Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Research Report

CompreTAP: Feasibility and reliability of a new language comprehension mapping task via preoperative navigated transcranial magnetic stimulation

Leonie Kram [a,c], *Ann-Katrin Ohlerth* [a,d], *Sebastian Ille* [a,b,c], *Bernhard Meyer* [a] and *Sandro M. Krieg* [a,b,c,*]

[a] *Department of Neurosurgery, Technical University of Munich, School of Medicine, Klinikum rechts der Isar, Germany*

[b] *TUM Neuroimaging Center, Technical University of Munich, School of Medicine, Klinikum rechts der Isar, Germany*

[c] *Department of Neurosurgery, Heidelberg University Hospital, Ruprecht-Karls-University Heidelberg, Germany*

[d] *Neurobiology of Language Department, Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands*

ARTICLE INFO

Article history:
Received 1 November 2022
Reviewed 23 January 2023
Revised 1 February 2023
Accepted 25 September 2023
Action editor Stephanie Forkel
Published online 22 November 2023

Keywords:
Navigated transcranial magnetic stimulation
Language mapping
Language comprehension
Severe expressive aphasia
Inter-rater-reliability

ABSTRACT

Objective: Stimulation-based language mapping approaches that are used pre- and intraoperatively employ predominantly overt language tasks requiring sufficient language *production* abilities. Yet, these production-based setups are often not feasible in brain tumor patients with severe expressive aphasia. This pilot study evaluated the feasibility and reliability of a newly developed language *comprehension* task with preoperative navigated transcranial magnetic stimulation (nTMS).

Methods: Fifteen healthy subjects and six brain tumor patients with severe expressive aphasia unable to perform classic overt naming tasks underwent preoperative nTMS language mapping based on an auditory single-word Comprehension TAsk for Perioperative mapping (CompreTAP). Comprehension was probed by button-press responses to auditory stimuli, hence not requiring overt language responses. Positive comprehension areas were identified when stimulation elicited an incorrect or delayed button press. Error categories, case-wise cortical error rate distribution and inter-rater reliability between two experienced specialists were examined.

Results: Overall, the new setup showed to be feasible. Comprehension-disruptions induced by nTMS manifested in no responses, delayed or hesitant responses, searching behavior or selection of wrong target items across all patients and controls and could be performed even in patients with severe expressive aphasia. The analysis agreement between both specialists was substantial for classifying comprehension-positive and -negative sites. Extensive left-hemispheric individual cortical comprehension sites were identified for all patients. Apart from one case presenting with transient worsening of aphasic symptoms,

Abbreviations: DES, direct electrical stimulation; nTMS, navigated transcranial magnetic stimulation; MRI, magnetic resonance imaging; SLT, speech and language therapist.

* *Corresponding author*. Department of Neurosurgery, Klinikum rechts der Isar, Technische Universität, München Ismaninger Straße 22, 81675 Munich, Germany.

E-mail address: sandro.krieg@tum.de (S.M. Krieg).

https://doi.org/10.1016/j.cortex.2023.09.023

pre-existing language deficits did not aggravate if results were used for subsequent surgical planning.

Conclusion: Employing this new comprehension-based nTMS setup allowed to identify language relevant cortical sites in all healthy subjects and severely aphasic patients who were thus far precluded from classic production-based mapping. This pilot study, moreover, provides first indications that the CompreTAP mapping results may support the preservation of residual language function if used for subsequent surgical planning.

1. Introduction

Language comprises a highly complex, interconnected neural network synchronizing numerous expressive and receptive functions (Chang et al., 2015; Friederici, 2017; Tremblay, Dick, & Small, 2011). Localizing functionally relevant areas necessary for language is one of the major objectives in the treatment of language-eloquent brain tumors in order to balance the overall survival, functional outcome and quality of life (Duffau & Mandonnet, 2013; Gogos et al., 2020; Ottenhausen, Krieg, Meyer, & Ringel, 2015). Whilst for this matter direct electrical stimulation (DES) during awake surgeries remains the gold standard, non-invasive navigated transcranial magnetic stimulation (nTMS)-based language mapping is increasingly employed preoperatively (Bährend et al., 2020; De Witt Hamer et al., 2012; Haddad et al., 2021; Ille, Sollmann, et al., 2016; Mandonnet et al., 2010; Picht et al., 2013; Szelényi et al., 2010; Tarapore et al., 2016). Both, nTMS and DES mapping allow localization of areas relevant for language function and, therefore, can guide preoperative planning and intraoperative resection, respectively.

Typically, overt *production* tasks are employed during language mapping (Hauck et al., 2015; Krieg et al., 2017; Rofes et al., 2015; Talacchi et al., 2013; Tarapore et al., 2013) as stimulation methods rely on identifying a causal link between a stimulation of a specific cortical area and the transient disruption of language function. The latter typically manifests in expressive language mistakes. However, these overt production tasks can be challenging for patients with language impairments affecting productive language abilities. Brain tumors located within language-eloquent areas can cause precisely these inabilities by affecting single or multiple stages of language (Faulkner et al., 2017; IJzerman-Korevaar et al., 2018). Expressive aphasia, one of the most widely studied and known language impairments, predominantly affects language production abilities whilst comprehension skills may be well preserved (Fridriksson et al., 2015). Studies suggest that preoperative language mapping can be confounded by distinct or severe aphasia as these impairments can lead to an increased number of errors during nTMS-based language mapping (Schwarzer et al., 2018). Especially severe manifestations of expressive aphasia can preclude patients completely from these overt production-based language mappings since patients are unable to name sufficient or any items repeatedly and correctly. Nonetheless, their comprehension skills may still be preserved enabling a comprehension-based language mapping. Fernandez Coello and colleagues stressed the importance of choosing stimulation tasks based on patient- and lesion-specific characteristics (Fernandez Coello et al., 2013). Hence, it is very important to develop tasks for patients with tumors affecting productive language skills and thereby to allow a language mapping pre- and intraoperatively in order to preserve unaffected language abilities.

More and more tasks and intraoperative testing batteries specifically target receptive functions (Alarcon et al., 2019; Bello et al., 2007; De Witte et al., 2015; Fernandez Coello et al., 2013; Gatignol et al., 2004; Martin-Monzon et al., 2022; Rofes et al., 2015; Rofes & Miceli, 2014). Yet, most of these receptive tasks still require an overt response by the patient. Only one pilot study tested the feasibility and the optimal stimulation parameters for an auditory sentence comprehension task during preoperative nTMS-based language mapping in three pediatric patients not requiring overt responses (Rejno-Habte Selassie et al., 2020). Reliable pre- and intraoperative language tasks entail the usage of items that a patient can respond to promptly and accurately during a time-restricted rapid presentation (Krieg et al., 2017; Rofes et al., 2015; Talacchi et al., 2013). However, these setups of *overt* or complex auditory comprehension tasks might be challenging for patients with expressive language impairments especially under the time-constrained conditions during language mapping.

