with subjective word-finding complaints and matched controls supports the test's potential value as a brief, focused clinical screener for assessing mild word-finding difficulties.



A: Schematic overview of the baseline (left) and interference (right) naming cards from the Picture Naming Cards Test. All pictures on one card are displayed at once. Words are translated from Dutch for illustrative purposes. **B:** Behavioral results on the Picture Naming Cards Test in Experiment 4: total naming duration in seconds (left panel) and accuracy scores (right panel) per card and group at group level (left-hand side) and individual level (right-hand side). Error bars in black indicate the standard error of the mean. Controls are displayed in purple, young-stroke in orange (light orange: no aphasia; dark orange: aphasia). BL = the baseline naming card; IF = the interference naming card.

Investigating the comprehension of discourse particles via a selection test in Dutch-speaking persons with aphasia

Imke Wets | Vitória Piai | Marina B. Ruiter | Lotte Hogeweg | Helen de Hoop

Introduction and aims

Aphasia is an acquired language disorder characterised by linguistic deficits, which cannot only limit a person's everyday communication activities but also affect their interactions and social relationships [1]. Decades of research in aphasia have mostly focused on language use in isolation by examining words or single sentences. Accordingly, the standard clinical assessment of persons with aphasia (PWA) includes decontextualized language measures, such as bare-picture naming (i.e. naming words), even though PWA themselves typically prefer to improve their ability to communicate in everyday life [1]. In recent years, however, aphasiology has seen an increase in research into language in context. Analysis of language in context can be seen as a supplementary form of assessment identifying deficits and/or compensatory abilities that standard assessments focusing on words and sentences in isolation might overlook [2]. Moreover, it could contribute towards a better understanding of the difficulties PWA encounter in communication, which in turn could guide interventions aimed at improving everyday communication [2]. Interpersonal discourse particles (e.g., 'actually', 'indeed') are linguistic tools that enable speakers to respond to contextual expectations. In two recent studies, Wets et al. [3,4] examined the use of six Dutch discourse particles in the language production of Dutch-speaking PWA and people with primary progressive aphasia (PPA). While some discourse particles (e.g., *ja* 'yeah') were produced more by PWA compared to neurotypically healthy speakers (NHS), this was not the case for some other discourse particles. The authors noted that interpersonal discourse particles (e.g., *eigenlijk* 'actually') occurred infrequently in general. Therefore, given that PWA may have reduced language processing capacity [5], it remains an open question whether they can make proper use of discourse particles like *eigenlijk* 'actually', which fulfil interpersonal functions that may require substantial mental effort [6]. Thus, it is currently unclear from a research or diagnostic perspective to what extent Dutch-speaking PWA can successfully integrate these words into a communicative context. Therefore, the aim of this study is to examine the performance of PWA and NHS in selecting an interpersonal discourse particle that best fits a given communicative context.

Methods

35 Dutch-speaking PWA and 20 NHS completed a discourse particle selection test, comprising both interpersonal discourse particles and nouns. Participants were presented with a simultaneously spoken and written dialogue accompanied by pictograms to overcome possible reading problems. Participants had to decide which word fitted best in the blank in the last sentence by choosing from four options. A cumulative link mixed model approach was used to analyse item scores in order to assess how PWA and NHS perform in selecting discourse particles. Furthermore, the relation between discourse particle selection scores and the scores on a standardised working memory task was assessed using a Spearman correlation analysis.

Results

PWA performed more poorly in selecting interpersonal discourse particles compared to NHS. Moreover, better sentence comprehension was related to better discourse particle scores (OR = 1.05, 95% CI [1.02-1.08]) (Figure 1). We also found that better performance on a working memory task was related to better discourse particle scores in PWA, $p = .021$.

Discussion

The present study investigated the use of Dutch interpersonal discourse particles in a selection test designed to shed light on PWA's ability to process interpersonal discourse particles. Although the importance of interpersonal discourse particles is well established in the linguistic field of pragmatics [6], it is currently unclear to what extent PWA can successfully integrate these words into a communicative context. Our results showed that PWA selected fewer interpersonal discourse particles correctly, despite understanding the task and the materials, thus indicating they have deficits in integrating interpersonal discourse particles in a communicative context. Furthermore, we found that PWA who obtain a lower discourse particle selection score obtained a lower score on a working memory task as well. This is in line with theoretical proposals of limited language processing capacity in PWA [7]. Taken together, these findings show that examining discourse particles can lead to a better understanding of aphasic discourse and may provide valuable diagnostic insights for clinicians.

References

[1] Wallace, SJ, Worrall, L, Rose, T, Le Dorze, G, Cruice, M, Isaksen, J, et al. Which outcomes are most important to people with aphasia and their families: an international nominal group technique study framed within the ICF. Disabil Rehabil. 2017;39(14):1364-1379.
[2] Bryant, L, Ferguson, A, & Spencer, E. Linguistic analysis of discourse in aphasia: a review of the literature. Clin Linguist Phon. 2016;30(7):489-518.
[3] Wets, I, Bleumink, K, van Lier, L, Opheij, E, Hogeweg, L, de Hoop, H, et al. Exploring the production of discourse particles by persons with aphasia. Aphasiology. 2025;39(11):1-21.
[4] Wets, I, Piai, V, Jiskoot, L, van den Berg, E, Janssen, N., Hogeweg, L, et al. The use of discourse particles in oral picture description by individuals with primary progressive aphasia. Cortex. 2025; 193:134-143.
[5] Hula, WD, & McNeil, MR. Models of attention and dual-task performance as explanatory constructs in aphasia. Semin Speech Lang. 2008;29(3):169-187.
[6] van Bergen, G, & Bosker, HR. Linguistic expectation management in online discourse processing: an investigation of Dutch *inderdaad* 'indeed' and *eigenlijk* 'actually'. J. Mem Lang. 2018;103:191-209.
[7] Caplan, D, Baker, C, & Dehaut, F. Syntactic determinants of sentence comprehension in aphasia. Cognition. 1985;21:117-175.

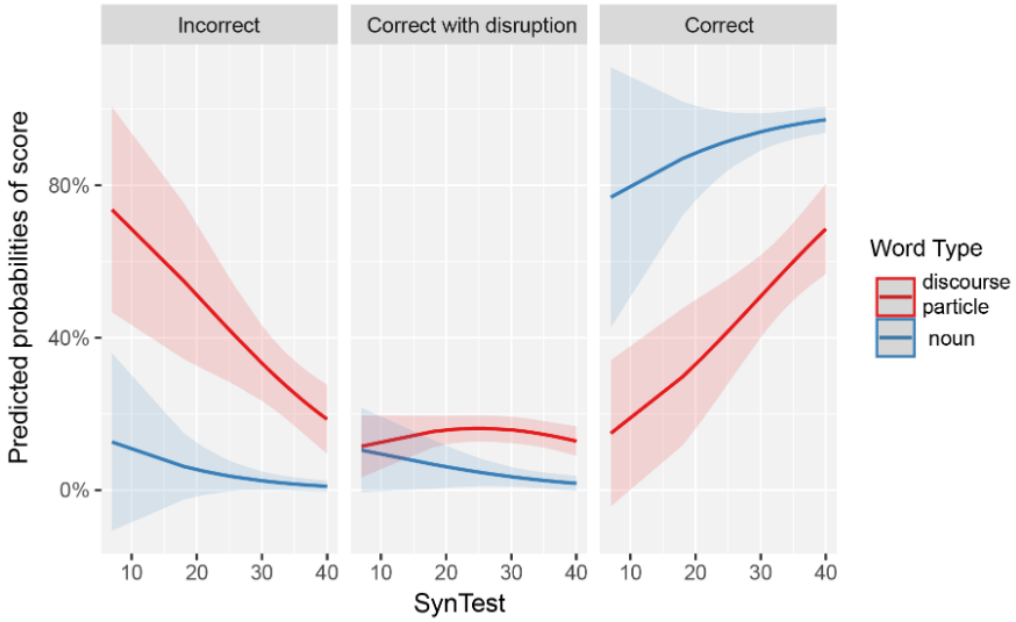


Figure 1. Predicted probabilities for each item score (i.e., incorrect, correct with disruption, correct) across SynTest scores, by Word Type (i.e., discourse particle, noun), including 95% confidence interval bands. SynTest scores: smaller numbers indicate worse performance. Estimates are given on the probability scale. Confidence intervals are calculated using normal theory, yielding confidence intervals beyond the bounds [0,1], indicating less certainty around those values.

A German Test for Thematic and Taxonomic Relations in Awake Brain Surgery: Standardization and Validation of the GB-WOBA

Manon Winkler | Madeline Philipsen | Roel Jonkers | Thomas Picht | Adrià Rofes

Introduction and Aims

Individuals with a brain tumor may experience language impairments after awake brain surgery, such as deficits in semantic processing (Collée et al., 2022) and comprehension (Ilmberger et al., 2008). While object naming is the current gold standard in intraoperative language mapping (Ilmberger et al., 2008), research indicates a behavioral and neural dissociation between nouns and verbs (Rofes et al., 2024). The possible postoperative deficits, along with this dissociation between word classes, demonstrate the need for a new test suitable for awake brain surgeries that includes both word classes and assesses comprehension and the semantic system in more depth to improve postoperative outcomes. Thematic (e.g., dog – leash, singing – dancing) and taxonomic (e.g., dog – bear, cutting – sawing) relations offer promising insights into the semantic system, as these relations show differences in terms of accuracy, reaction times (RT), and neural correlates (Mirman et al., 2017).

Building on this, the Groningen-Berlin Word-Picture-Association test ("*Groningen-Berlin Wort-Bild-Assoziationstest*"; GB-WOBA), including a noun and a verb task, was developed, standardized, and is soon to be validated in people with post-stroke aphasia and in individuals with a brain tumor. A preliminary version of the verb test was administered to eight people with a brain tumor after surgery. Individual differences in thematic and taxonomic processing were revealed in this group, whereas no such differences were found in healthy controls. This suggests that the test can capture important differences in semantic processing and supports the need to further develop, standardize, and validate it. The main aims of this study are (i) to standardize and validate the GB-WOBA and (ii) to assess differences in thematic and taxonomic relations among individuals with a brain tumor.

Methods

The GB-WOBA is designed to assess semantic association and comprehension. Participants first silently read a word displayed in the middle of the screen (i.e., the prime; e.g., leash or singing) and then tap one of four pictures that best match the word (i.e., the target; e.g., dog or dancing). The relation between the prime and the target is either thematic or taxonomic. The other three pictures are semantically related to the target (e.g., cat or jumping), phonologically related to the prime (e.g., leaf or sinking), and unrelated to the target and distractors (e.g., toaster or sawing).

The study followed a minimal standardization protocol (Rofes et al., 2015). For the standardization, 51 healthy native German speakers (26 female, 25 male; mean age = 48.63 years, SD = 16.93; mean years of education = 17.05, SD = 3.33) participated and were equally divided into five age groups (18–29, 30–41, 42–53, 54–65, 66+).

The test items were systematically matched on several psycholinguistic variables, including word frequency, imageability, age of acquisition, number of graphemes and syllables, Levenshtein distance, and semantic similarity, across thematic and taxonomic conditions.

Cut-off scores were calculated using SINGLIMS (Crawford & Garthwaite, 2002) by identifying the lowest score that did not yield a significant result relative to each age group. Differences in accuracy and RT between thematic and taxonomic relations in nouns and verbs were calculated using (generalized) linear mixed models.

The validation is soon to start. By the time of the conference, data are expected from 20 individuals with post-stroke aphasia, a preliminary sample of individuals with a brain tumor, and additional healthy controls. Convergent validity will be assessed by correlating the performance of individuals with post-stroke aphasia on the GB-WOBA with their performance on established tests of language comprehension and semantic association.

In individuals with a brain tumor, the GB-WOBA will be administered preoperatively with transcranial magnetic stimulation (TMS) and intraoperatively with direct electrical stimulation (DES). Performance patterns and stimulation sites will be compared with standard object and action naming. Differences between thematic and taxonomic relations will be analyzed at the group level and with single-case statistics to capture possible individual variability.

Results

The standardization indicates that age significantly influences accuracy and RT, whereas years of education do not affect test performance. These findings support considering age groups. Thematic and taxonomic conditions were matched ($p > .050$), with only minor differences observed in the noun condition for semantic similarity (prime – target $p = .003$, target – semantic distractor $p = .006$). The matched version of the test includes 46 noun and 46 verb items with 23 thematic and 23 taxonomic items each. Between thematic and taxonomic conditions, no significant differences in accuracy (nouns: $p = .740$, verbs: $p = .113$) and RT (nouns: $p = .377$, verbs: $p = .742$) were found in healthy participants during standardization.

Discussion

The GB-WOBA is a novel semantic association test in German that may complement existing language tasks and improve current assessment procedures during awake surgery and postoperative outcomes. It extends beyond standard naming by including different word classes and types of semantic relations, potentially allowing deeper insights into behavioral and neuroanatomical differences of thematic and taxonomic relations.

The non-significant differences in thematic and taxonomic relations among healthy participants during standardization may be ascribed to the matching of thematic and taxonomic conditions across seven psycholinguistic variables. The ongoing validation will determine the sensitivity of the GB-WOBA to individual differences in thematic and taxonomic processing among two neurological populations. We predict that validation analyses in individuals with post-stroke aphasia will show correlations between the GB-WOBA and established comprehension and semantic association tests. Furthermore, individuals with a brain tumor will perform worse than healthy controls on the GB-WOBA. Differences between thematic and taxonomic relations may be more evident at the individual level (i.e., single-case statistics) than at the group level. Pilot data of a preliminary version of the test suggest that some individuals with a brain tumor demonstrate selective impairments in

thematic or taxonomic relations, or in both. Therefore, the GB-WOBA may be appropriate for improving postoperative outcomes.