Thus far, no study in adult patients with language-eloquent tumors and language deficits employed a receptive language test for preoperative stimulation-based language mapping. Therefore, developing a task suitable for patients with severe expressive aphasia unable to perform classic overt language mapping tasks is highly valuable. To this end, this pilot-study aims to evaluate the feasibility and reliability of a newly developed language Comprehension TAsk for Perioperative mapping (CompreTAP) in brain tumor patients with severe expressive aphasia and in healthy controls. Prior to testing this new setup in the operating room under more challenging and time-restricted conditions, this study examines its effectiveness and utility for preoperative nTMS language mapping. Since it is yet unknown in which ways comprehension-errors manifest in this new setup, we, moreover, assessed the analysis agreement of error evaluation between a neurolinguist and a trained speech and language therapist.

2. Material and methods

We report all data exclusions, all inclusion/exclusion criteria, whether inclusion/exclusion criteria were established prior to data analysis, all manipulations, and all measures in the study. Since this pilot study tested the feasibility of this new comprehension task within a first small cohort, no a priori sample size calculations were performed.

2.1. Patient and healthy subject population

Six brain tumor patients and 15 healthy controls were prospectively included between July 2021 and June 2023. All subjects needed to be at least 18 years old, German native speakers, and present without contraindications for magnetic resonance imaging (MRI) or nTMS such as cochlear implants or cardiac pacemakers. The absence of any neurological or psychiatric history was additionally required for the inclusion of healthy controls, patients needed to present with a severe expressive aphasia. This was attributed by a trained speech and language therapist (SLT) based on the individual performance on an object naming task typically used for preoperative nTMS language mapping (Krieg et al., 2017). Aphasia severity was rated from 0 = no aphasia to 5 = very severe aphasia. This rating is a modified version of a rating based on the Aachener Aphasie Test (Huber et al., 1983) used in former publications (Ille et al., 2021; Ille, Kulchytska, et al., 2016; Picht et al., 2013). Two additional severity points were added to the scale: No deficit (0), minimal symptoms such as occasional word finding difficulties with no impact on daily communication (1), light aphasic symptoms with a small impact on daily communication (2), moderate aphasic symptoms impacting but not limiting daily communication (3), severe aphasic symptoms with a profound impact on daily communication but simple communicative tasks still possible (4), Extremely severe aphasia precluding patients from daily communication (5). This allows to differentiate aphasia severity more thoroughly, particularly light symptoms. While the Aachener Aphasie Test is very useful in identifying moderate and severe aphasia, it does not differentiate minimal aphasic symptoms from no aphasic symptoms which may not adequately reflect the wide severity spectrum observed in clinical routine. Handedness was tested with the Edinburgh Handedness Inventory (Oldfield, 1971). Moreover, standardized language eloquence levels (low: 0–2, moderate: 3–5, high: 6–9) were determined for patients' tumors based on a recently published classification system (Ille et al., 2021).

Participants provided written informed consent. The study was approved by the local ethics committee of the Institutional Review Board (reference number: 192/18S) and followed the guidelines of the Declaration of Helsinki.

No part of the study procedures or analysis plans was preregistered prior to the research being conducted. The data is stored in an institutional repository and not publicly available due to hospital legislation and medical ethical objections. All data presented in this study are available upon reasonable request, access will be granted to named individuals in accordance with ethical procedures governing the reuse of sensitive data, i.e., if the ethical committee approves and if the data sharing agreement is signed by both a demanding and providing party. Readers seeking access to the data are advised to contact the corresponding author, Prof. Dr. med. S.M. Krieg.

2.2. MR image acquisition

Patients and healthy controls underwent a standardized structural MRI protocol in the department of neuroradiology on a 3-T MRI scanner (Achieva dStream or Ingenia; Philips Healthcare, Best, Netherlands) with an 8- or 32-channel phased-array head coil (Sollmann et al., 2016, 2018). This comprised at least a Diffusion Tensor Imaging (DTI) sequence with 32 diffusion sensitizing gradient directions as well as a three-dimensional T1-weighted gradient echo sequence, for healthy controls without contrast agent administration and for patients with and without contrast agent, respectively. These structural scans were subsequently used for neuronavigation to individually guide nTMS-coil positioning during stimulation.

2.3. CompreTAP task

Since one of our primary objectives was to construct a test suitable for patients with aphasia, classic diagnostic tools for rehabilitation planning as well as symptom and severity estimation acted as an initial orientation (Huber et al., 1983; Kertesz, 2007). These aphasia diagnostic instruments typically include an auditory language comprehension section which tests single-word as well as sentence comprehension. Based on one commonly employed tool in Germany, the Aachener Aphasie Test (Huber et al., 1983), our test consists of auditorily presented target items which were simultaneously shown in sets of four picture stimuli from which the correct target item had to be selected. In contrast to these commonly used tests, we opted for presentation of items without semantically or phonologically related distractors in each set of four black-and-white drawings in order to fit the task to the time-restricted presentation mode that can be time-locked to the stimulation application. The 62 items of everyday objects and animals stem from the object naming test of “the Verb And Noun Test for Peri-Operative testing (VAN-POP)” (Ohlerth et al., 2020). Thus, word frequencies, age of acquisition, syllable lengths and livingness of objects were balanced for in our item set. Legal copyright restrictions do not permit us to publicly archive or share the full set of stimuli used in this experiment in a trusted digital repository. All items used stem from the “verb and noun test for peri-operative testing (Van-POP)”, the copyright of the pictures belongs to the Rijksuniversiteit Groningen, the Netherlands. Readers seeking access to the stimuli are advised to contact the author, Dr. A.-K. Ohlerth [Ann-Katrin.Ohlerth@mpi.nl; ann.katrin.ohlerth@gmail.com]. Stimuli will be released if the declaration of usage agreement is signed, and all points of the agreement are followed closely. On the basis of these 62 items, sets of four figures were created, no additional masker figures were integrated. The items were randomly combined, while the only constraint applied was to control for phonological and semantic similarities precluding any distractor items. Each item was used on average 3.94 times for 28 different item sub-sets, which in turn each appeared on average twice. Moreover, the position of items was varied throughout, none of the item subsets re-appeared in identical order. All these sets were presented via PowerPoint on a computer screen together with non-synthesized pre-recordings of the target item (mean duration of prerecording: 1.0 sec, range: .5 sec–1.6 sec). The background of each of the four images on a slide were color-matched to four colored buttons and their respective position (left or right column, upper or lower row) and were placed between the participant and the computer screen. Fig. 1 shows the setup in an example item. As soon as the four images appeared on screen, participants heard the auditorily presented target word and were asked to select the matching target item by pressing the button with the corresponding color. To minimize hand motor difficulties during button press while stimulation is applied, patients were instructed to use their fingers of the left hand for pushing the button – ipsilateral to the subsequently stimulated left hemisphere. The presentation of the picture stimuli and the auditory stimulus were onset-aligned. Picture presentation lasted for 4 sec. Moreover, Big-Point recordable buttons (TTS, Nottinghamshire, UK) which are typically employed in context of alternative communication were used as these allowed to pre-record the color label of each button. Consequently, each button press elicited the corresponding pre-recorded color label of the respective button chosen. This acted as an auditory control for the nTMS operator during the analysis of the reaction times post-mapping and allowed to monitor closely patient's performance and attention to the task during the mapping.

2.4. *nTMS-based CompreTAP language mapping*

Participants underwent language mapping using the Nexstim eXimia NBS system, version 5.1 with a NEXSPEECH® module (Nexstim Plc, Helsinki, Finland). Prior to performing the comprehension task under stimulation, subjects and patients completed two baseline trials without nTMS to preclude any items that could not be identified promptly and correctly by the individual. Based on numerous neuroimaging findings demonstrating that acoustic-phonetic, lexical, morphologic as well as syntactic and semantic comprehension processes are all performed within under 1 sec after auditory stimulus onset (Bornkessel-Schlesewsky et al., 2016; Eckstein & Friederici, 2006; Friederici, 2002, 2011; Getz & Toscano, 2021; Hagoort et al., 2004), stimulation during nTMS mapping was applied for 2 sec (10 repetitive pulses) covering the entire duration of the auditory stimulus presentation and presumed comprehension processes. The inter-stimulation interval was set to 4000 msec matching the slide presentation duration. Following our standard object naming protocol, the stimulation frequency was applied at 5 Hz, the intensity set at 110% of the ipsilateral resting motor threshold (Krieg et al., 2016, 2017). The resting motor threshold is defined as the minimum necessary stimulation intensity needed for eliciting a motor evoked potential in the abductor pollicis brevis. The nTMS comprehension mapping targeted the majority of the frontal, parietal and temporal lobe of the left, tumor-hemisphere as these results were substantial for the subsequent surgical workflow. Each of 46 predetermined left-hemispheric stimulation target sites was stimulated three times with repetitive nTMS (Fig. 2, abbreviations Table 1).

Fig. 1 – Exemplary item setup. Each item comprises an auditorily presented target item (calendar) and four visually presented items shown simultaneously on the computer screen with the respectively colored background matching four colored buttons.

2.5. *Identification of language comprehension positive nTMS sites*

The camera of the nTMS device was positioned in such a way that it could record the hand movement of the subject or patient reaching for the button. Moreover, the auditorily presented target item and button-press sounds were recorded on video. Two nTMS operators specialized on language and language impairments, a trained SLT (rater 1, LK) and a neurolinguist (rater 2, AKO), with extensive nTMS mapping experience of standard overt naming protocols (rater 1: ~100, rater 2: >200 language mappings), identified stimulation-induced comprehension errors based on video recordings of the stimulation exam. Videos were scanned for deviant behavior during button press compared to baseline behavior, such as delayed or incorrect responses or change of hand positioning. Blinded to the stimulation site, both raters finally marked all comprehension errors following stimulation-caused disruptions of the language network as comprehension-positive. These were then transferred to the neuronavigation system indicating specific cortical sites at which stimulation elicited a comprehension error, allowing to delineate comprehension-positive and comprehension-negative cortical sites. Error rates were defined as the number of errors divided by number of stimulations applied for each cortical region as parcellated based on Corina's cortical system (Corina et al., 2005, Fig. 2, Table 1). All error rates reported prior to inter-rater assessment were based on the analysis of rater 1.

Since the introduction of synchronous video recordings and nTMS mappings the video-based analysis became standard and is widely applied across centers in the field of pre-operative nTMS-based language and cognitive mappings (Lioumis et al., 2012). It allows to identify a wide variety of

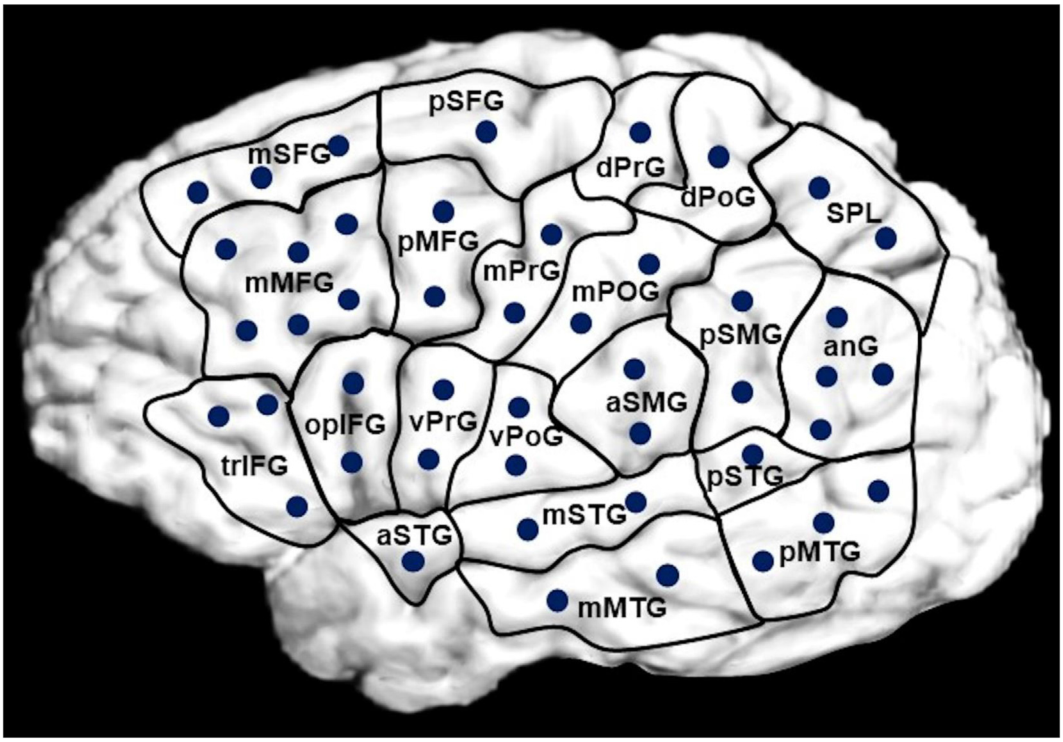

Fig. 2 – Stimulation target template. Depicted are 46 left-hemispheric targets based on cortical parcellation system (CPS) regions. See Table 1 for corresponding abbreviations of CPS regions.