References

Collée, E., Vincent, A., Dirven, C., & Satoer, D. (2022). Speech and Language Errors during Awake Brain Surgery and Postoperative Language Outcome in Glioma Patients: A Systematic Review. *Cancers*, *14*(21), Article 21. https://doi.org/10.3390/cancers14215466

Crawford, J. R., & Garthwaite, P. H. (2002). Investigation of the single case in neuropsychology: Confidence limits on the abnormality of test scores and test score differences. *Neuropsychologia*, *40*(8), 1196–1208. https://doi.org/10.1016/S0028-3932(01)00224-X

Ilmberger, J., Ruge, M., Kreth, F.-W., Briegel, J., Reulen, H.-J., & Tonn, J.-C. (2008). Intraoperative mapping of language functions: A longitudinal neurolinguistic analysis: Clinical article. *Journal of Neurosurgery*, *109*(4), 583–592. https://doi.org/10.3171/JNS/2008/109/10/0583

Mirman, D., Landrigan, J.-F., & Britt, A. E. (2017). Taxonomic and thematic semantic systems. *Psychological Bulletin*, *143*(5), 499–520. https://doi.org/10.1037/bul0000092

Rofes, A., D'Agata, F., Guerrini, F., & Spena, G. (2024). Speech and lexico-semantic errors during direct cortical stimulation mapping of the language-dominant hemisphere: Effects of object and action naming. *Journal of Neurosurgery*, *140*(6), 1641–1659. https://doi.org/10.3171/2023.10.JNS231493

Rofes, A., de Aguiar, V., & Miceli, G. (2015). A minimal standardization setting for language mapping tests: An Italian example. *Neurological Sciences*, *36*(7), 1113–1119. https://doi.org/10.1007/s10072-015-2192-3

A usage-based approach to language in Alzheimer's disease: the effect of multiword-frequency variables on patterns of impairment

Vitor Zimmerer | Natasha Adamczyk | Dicky Lim | Chui Thing Kiew | Rinn Wang | Fiona Yang | Sebastian Bello Lepe | Rosemary A. Varley

Background

Decline in language capacity is a core feature of Alzheimer's disease (AD). Identification of language change in dementia serves the design of clinical tools for detection and tracking and helps test linguistic theory (Kavé & Goral, 2018).

Many language studies carried out on dementia are informed by generativist "word and rules" approaches, which focus on properties of words and formal grammatical structure. Usage-based approaches on the other hand consider more strongly the effects of linguistic experience on language organisation, and challenge the distinction between lexicon and grammar. For instance, previous work found speakers with amnestic AD (Zimmerer et al., 2016) and other dementias (Zimmerer et al., 2020) produced stronger word collocations (specific combinations of words that appear together often) than controls, suggesting that strong collocations are affected less by impairment and may be processed as holistic language formulas (Bridges & Van Lancker Sidtis, 2013). Such approaches can capture additional dimensions of language production which respond to cognitive change.

In the current study, we compared usage variables in AD with more conventional measurements of grammatical complexity, to determine the contribution of each approach, and investigated the effect of collocation strength on speech pauses as a measure of online cognitive effort.

Methods

We investigated "Cookie Theft" picture descriptions of 18 people with a diagnosis of AD (mean age = 74.7; mean MoCA score = 13.7) and 21 healthy controls (mean age = 67.7; mean MoCA score = 28). All participants were English speakers. We used computerized methods to determine word frequency of each word, and collocation strength of each word combination (bigrams), using large-scale English corpora as reference. We also applied the Northwestern Narrative Language Analysis (Thompson, 2013) to determine grammatical complexity of samples. For pause duration, we measured the duration of silence or fillers (e.g. um) preceding each individual word.

Results

Higher MoCA scores (better cognitive capacity) were associated with lower average content word frequency, $t(37) = -2.51$, $p = .017$ and lower average bigram collocation strength, $t(37) = -3.573$, $p = .001$, indicating less reliance on frequent (or formulaic) language forms. When collocation strength was entered into models, adding grammatical complexity did not improve model fit.

A hurdle model built around lexical frequency found it had a significant effect ($p < .001$), full model $R2 = .015$. A hurdle model based on collocation strength found it had a significant effect ($p < .001$), full model $R2 = .037$. Pauses were longer before less common words and

within weaker collocations; however, collocation strength served as a better predictor of pauses.

Conclusions

Findings show that language processing efficiency is dependent on patterns of word co-occurrence. Collocation strength of word combinations can serve as an important predictor of language impairment in AD, and may be more sensitive to disorder than grammatical complexity. It also serves as a predictor for pause duration as an indicator of cognitive effort during production. The work supports the value of experience-based measures for characterising and tracking language impairment.

References

Bridges, K. A., & Van Lancker Sidtis, D. (2013). Formulaic language in Alzheimer's disease. *Aphasiology*, *27*(7), 799-810.

Kavé, G., & Goral, M. (2018). Word retrieval in connected speech in Alzheimer's disease: a review with meta-analyses. *Aphasiology*, *32*(1), 4-26.

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, 53(4), 695–699.

Thompson, C. K. (2013). *Northwestern narrative language analysis (NNLA) theory and methodology*. Aphasia and Neurolinguistics Research Laboratory Northwestern University, Evanston, IL.

Zimmerer, V. C., Hardy, C. J., Eastman, J., Dutta, S., Varnet, L., Bond, R. L., ... & Varley, R. A. (2020). Automated profiling of spontaneous speech in primary progressive aphasia and behavioral-variant frontotemporal dementia: An approach based on usage-frequency. *Cortex*, 133, 103-119.

Zimmerer, V. C., Wibrow, M., & Varley, R. A. (2016). Formulaic language in people with probable Alzheimer's disease: A frequency-based approach. *Journal of Alzheimer's Disease*, *53*(3), 1145-1160.

Email addresses first or corresponding authors

- sophie.arheix-parras@sc.edu,
- Karine Marcotte <karine.marcotte@umontreal.ca>,
- zannatnoor63@gmail.com,
- Macarena Martínez Cuitiño <mariamacarenamartinez@gmail.com>,
- Yuval Zecharia Katz <yuvalzkatz@gmail.com>,
- amelie.sanfacon-verret.1@ulaval.ca,
- silvia.marchesotti@unige.ch,
- d.j.garcia.inclan@student.rug.nl,
- mscimeca@bu.edu,
- v.zimmerer@ucl.ac.uk,
- michal.laznicka@ff.cuni.cz,
- Alycia Laks <alycia.laks@georgetown.edu>,
- claudia_penaloza@ub.edu,
- Vasiliki Koukoulioti <vasilikikou@del.auth.gr>,
- Mauro Viganò <dott.mauro.vigano@gmail.com>,
- Dörte de Kok <d.a.de.kok@rug.nl>,
- m.c.winkler@rug.nl,
- m.m.f.philipsen@rug.nl,
- papakyritsis <papakyritsis@upatras.gr>,
- vasarivivien@gmail.com,
- riyad.abdullah.11339@gmail.com,
- solene.hameau@uclouvain.be,
- CASSIADY.PERARD66@login.cuny.edu,
- sdemarchi@gc.cuny.edu,
- ana.matic@erf.unizg.hr,
- zcjtise@ucl.ac.uk,
- Georgia Roumpea <georoubea@gmail.com>,
- irina.chupina@donders.ru.nl,
- l.rybka@rug.nl,
- f.tsiwah@rug.nl,
- anaelle.filisetti@univ-lille.fr,
- d.satoer@erasmusmc.nl,
- vitoria.piai@donders.ru.nl,
- ana.estrada.pineda@udc.es,
- lotta.heidemann@uni-bielefeld.de,
- jean-marie.annoni@unifr.ch,
- greg.python.logo@gmail.com,
- j.h.van.der.sluijs@rug.nl,
- Rinn Wang <rinn.wang@ucl.ac.uk>,
- sophotsagareishvili@gmail.com,
- kalditamas@gmail.com,
- io.salmons@uab.cat,
- "Onofri, Ludovica" <ludovica.onofri.19@ucl.ac.uk>,
- johanna.scholtz@evkb.de,
- Laure Salpéteur <laure.salpeteur@uclouvain.be>,