Table 1 – Overview of cortically parcellated areas.

Abbreviation	Cortical anatomical area
anG	Angular gyrus
aSMG	Anterior supramarginal gyrus
aSTG	Anterior superior temporal gyrus
dPoG	Dorsal postcentral gyrus
dPrG	Dorsal precentral gyrus
mMFG	Middle middle frontal gyrus
mMTG	Middle middle temporal gyrus
mPoG	Middle postcentral gyrus
mPrG	Middle precentral gyrus
mSFG	Middle superior frontal gyrus
mSTG	Middle superior temporal gyrus
opIFG	Opercular inferior frontal gyrus
pMFG	Posterior middle frontal gyrus
pMTG	Posterior middle temporal gyrus
pSFG	Posterior superior frontal gyrus
pSMG	Posterior supramarginal gyrus
pSTG	Posterior superior temporal gyrus
SPL	Superior parietal lobe
trIFG	Triangular inferior frontal gyrus
vPoG	Ventral postcentral gyrus
vPrG	Ventral precentral gyrus

The table provides an overview of the cortical parcellations system shown in Fig. 1, based on Corina et al. (2005) and the corresponding abbreviations used.

stimulation-induced errors with a high accuracy and reliability, such as no responses or semantic and phonologic paraphasias (Lioumis et al., 2012; Sollmann et al., 2013). Still, while accounting for a large proportion of identified errors, hesitant responses are typically considered as the most subjective and least reliable error category induced by nTMS (Krieg et al., 2016; Ohlerth, Bastiaanse, Negwer, et al., 2021). The initial identification of hesitant responses by both nTMS operators was based on the video recordings, as this is thus far the standard approach in preoperative nTMS-based language mappings during clinical routine. Moreover, additional reaction time analyses were performed to ascertain the reliability and accuracy of this most subjective error category based on an approach proposed by Schramm et al. (2020). All video recordings available in .asf format were copied to an external computer to extract the audio track in .wav format. For this, the Python-module MoviePy version 1.0.3 (Zulko, 2020) was employed. Subsequently Praat version 6.3.04 (Boersma & Weenink, 2023) was used to measure the response times for each item separately. The reaction time between each auditory stimulus onset and the onset of the respective prerecorded color label elicited with each button press were measured and documented. All items at which any other error type occurred, i.e., press of the wrong target button, searching behavior or no responses were excluded. Subsequently, hesitations in reaction time were defined as all responses that exceeded two standard deviations of the mean response time per individual and compared to the analysis of rater 1. Moreover, since reaction times tend to vary over the course of the stimulation examination, the mean and standard deviation of the last five error-less items preceding the item marked as hesitant by rater 1 were examined separately.

2.6. Function-based tractography of the subcortical language network

All language-positive cortical sites were subsequently used as seeds for an individual tractography of the functional language network. DTI-based tractography was conducted with a deterministic tracking algorithm embedded into Brainlab Elements (version 3.2.0.281; Brainlab AG, Munich, Germany) following a standard protocol (Negwer et al., 2017; Sollmann et al., 2018; Sollmann, Zhang, Schramm, et al., 2020). These results were subsequently used to guide preoperative surgical planning and intraoperative resection of the tumor. Albeit the present nTMS-based tractography results allow not to draw any conclusions about the benefit of this comprehension-based functional tractography compared to other possible variants, they offer preliminary insights into the utility of this mapping setup for preserving residual language function as the combined results of comprehension-positive cortical sites and subsequent tractography were used perioperatively to support the preservation of language skills in the present cohort.

2.7. Inter-rater reliability and statistical analysis

All statistical analyses were performed with R (R Core Team, 2020). A p-value <.05 was considered statistically significant. To ascertain the reliability of this newly developed task, the two raters individually analyzed the video recordings of the stimulation exams. Cohen's kappa was used to compute inter-rater agreement between the two nTMS operators (Gamer et al., 2019) as well as between the analysis of hesitant responses by rater 1 and hesitations identified with the reaction time analysis. A kappa of 1 was considered almost perfect (Landis & Koch, 1977). Bangdiwala's agreement chart for categorical data was used to additionally graphically compare inter-rater reliability (Bangdiwala, 1988; Bangdiwala & Shankar, 2013).

To ascertain the impact of noise caused by the nTMS application on participant's ability to hear the auditory stimulus presented simultaneously, a post-hoc comparison of the respective intensities in isolation was conducted. For this, audio recordings were taken of exemplary stimulation-noises of 20 pulses for different stimulation intensity settings as well as of 10 exemplary item presentations, with a distance of 3 cm between recording device and the stimulation coil/the PC-screen. Mean intensity values were extracted with Praat (Boersma & Weenink, 2023) and analyzed descriptively.

3. Results

3.1. Identifiable error categories

For all patients and controls, nTMS application led to recognizable comprehension-errors. Across patients and controls, the following comprehension error categories were identified and subsequently marked as comprehension-positive:

No response errors comprise stimulation-induced errors in which subjects did not select any item – similar to no response naming errors observed during naming-based language mapping (Corina et al., 2010).

Searching behavior were responses during which a subject did not select the target item promptly but was heading for different buttons with an obvious uncertainty while eventually pressing a correct target item.

Selection of wrong target item were errors in which a subject pressed a colored button corresponding to any of the other three visually presented items, e.g., "key" (red button) was pressed instead of the target "saw" (blue button). These errors correspond to the semantic error category, found in production-based language mapping, where an incorrect label is uttered during naming.

Hesitations/delayed responses were classified if subjects showed an obvious hesitant selection of the target item or chose the correct button with significant delay compared to baseline behavior.

All of these errors attributed to comprehension difficulties induced by stimulation were differentiated from hand-motor or coordinative difficulties. The latter was assigned on the basis of non-directed hand motor activation not clearly aimed at a specific button. Still, these error types appeared only very rarely since the left hand, ipsilateral to the stimulated hemisphere was used for button pressing.

3.2. Healthy subject characteristics and comprehension mapping results

Fifteen healthy subjects completed the comprehension language mapping. All subjects were able to perform the language comprehension task during baseline at ceiling levels, with 100.0% of items being identified correctly and promptly twice prior to nTMS application. Moreover, the language comprehension task was feasible during nTMS application in all subjects and nTMS-induced comprehension errors could be identified in all individuals. For a detailed overview of subject characteristics and individual error rates during nTMS see Table 2. The individual, illustrative error rates across CPS regions for the first six healthy controls (C1–C6) are additionally depicted in Fig. 3.

Table 2 – Overview of subject characteristics and error rate during nTMS application.

Subject	Age	Sex	Handedness	Error rate during nTMS
C1	31	Female	R	10.1%
C2	26	Male	R	3.0%
C3	25	Female	R	5.1%
C4	25	Female	R	8.0%
C5	24	Female	R	16.7%
C6	33	Male	R	12.3%
C7	26	Male	R	11.6%
C8	20	Female	R	9.4%
C9	24	Male	R	4.3%
C10	28	Female	R	7.2%
C11	24	Male	R	20.3%
C12	27	Male	R	7.2%
C13	23	Male	R	12.3%
C14	25	Female	R	12.3%
C15	22	Female	R	9.4%

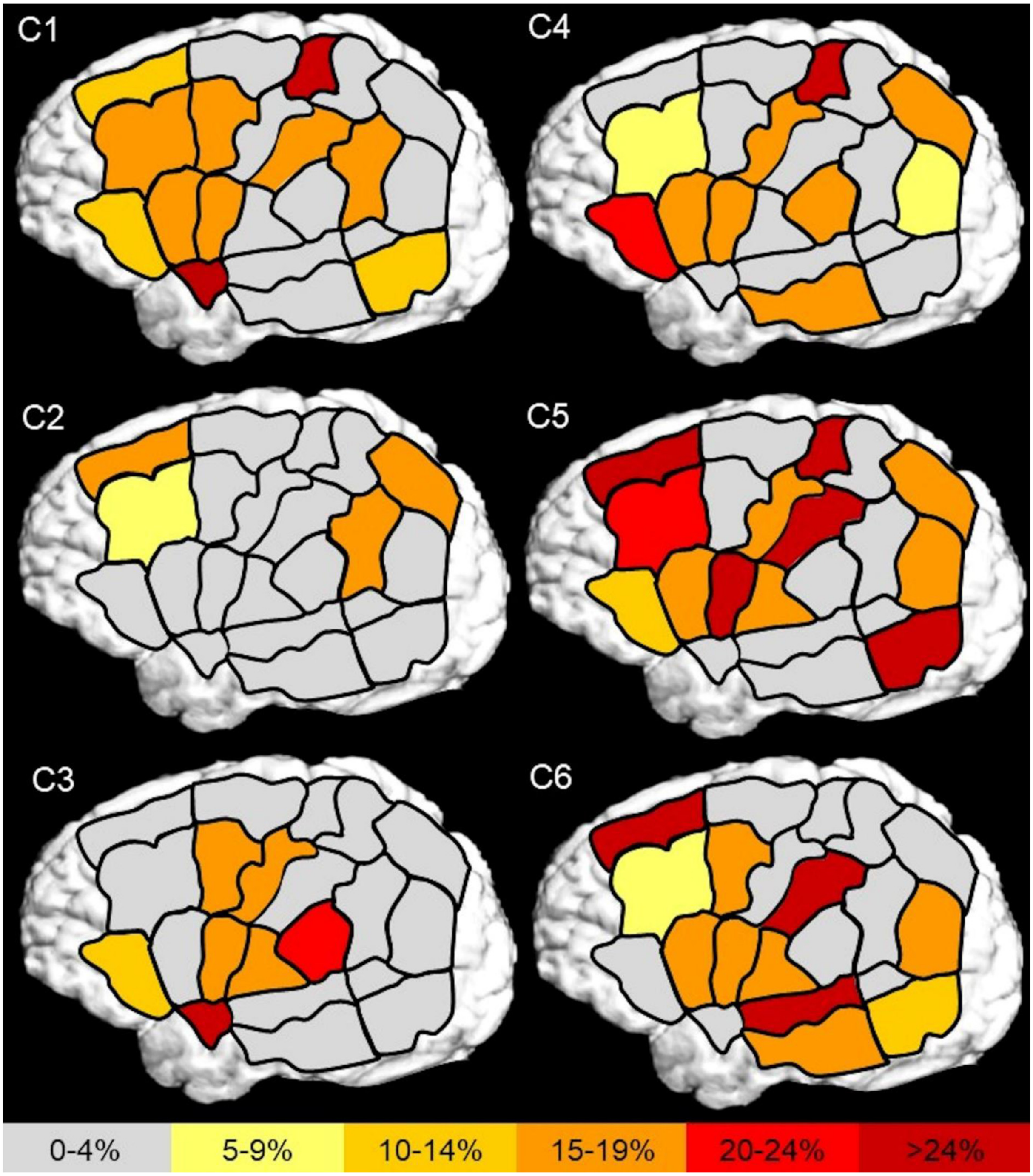

Fig. 3 – Error rate distribution in six illustrative healthy controls. Case-wise presentation of cortical error rate distribution in percent across predefined CPS regions for each healthy control (C1–C6).

3.3. Patient characteristics and comprehension mapping results

The new comprehension-based language mapping setup was piloted in six patients with language-eloquent intracranial lesions (overview of patient characteristics in Table 3). All patients presented with a severe expressive language deficit, i.e., classic object-naming-based language mapping was not feasible. Apart from case 2 who was still able to name numbers, none of the patient could perform any other expressive language task. Thus, for the latter five patients, the comprehension-positive nTMS results were used to prepare functionally relevant language tractography.

All patients were able to understand the instructions to the comprehension task and could reliably and reproducibly identify a sufficient number of items – on average 62.8% (±21.6%) – via button-press during the two baseline trials. The individual item set, which finally only comprised the items a patient identified correctly, was used during stimulation. A detailed overview of the patients' preexisting language deficit,

Table 3 – Overview of patient characteristics, tumor entity and location, as well as standardized language eloquence levels.

Patient	Age	Sex	Handedness	Tumor entity	Tumor location	Language eloquence
P1	63	Male	A	GBM Re	Parieto-occipital	7
P2	69	Male	R	GBM Re	Limbic	7
P3	71	Male	R	GBM	Temporo-occipital	8
P4	60	Female	R	GBM Re	Temporo-occipital	8
P5	43	Female	R	M	Frontal, temporal	5
P6	72	Female	R	GBM	Frontal	8

The table provides an overview of age, sex, handedness (R = right-handed, A = ambidextrous), tumor entity (GBM = glioblastoma, M = metastasis, Re = recurrence), tumor location and language eloquence level (0–9).

number of errors and error rate during the baseline prior to nTMS, number of stimulations applied in total and errors during nTMS, as well as the resulting error rate during nTMS are provided in Table 4. The individual error rates across CPS regions are additionally depicted in Fig. 4. In some of the cases, the placement of the individual targets was adjusted due to tumor location, edema or a previous resection. Moreover, as Appendix A shows, some patient cases did not tolerate the stimulation within some frontal or anterior temporal areas due to increased pain levels.