- mareike.hartmann@uni-bielefeld.de,
- Caroline.rogue@unige.ch,
- marina.ruiter@ru.nl,
- imke.wets@ru.nl,
- anais.dacostamartinez@unige.ch
- Strinzel, Anna-Maria <anna-maria.strinzel@uni-bielefeld.de>
- nicole.gerlak@lc.cuny.edu
- jean-marie.annoni@unifr.ch (jean-marie.annoni@unifr.ch)

Symposium 1

Living well with aphasia: research and topics priorities identified by stroke survivors and aphasia associations

Introduction: Cocreation of Aphasia Topics Priorities.

Jean Marie Annoni [1, 2]
[1]Department of Neurosciences, University of Fribourg, Chemin du Musée 8, 1700 Fribourg, Switzerland; jean-marie.annoni@unifr.ch, [2]Aphasia International Association (AIA) Maison Des Associations Internationales Rue Washington 40 1050 Bruxelles Belgique Start - Aphasie International

Aphasia is a condition that makes it challenging for someone to speak, understand, read, or write, most often after a stroke. But the impact goes far beyond the language impairment. People with aphasia (PWA) often struggle with loneliness, emotional difficulties, and being misunderstood by their close environment. The aim of this symposium is to coin the priorities suggessed by PWA in life, research, and care.
In the introduction we will present an international study promoted by AIA[1] and LIHA[2] with nearly 100 people living with long-term aphasia and health professionals from 15 countries explored what really matters to them.
The study took three steps: first, online meetings gathered ideas; then, small group sessions in each country discussed and ranked them; finally, everyone voted across countries to choose the top priorities. To ensure everyone took part, the team used aphasia-friendly materials, live and written translations, and simple communication tools. Despite challenges of working across different languages, participation was high, with most people staying involved throughout.
In the end, four key priorities were chosen:

(1) Raising awareness - so families, friends, and society understand aphasia better.
(2) Emotional challenges - such as changes in relationships and loss of closeness.
(3) Rebuilding confidence - after the life-changing effects of aphasia.
(4) Improving therapy - especially how hospitals and services support treatment

This study shows that with the right support, people with aphasia can speak up about what matters the most. Their voices highlight not just medical needs but the full picture of living with aphasia, guiding future care, research, and policies worldwide (see Figure 1). The paper is in minor revisions for publication.
To illustrate and develop this observation, three PWA will present different aspects of their implication in aphasia expertise, research, and dissemination of informations, and possible collaboration with scientists.

---

[1] **Association Internationale Aphasia** (AIA) are a non-profit international organisation for persons with an acquired communication and language disorder caused by brain damage and their families and caregivers.
[2] **Liga Hispanoamericana de Afasia**, created by Prof Silvia Rubio from Argentina in 20219

Presentation 1 : Aphasique et heureuse (Aphasic and happy)

Margot Turcat[1]

[1] Margot Turcat, Bordeaux, GARB (Groupe des Aphasiques de la Région de Bordeaux), enseignante, autrice de *Mon Petit AVC* et *Ça Me prend la tête !* (Éditions Larousse), patiente experte, patiente formatrice et créatrice du compte Instagram « @mon.petit.avc » mon.petit.avc@gmail.com

My talk will focus on the idea of being 'aphasic and happy'. It will be presented partly in English and partly in French with translation.

I am now 40 years old, a mother to an eight-year-old boy, a qualified art teacher and an author. However, almost eight years ago, my very orderly and structured life was shattered when I had a stroke. After an initial disastrous response, events unfolded rapidly, and unsurprisingly, I discovered what I had suspected from the outset: I was having an ischaemic stroke.

I came to terms with this very quickly and realised that I was entering the world of disability. In order to explain this to my son, who was just 14 months old at the time, I decided to tell him my story. Unable to write properly, I turned to the pencils available at the rehabilitation centre and began to draw my experiences for him. At the same time, I created an Instagram account, @mon.petit.avc, to keep my family, friends and colleagues updated on my new daily life.

Spotted by Larousse Publishing at the end of 2020, this collaboration resulted in two comic books: Mon Petit AVC, released in 2021, and Ça me prend la tête! in 2025. These books are now featured in numerous health training bibliographies, candidly recounting the daily life of someone who suffered a stroke at a young age. The first book recounts the 'hard' part, from the day of the stroke to leaving the rehabilitation centre six months later. The second tackles the issue of daily life when you return home and everything has changed. Everything then needs to be rebuilt: personal life, family life, social life and professional life. The books are written in a realistic tone and feature wonderful human encounters, as well as good times and more delicate situations.

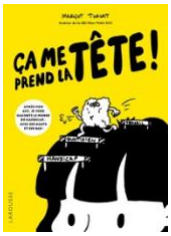
My aim is to raise awareness of this condition and help as many victims as possible express their experiences.

Original French : Ma présentation portera sur être « aphasique et heureuse ».

J'ai aujourd'hui 40 ans, je suis maman d'un petit garçon de 8 ans, professeur certifiée d'arts plastiques et autrice mais il y a bientôt 8 ans, ma vie très rangée et cadrée a volé en éclats quand l'AVC a décidé de s'inviter dans mon quotidien... Après une prise en charge catastrophique, tout s'enchaîne et je découvre sans surprise ce que je soupçonnais depuis le début, j'étais entrain d'avoir un AVC ischémique.

J'encaisse très vite la nouvelle et comprends que désormais, je tombe dans le monde du handicap. Afin d'expliquer tout cela à mon fils qui était alors un bébé de 14 mois, je décide de lui raconter mon histoire. Incapable d'écrire correctement, je me tourne alors vers les crayons que j'ai à disposition en centre de rééducation et je commence à lui dessiner mon vécu. En parallèle, je décide de poster sur cela sur un compte Instagram créé pour l'occasion @mon.petit.avc qui était initialement à destination de ma famille, mes amis et mes collègues afin de les informer de mon nouveau quotidien.