In the five cases, in which comprehension-based nTMS results and subsequent functionally-relevant tractographies were used to guide surgical planning and resection, four patients did not develop any new language deficits post-surgery. Case 6, however, showed a transient worsening of aphasic symptoms directly post-surgery which improved during the postsurgical one-week hospitalization.

In the following section, two illustrative cases will be described in detail. For these two patients, case 1 and case 6, the comprehension-based DTI-tractography results are illustrated in Fig. 5.

3.3.1. *Case 1*

This ambidextrous 63-year-old male patient was referred to our department after an external clinic suspected a left, parieto-occipital glioblastoma recurrence. The patient presented with a worsening of a pre-existing aphasia eight months after his first resection and subsequent radiochemotherapy. He had a high grade of language eloquence (grade: 7). The patient was unable to perform object or number naming preoperatively due to the highly severe expressive aphasia. However, he was able to understand the instructions to the comprehension task and could complete 47.6% of items correctly during the two baseline trials. Disruptions of nTMS predominantly elicited searching behavior or no responses, rarely selection of a wrong target item. The error rate under nTMS was 14.9%. Cortical areas with the highest error rates comprised parietal and temporal ones (Fig. 4, P1): anG, pSTG, pMTG and SPL. Since this was the only feasible task for case 1, the results of the nTMS language mapping were used to prepare functionally relevant language tractography and to guide intraoperative resection. Histopathology confirmed a WHO grade IV glioblastoma. Clinical assessment indicated no new language deficits post-surgery.

3.3.2. *Case 6*

Case 6, a 72-year-old right-handed, female patient, presented with increasing aphasic symptoms since 3–4 weeks. Preoperative imaging indicated a left glioblastoma within the opercular inferior frontal gyrus. Analysis indicated a high language eloquence (grade: 8). This diagnosis was supported by histopathological results post-surgery. Her language production abilities were severely impaired prior to craniotomy. The production of single-words alone was possible. However, rapid and prompt naming as needed for overt language mapping was not feasible. Additionally, the tumor also seemed to affect language comprehension skills, as she was only able to select 37.1% of comprehension items correctly. Since these, however, were possible reliably and repeatedly, the comprehension task was used for nTMS language mapping and its results for function-specific tractography and neurosurgical guidance. Under stimulation, case 6 made 25.4% comprehension errors, including no response errors and delayed responses, selection of wrong target item and searching behavior. P6 had overall high error rates particularly in parietal and temporal regions and in the opIFG (Fig. 4, P6). Directly post-surgery, her aphasic symptoms were transiently stronger pronounced, but improved substantially during clinical stay.

Table 4 – Overview of the individual language status and error rate during nTMS application.

Patient	Aphasia[a]	Number (percentage) of baseline errors	Number of repetitive stimulations applied	Number of errors during nTMS	Error rate during nTMS
P1	E5, R3	33 (52.4%)	121	18	14.9%
P2	E4, R0	3 (4.8%)	134	26	19.4%
P3	E5, R3	31 (50.0%)	137	24	17.5%
P4	E5, R1	6 (9.7%)	138	16	11.6%
P5	E5, R2	26 (41.9%)	125	26	20.8%
P6	E5, R3	39 (62.9%)	138	35	25.4%

[a] Severity 0–5, E = expressive, R = receptive.

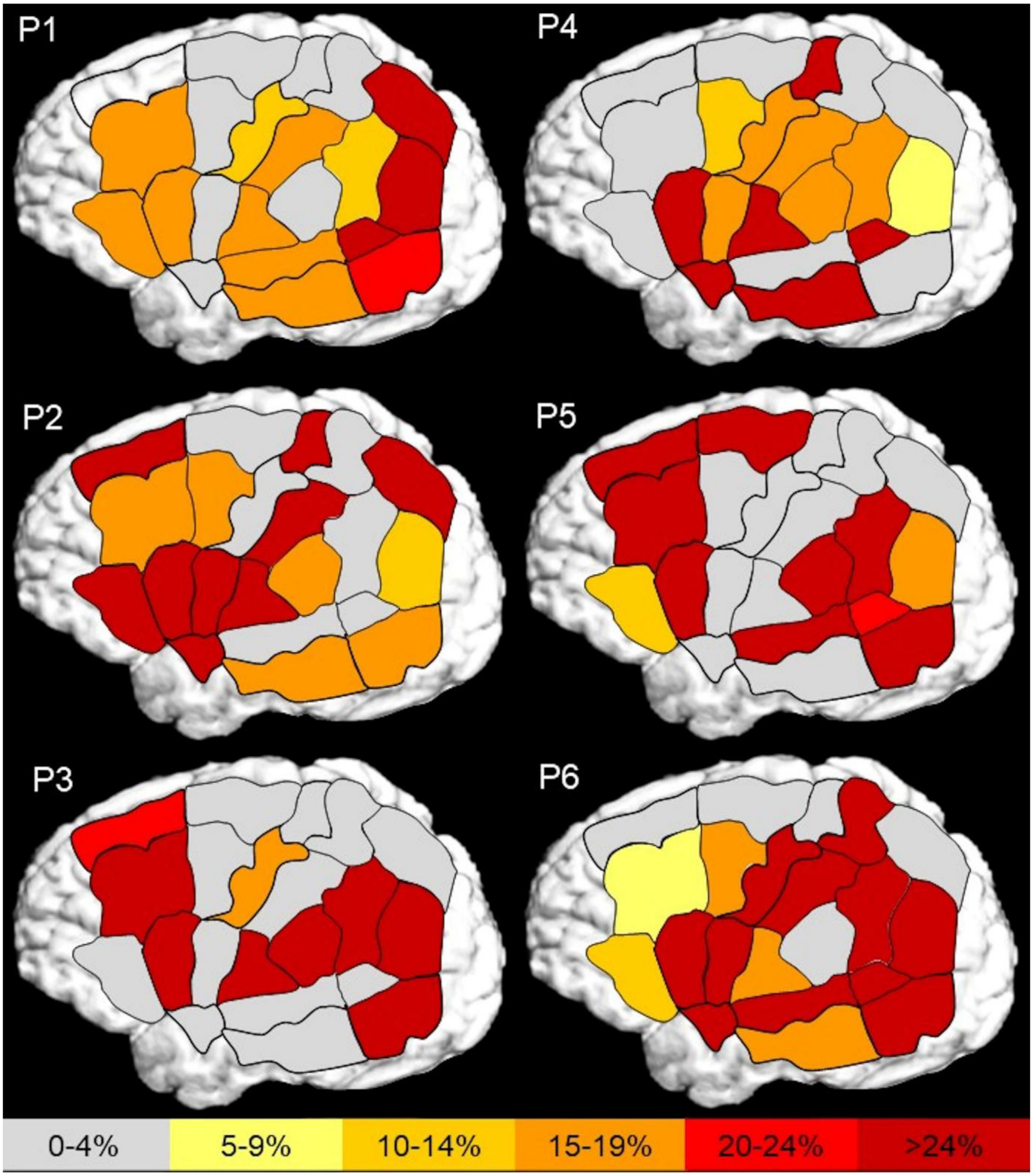

Fig. 4 – Error rate distribution in patient cohort. Case-wise presentation of cortical error rate distribution in percent across predefined CPS regions for each patient (P1–P6).

3.4. *Group-wise comparison of different comprehension error types*

While the sample sizes of patients included in this pilot study does not warrant any statistical group-wise comparisons, the preliminary mean cortical error rate for each cortically parcellated area was compared graphically between patients and controls (Fig. 6). As can already be seen in the group-wise comparison across comprehension errors (Fig. 6A), the distribution of comprehension errors elicited by nTMS were wide-spread across the entire left hemisphere. Overall, across categories, higher error rates were observed in patients. The error rate pattern for each error category (no response, searching behavior, selection of wrong target item and hesitant responses) specific to patients or controls is shown in Fig. 6(B–E). Of note, across controls no response errors as well as selection of wrong target items manifested only rarely and thus, did not allow a detailed error pattern analysis.

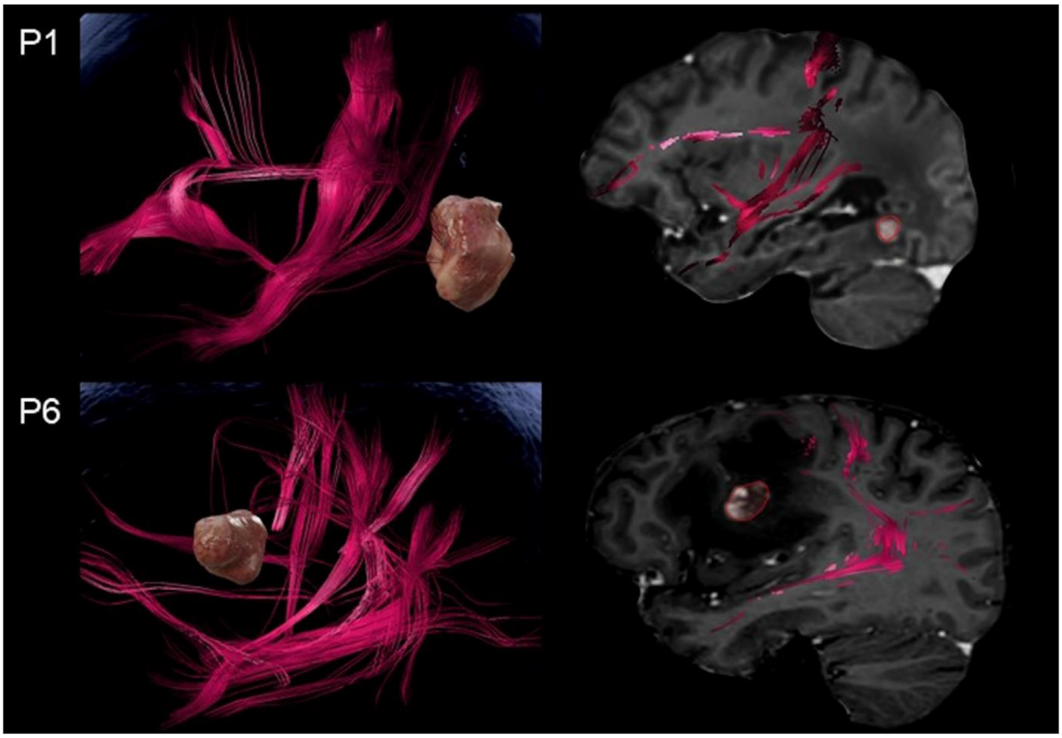

Fig. 5 – Exemplary function-based tractographies. Illustrations show the individual functional language network (pink) in relation to the respective glioblastoma (left column: brown, right column: red outline) of case 1 (P1) and case 6 (P6).

3.5. *Inter-rater reliability*

We compared the inter-rater agreement for stimuli identified as comprehension-positive or comprehension-negative between a neurolinguist and a SLT for patient and control data separately. Both groups showed a significant, substantial agreement strength (control group K = .65, $p < .001$; patient cohort K = .66, $p < .001$), across all comprehension-errors and comprehension-negative stimuli during the stimulation exam (Fig. 7). Additionally, we assessed the agreement between raters for each classified error category. Table 5 summarizes these results. For patients, both raters had a highly significant, substantial agreement for no response errors, searching behavior and selection of wrong target item (all $p < .001$), but not for hesitations. For controls, the inter-rater agreement was substantial for searching behavior, almost perfect for selection of wrong target item and fair for hesitations and no responses (all $p < .001$). Still, rater 2 classified a single and rater 1 five clear no responses across all healthy subjects. Three out of the latter five no responses attributed by rater 1 were classified as hesitations by rater 2.

3.6. *Reaction time-based analysis of delayed responses*

To establish whether the more subjective, video-based analysis is concordant with objective reaction time analyses, reaction times were measured for each item during the stimulation examination. Since no procedure to analyze reaction times systematically is readily available within the nTMS system in use, the manual extraction based on the respective audio track was performed within a third-party program. This additional analysis took on average 79.3 ± 8.9 min for patients and 53.3 ± 6.5 min for controls. Since inter-individual variability in naming latencies is well established (Jodzio et al., 2023), the mean and standard deviation of response times during the stimulation exam were determined for each individual. The descriptive results are summarized in Table 6. Durations exceeding two standard deviations of the individual mean response time were classified as delayed responses. The subsequent agreement analysis between this duration-based identification of hesitations and the analysis of rater 1 revealed a significant, yet slight agreement for patients (K = .132, $p < .001$) and for controls (K = .118, $p < .001$). Still, of 66 hesitations classified by rater 1 across patients and healthy subjects, only 22.7% were assigned on the basis of a response delay, whilst 71.2% were attributed based on hesitant hand motions such as halting or indecisive movements prior to a button press and 6.1% based on a combination of hesitant hand motions and delays. Out of all subjectively classified delayed responses by rater 1 (n = 15), 53.3% exceeded the mean individual reaction time by more than two standard deviations.