Repérée à la fin de l'année 2020, par les éditions Larousse, deux BD sont nées de cette collaboration : Mon Petit AVC, sorti en 2021 et Ça me prend la tête ! En 2025. Ces deux livres sont aujourd'hui référencés dans de nombreuses bibliographies de formations en santé car ils racontent sans fard le quotidien d'une personne atteinte par un AVC dans son jeune âge. Le 1er ouvrage raconte le « dur », du jour J à la sortie du centre de rééducation, six mois plus tard. Le second aborde la question de la vie quotidienne quand on rentre chez soi et que rien n'est plus pareil. Tout est alors à reconstruire : vie personnelle, familiale, amicale et professionnelle. Le tout dans un ton réaliste avec de très belles rencontres humaines, de bons moments mais aussi des situations plus délicates.



Mon but est de vulgariser cette pathologie et d'aider un maximum de victimes afin qu'ils puissent poser des mots sur leurs maux.

Presentation 2: Contributing to Research Through Lived Experience

Author Alexia Kountouri [1]

[1] President AIA, Aphasia International Association (AIA) Maison Des Associations Internationales Rue Washington 40 1050 Bruxelles Belgique; Stroke Ambassador, Cyprus Stroke Association, Limassol, Cyprus; kountourialexia19@gmail.com

Introduction: Alexia experienced a stroke in 2015 that affected the right side of her body. Like many survivors, she faced significant physical, emotional, and social challenges. Rather than allowing these obstacles to silence her, she chose to transform her lived experience into a meaningful contribution to research.

For Alexia, having a stroke initially felt like losing her voice. Participating in research gave that voice back. Through her involvement, she brought researchers the true value of stroke survivors' experiences and perspectives.

Why She Became Involved in Research: Alexia chose to engage in research because she wanted her experience to matter. She believed that research should also reflect the real-life challenges faced by stroke survivors and that those who live with these realities should have a role in shaping research questions and outcomes. Her motivation was to ensure that the voices of stroke survivors—including those living with aphasia—are heard and represented.

Overcoming Challenges: One of the main difficulties Alexia faced was communication. Aphasia created barriers that sometimes made it harder to process information or express her thoughts clearly. Despite these challenges, she remained determined to contribute.

At the beginning, she had concerns about whether her aphasia would prevent her from contributing effectively. She wondered if others would understand her and whether her input would truly matter. Nevertheless, she chose to move forward and take a leap of faith.

Shaping the Research Process: As a research partner, Alexia actively collaborated with researchers to refine the study's research questions. Her input helped ensure that these questions were understandable and accessible to stroke survivors, including those living with aphasia. She also provided insights that made the research more inclusive and better represented the realities faced by survivors.

Learning about research methodology required time and effort, as she became more familiar with the research process, her confidence grew, and she was able to contribute more actively to discussions and decisions.

Working Through Communication Barriers: Because of aphasia, Alexia sometimes needed extra time to process information and express her thoughts. She learned to ask for support when necessary, and the willingness of the team to accommodate these needs made a significant difference.

Making an Impact: Alexia's contributions helped ensure that the research remained grounded in the real experiences of stroke survivors. Her perspective helped researchers better understand the practical and emotional challenges that survivors face in everyday life.

Although her stroke changed many aspects of her life, it did not take away her strength or determination. She continues to fight, grow, and redefine what is possible.

Advice for Other Research Partners: Based on her experience, Alexia encourages other patient partners to advocate for themselves during the research process. She believes it is important to speak up when support is needed—for example by asking for breaks, requesting additional time, or asking for clarification. Being honest about personal challenges helps create a more inclusive and supportive research environment.

Presentation 3: "How I turned my disability into my purpose. My experience in improving aphasia awareness."

Hanka Mayhew 1

[1]Say Aphasia and AIA ambassador, UK (I have aphasia), 9 Gatton Park Lane, Brighton, Bn1 5BQ; and AIA T: +44 7977 321141 https://www.sayaphasia.org

Hanka Mayhew lived in Spain for 11 years with her husband and three children, speaking five languages and working as an English teacher and translator until a massive brain bleed left her with aphasia and unable to speak. After three months in a hospital in Spain and three more months in England, she spent a year in a wheelchair relearning how to live. Hanka Mayhew faced severe consequences after her stroke eleven years ago. As a teacher of English for foreign students and a translator, she struggled immensely with speech, comprehension, and writing difficulties. Her principle was to stay engaged in activities that distracted her from the effects of the illness. Initially, horticulture played a crucial role in her recovery. Through regular physical exercise, psychological and speech therapy and participation in aphasia support groups she made steady progress. Today her speech is nearly flawless, she travels and she is independent in daily life. The loss of her husband three years ago was another devastating blow. She explains that English and German are easiest, Spanish is slowly returning with little practice, and she can understand some French and Russian but cannot speak them; she retained reading ability but reads more slowly. In February 2024, she joined the Aphasia International Association and became a Say Aphasia ambassador, traveling to assemblies, meetings and conferences across Europe and giving talks alone or with clinicians/ laypersons support to raise awareness and build an international aphasia

community. Say Aphasia (in England) and AIA ambassador she travels around Europe to raise awareness of aphasia. Her first presentation was at European Life After Stroke Forum in Prague on 10th and 11th March last year. She practiced her talk 30 times or more before the conference but could convince the participants of the importance to develop awareness of aphasia in the community. Her children say “Mummy, you are still a teacher but not in English, but in Raising Awareness of Aphasia.”

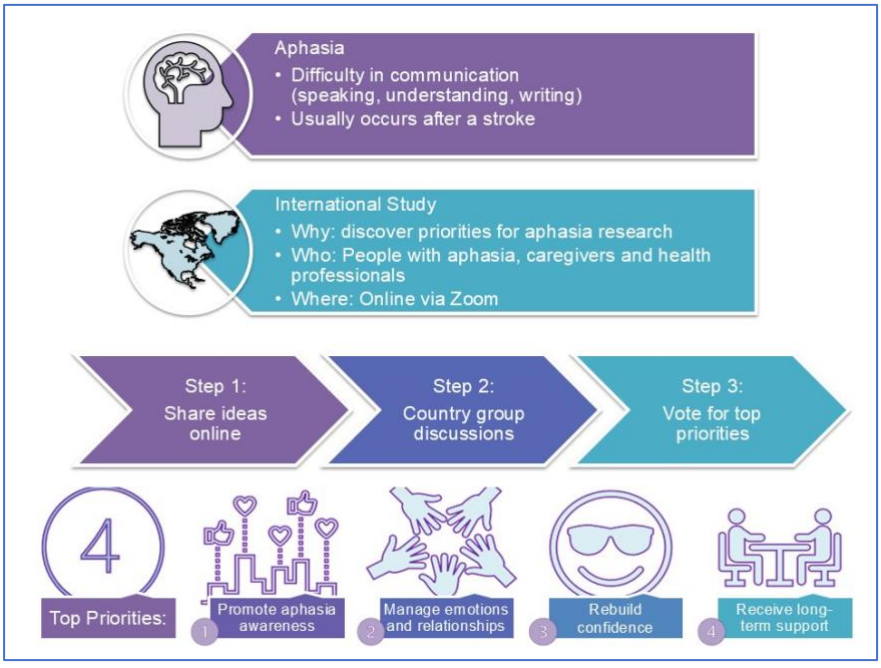
References

Charalambous M, Kountouri A, Phylactou P, Triantafyllidou I, Annoni JM, Kambanaros M. The views of people living with chronic stroke and aphasia on their potential involvement as research partners: a thematic analysis. *Res Involv Engagem. 2022 Sep 5;8(1):48. doi: 10.1186/s40900-022-00379-1. PMID: 36064623; PMCID: PMC9446531.*

Charalambous M, Kountouri A, Schwyter JR, Annoni JM, Kambanaros M. The development of the People with Aphasia and Other Layperson Involvement (PAOLI) framework for guiding patient and public involvement (PPI) in aphasia research. *Res Involv Engagem. 2023 Sep 1;9(1):74. doi: 10.1186/s40900-023-00484-9. PMID: 37658465; PMCID: PMC10474738*.

Jean Marie Annoni Marina Charalambous Jernej Sluga Clotilde Marteel Javier Gil Reto Strähler Fanni Eckhardt Colin Lyall Davide Crovetti Hanka Mayhew Alexia Kountouri Claire Bennington Silvia Rubio. A Global Multilingual Co-Creation of Aphasia Topics Priorities through Patient and Public Involvement. *Research Involvement and Engagement In revision*

Fig 1 : Visual lay summary of the study presented in introduction



Science of Aphasia

Symposium submitted to

Science of Aphasia 2026 (22-25 septembre) Bordeaux (France)

Rethinking Assessment in Aphasia: Discourse Ecosystems, Interpretable Metrics, and Clinical Relevance

Maria Varkanitsa[1], Brielle C. Stark[2], Haris Themistocleous[3]

[1]*Boston University*
[2]*Indiana University Bloomington*
[3]*University of Oslo*

Introduction

Discourse is widely recognized as one of the most ecologically valid windows into communication, but discourse assessment is still too often constraint by legacy task choices, labor-intensive manual analysis, and metrics whose clinical meaning is not always clear. At the same time, computational approaches are advancing rapidly, creating new opportunities to analyze language at scale, but also new challenges regarding interpretability, task validity, and clinical usefulness.

This symposium argues that the field is at a pivotal moment. To move aphasia assessment forward, we need approaches that are simultaneously ecologically grounded, computationally rigorous, and clinically interpretable. The three talks in this symposium address that need from complementary perspectives. The first reframes discourse not as a single elicitation product, but as the outcome of a broader discourse ecosystem shaped by speaker abilities, partner behavior, genre, goals, and context. The second asks what computational metrics should mean if they are to become clinically actionable, proposing information-theoretic indices that quantify predictive dysregulation and communicative strain rather than simply tallying linguistic errors. The third provides an empirical demonstration that clinically meaningful macrolinguistic features, including main concepts, coherence, sequencing, and semantic distance from the narrative core, can be extracted automatically and used to distinguish aphasia from healthy aging and related clinical populations.

Together, these talks chart an exciting path for the next generation of aphasia assessment. Rather than treating communication as a collection of isolated tasks or error counts, they point toward a more powerful framework that is ecologically grounded, computationally interpretable, and clinically meaningful. In doing so, this symposium highlights not just new methods, but a broader shift in the field: toward assessment approaches that can capture communication as it is truly lived, experienced, and supported in aphasia.

Science of Aphasia

Title presentation 1

Brave New World: Discourse Ecosystems and a Better Science of Discourse Assessment

Brielle C. Stark, Indiana University Bloomington [1]

Abstract (Max 500 words)

Discourse is the outcome that most closely reflects everyday communication in aphasia and primary progressive aphasia (PPA)—yet discourse assessment is still too often treated as a "task" rather than a complex, context-sensitive phenomenon. This talk argues for a big-picture reframing: we should stop treating discourse samples as interchangeable elicitation products and instead view them as outputs of *discourse ecosystems*—systems in which speaker abilities, partner behaviors, genre demands, goals, and environmental supports jointly shape what communication looks like. In both acquired aphasia and PPA, variability across tasks and settings is not simply measurement noise; it is a signature feature of discourse. Our science will progress fastest when assessment practices are designed to model that adaptiveness, rather than unintentionally suppress it through defaults, legacy methods, or one-size-fits-all metrics.

Building from the field's classic discourse tasks (e.g., picture descriptions, story retells, procedural monologues), I will highlight why these paradigms have been so productive—offering feasibility, comparability, and scalable sampling—while also naming their limitations: task choices can quietly define constructs, genre constraints can reshape linguistic organization, and familiar metrics can be treated as portable even when their meaning is task-dependent. The result is a recurring disconnect between what we *measure* in discourse and what we often *intend* discourse to represent: functional communication across contexts and meaningful change over time. Rather than abandoning foundational tasks, this presentation proposes using them as anchors within a broader, more explicit ecosystem approach—one that links task design, measurement, and interpretation to purpose.

To make this shift actionable, I will offer nine key talking points that together define a near-term roadmap for improving the "science of discourse science":

1. **Define the Purpose**: state what the task is meant to measure and why—avoid defaulting to familiar elicitation without rationale.
2. **Match Task to Population**: select/design tasks that fit the communication profiles and trajectories of the target group (e.g., post-stroke aphasia vs PPA).
3. **Use Meaningful Metrics**: choose outcomes that are sensitive, decision-relevant, and supported by the task's structure; metrics only matter if the task can sustain them.
4. **Ensure Psychometric Quality**: evaluate reliability, validity, and responsiveness to build trust and enable clinical-trial readiness.
5. **Support Replication and Sharing**: document instructions, scoring, and procedures; share materials to enable cumulative, multi-site progress.
6. **Use Technology Thoughtfully**: automation can accelerate analysis but cannot rescue weak task design; technology-focused advances (e.g., OpenBrainAI) are addressed elsewhere in the symposium.

7. **Choose the Right Genre**: select genres (narrative, procedural, conversation) based on goals because each foregrounds different abilities.
8. **Engage End Users Early**: partner with clinicians and people with lived experience so tasks are meaningful, usable, and ecologically grounded.
9. **Be Transparent About Limitations**: specify what tasks and metrics can—and cannot—conclude to strengthen interpretation and field credibility.

Together, these points position discourse assessment not as a single method, but as an ecosystem-informed enterprise—one that can better align discourse science with real communication in aphasia and PPA. [Figure 1]

References

Author, X. Y. (2012). Title of paper. *Title of Journal*, *1*, 1-2.

Figure or Table *(optional, max 3)*

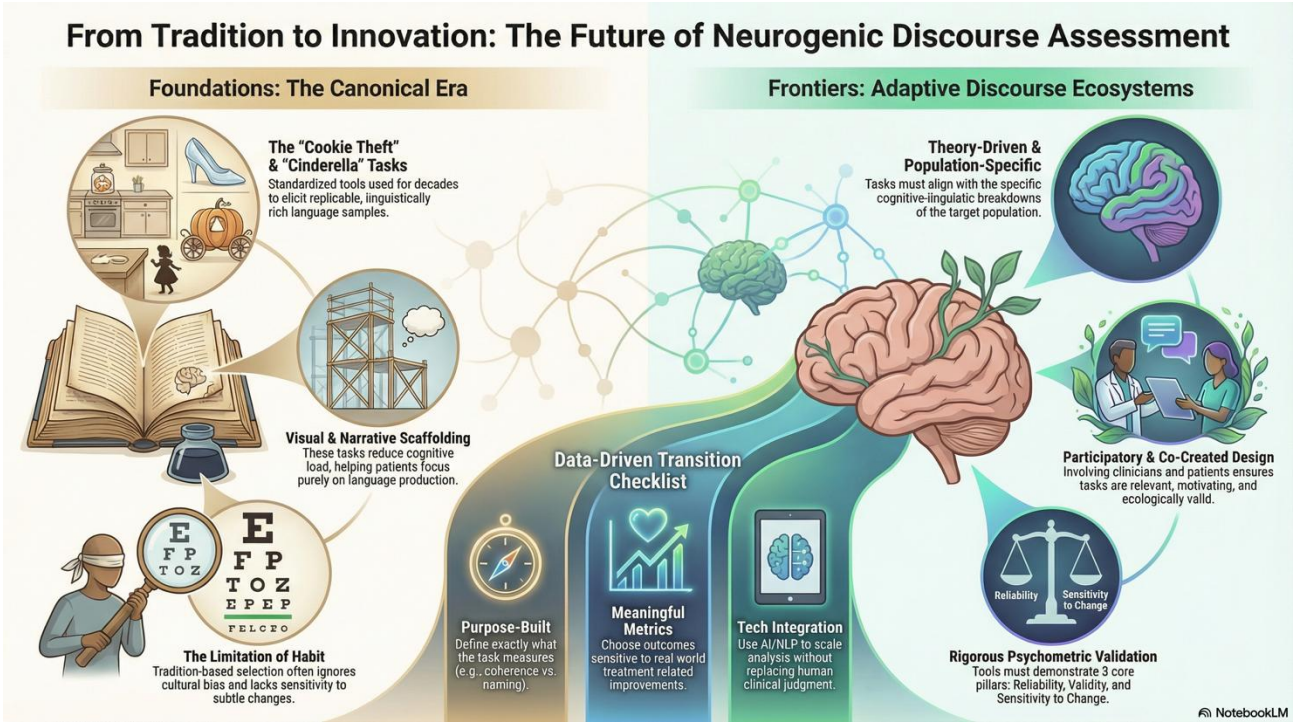


Figure 1: Infographic of how author BC Stark envisions moving toward discourse ecosystem science; developed by partnering with NotebookLM AI.

Science of Aphasia

Title presentation 2

Beyond Speech Errors: Quantifying the Communicative Channel in Aphasia

Haris Themistocleou, University of Oslo

Abstract (Max 500 words)

Computational Language Assessment (CLA) holds immense promise for automating aphasia diagnosis, tracking disease progression, and estimating treatment efficacy. However, a fundamental methodological hurdle remains: *the interpretability problem*. Currently, a plethora of computational measures is being employed to explain the neurocognitive foundations of language impairment. This talk argues that to make CLA clinically actionable, we must move beyond merely cataloging linguistic errors or comparing surface-level deviations against normative populations. Instead, we must employ reconceptualize aphasic language output as the behavioral signature of a disrupted neurocognitive closed-loop system.

We advance the idea that the brain operates by constantly generating internal predictions and testing them via closed-loop feedback control—an "inside-out" approach[1]. In aphasia, this predictive mechanism breaks down; the inability to generate stable, redundant internal models collapses the forward-looking pre-calculation necessary for fluent production. This intrinsic breakdown of the language-specific sensorimotor architecture—largely independent of general cognitive ability[2] —requires new tools for measurement. To quantify this internal predictive dysregulation via external output ("outside-in"), we propose the use of Large Language Model (LLM)-derived information-theoretic metrics. These metrics offer an interpretable window into the functional equilibrium of human communication.

Healthy speakers naturally balance the Principle of Least Effort with linguistic redundancy, providing predictive cues that narrow the cognitive search space (lowering Entropy) for the listener. Conversely, non-fluent aphasia is characterized by maladaptive morphosyntactic "over-economization." By omitting highly probable grammatical markers, patients strip away essential redundancy. This forces the listener's predictive system into a state of chronic high Surprisal, maximizing communicative strain. The Surprisal-Entropy Ratio specifically formalizes this equilibrium, capturing pathological spikes in unexpected linguistic units exactly where the functional pressure for redundancy is highest.

To demonstrate the approach, in the talk we provide examples from information-theoretic metrics, showing that they correlate strongly with the Aphasia Quotient, and validate their clinical utility. Ultimately, these indices solve the interpretability problem by modeling the interpersonal reality of the communicative exchange. In Shannon (1948)'s terminology[3], the loss of redundant morphosyntactic cues acts as "internal noise," severely reducing the Channel Capacity of the interaction. Information-theoretic metrics are not just passive quantifications of a patient's speech; they are direct, objective measures balancing performance and communicative bandwidth. To make this computational shift actionable, I will offer four key principles that define a near-term roadmap for implementing information-theoretic metrics in aphasia assessment:

1. Transition from tallying static linguistic errors to measuring dynamic predictive dysregulation and the actual communicative strain placed on the listener.
2. Utilize the S/E Ratio to isolate pathological unpredictability, evaluating a patient's errors strictly within the mathematical constraints of their surrounding narrative context.
3. Document LLM architectures, tokenization methods, and probability extraction pipelines to guarantee metric reproducibility and trust across multi-site clinical trials.
4. Align automated metrics with the clinically informed task designs (as discussed elsewhere in this symposium) to ensure CLA enhances, rather than replaces, the clinician's expert perspective.

Together, these points position Computational Language Assessment not merely as an automated scoring mechanism, but as a theoretically grounded, ecologically valid paradigm for understanding the predictive brain in aphasia.

References

[1]Buzsáki, G. (2019). *The brain from inside out*. Oxford University Press.
[2]Hickok, G. (2025). *Wired for Words. The neural architecture of language*. MIT Press.
[3]Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, *27*(3), 379-423. https://doi.org/https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Science of Aphasia

Title presentation 3

Automated Macrolinguistic Discourse Analysis for Transdiagnostic Detection of Language Impairment

Maria Varkanitsa[1], Seung Hee Lee[1], Sharon Wang[1], Swathi Kiran[1]

[1]Boston University Center for Brain Recovery

Abstract (Max 500 words)

As the field moves toward more naturalistic language assessment, discourse has emerged as one of the most informative windows into real-world communication. Yet much of what matters clinically in discourse is still difficult to measure at scale. In particular, macrolinguistic abilities such as conveying the central message, maintaining coherence, and preserving narrative structure are often disrupted across neurological populations, but they remain far less accessible to large-scale analysis than microlinguistic features because they typically rely on labor-intensive manual coding. In this study, we developed an automated natural language processing pipeline to extract macrolinguistic discourse features from narrative speech and evaluated their utility for distinguishing healthy controls (HC), people with dementia (PWD), and people with aphasia (PWA).

We analyzed Cinderella story retellings from two publicly available English-language discourse corpora, AphasiaBank[1] and DementiaBank[2], including HC (n=113), PWA (n=102), and PWD (n=94). Audio recordings were automatically transcribed using Amazon Transcribe and segmented into utterances based on terminators and conjunctions. We then generated semantic embeddings for participant utterances and for a reference inventory of Cinderella main concepts[3] using the all-mpnet-base-v2 Sentence-Transformers model[4]. Main concept embeddings were averaged to derive a centroid representing the semantic core of the narrative; utterances falling within one standard deviation of this centroid were classified as main concepts and matched to the closest target concept. Building on this framework, we extracted 23 automated macrolinguistic features indexing distance to centroid, main concept production, coherence, and sequencing. We then evaluated the diagnostic utility of these features using three-way multinomial logistic regression across HC, PWA, and PWD, together with one-vs-one pairwise classifications, and identified the most discriminative features based on model coefficients.

Relative to human-determined labels, the automated system achieved strong accuracy (84%) in identifying main concept utterances, supporting the validity of the automated MCA pipeline. Using a reduced set of 10 features after removing highly correlated variables ($|r| > .8$), logistic regression achieved a macro-averaged ROC AUC of 0.725 for the three-way classification. Pairwise performance was strongest for PWA versus HC (mean AUC = 0.938), more moderate for PWD versus HC (mean AUC = 0.661), and poor for PWD versus PWA (mean AUC = 0.575). In the three-way model, HC were classified most accurately, followed by PWA and then PWD, with PWD misclassified roughly equally as HC and PWA, suggesting that dementia occupies an intermediate position in macrolinguistic performance relative to the other groups. Distance to centroid increased progressively from HC to PWD to PWA, indicating decreasing semantic alignment to the narrative core and reduced narrative informativeness across groups. Across models, distance to centroid emerged as the strongest predictor, suggesting that greater semantic deviation from the narrative core is a consistent and interpretable marker of discourse-level impairment.

Fully automated macrolinguistic discourse analysis is feasible and clinically informative. The automated pipeline captured meaningful group differences and showed particularly strong utility for distinguishing aphasia from healthy language performance, while also highlighting discourse features that vary across clinical populations. These findings support automated discourse analysis as a scalable and ecologically valid approach to language assessment.

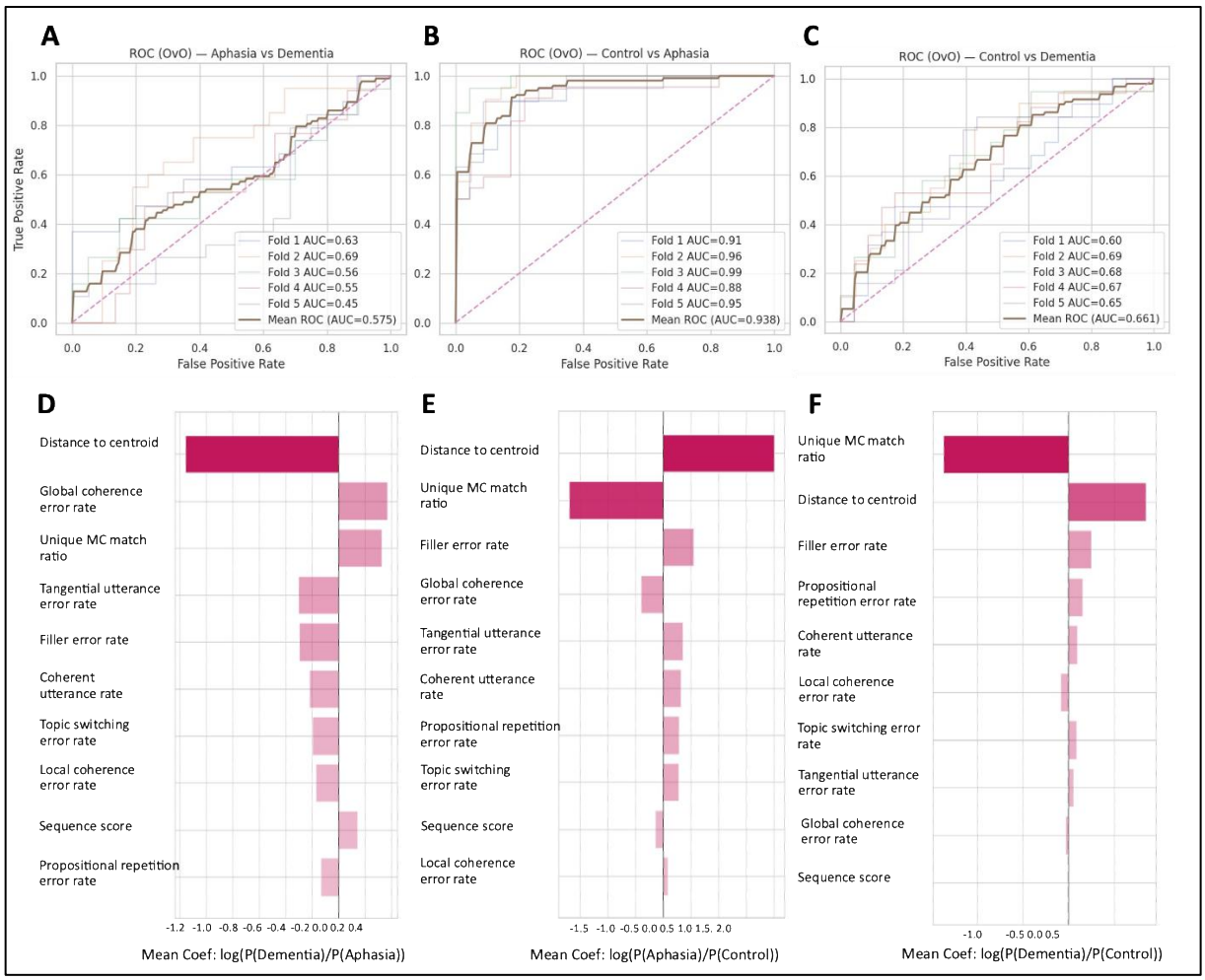


Figure 2. ROC curves (top) and corresponding feature importance plots (bottom) for pairwise logistic regression models classifying aphasia vs dementia (A/D), control vs aphasia (B/E), and control vs dementia (C/F) using automated macrolinguistic discourse features.

References

[1] MacWhinney, B., Fromm, D., Forbes, M., & Holland, A. (2011). AphasiaBank: Methods for studying discourse. *Aphasiology, 25*(11), 1286–1307. https://doi.org/10.1080/02687038.2011.589893

[2] Lanzi, A. M., Saylor, A. K., Fromm, D., Liu, H., MacWhinney, B., & Cohen, M. L. (2023). DementiaBank: Theoretical rationale, protocol, and illustrative analyses. *American Journal of Speech-Language Pathology, 32*(2), 426–438. https://doi.org/10.1044/2022_AJSLP-22-00281

[3] Richardson, J. D., & Dalton, S. G. (2016). Main concepts for three different discourse tasks in a large non-clinical sample. *Aphasiology, 30*(1), 45–73. https://doi.org/10.1080/02687038.2015.1057891

[4] Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence embeddings using Siamese BERT-networks *(arXiv:1908.10084)*. *arXiv*. https://doi.org/10.48550/arXiv.1908.10084