The additional analysis of the subset of items preceding each hesitation revealed that 60.0% of ten hesitations identified for patients and 33.9% of 56 hesitations identified for controls had a delay of more than two standard deviations from the mean of the five error-less items preceding each identified hesitation by rater 1. Moreover, if only the fifteen hesitations identified by rater 1 on the basis of seemingly delayed response behavior were considered, 66.7% of 12 delayed responses for controls and 100.0% of 3 delayed responses for patients were concordant with the objective reaction time cut-off criteria.

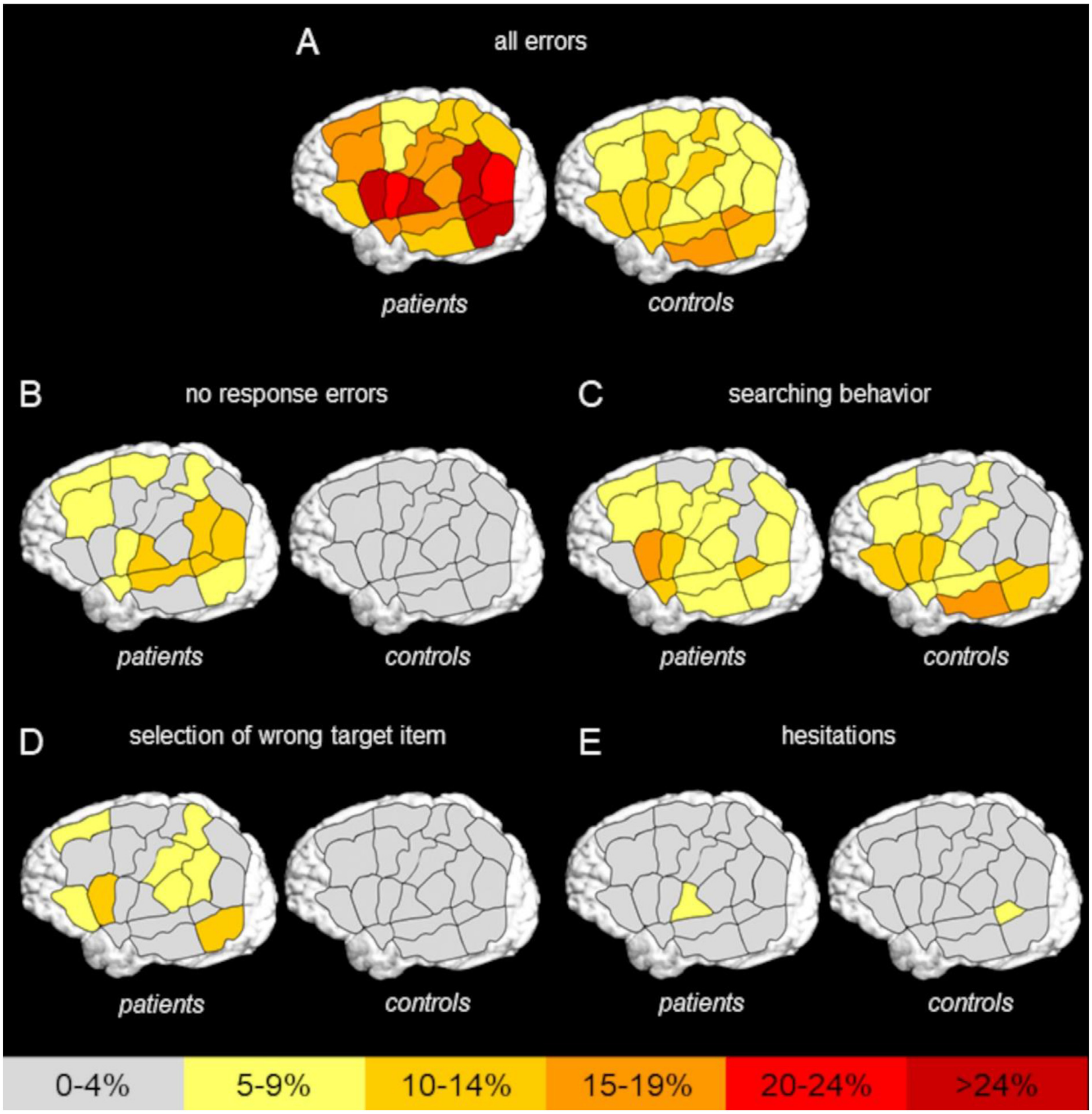

Fig. 6 – Group-wise presentation of cortical error rate distribution in percent across predefined CPS regions for all stimulation-induced errors (A) and specific error types: no response errors (B), searching behavior (C), selection of wrong target item (D) and hesitations (E).

3.7. *Analysis of noise impact on capability of hearing auditory stimuli*

Finally, the impact of the noise arising from the nTMS system and the applied pulses on the capability to hear the auditory stimuli presented was analyzed. This post-hoc analysis considered recordings of stimulations applied with a non-cooled stimulation coil at a stimulation intensity of 20–50% (in steps of 10%) based on a mean motor threshold of 31.7% (range: 25–39%) across patients and controls. The mean intensity in dB for the exemplary 20 pulse recordings was 57.3 dB while the one for the exemplary 10 items was higher with a mean of 73.4 dB. Moreover, the mean number of correctly identified target items out of 138 stimulations applied per participant were 119.5 across patients and controls, demonstrating that target items could be heard while stimulation was applied simultaneously.

4. Discussion

Preoperative nTMS-based language mapping is constantly gaining importance in neurosurgical context due to its ability

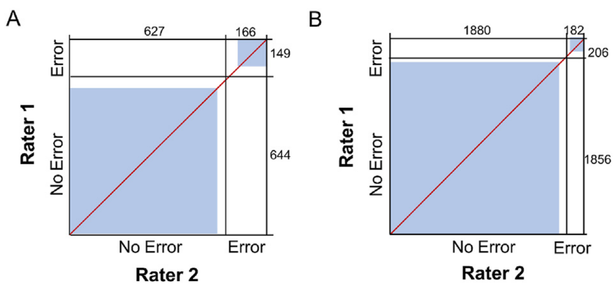

Fig. 7 – Bangdiwala's agreement charts. Charts compare the analysis agreement of stimuli identified as comprehension-positive (error) or comprehension-negative (no error) by rater 1 (SLT) and rater 2 (neurolinguist) for patients (A) and controls (B).

to identify areas necessary for language function, to inform surgical planning and to guide resections of language-eloquent brain tumors (Haddad et al., 2021; Ille, Sollmann, et al., 2016; Sollmann, Zhang, Fratini, et al., 2020). However, thus far the pre- and intraoperative language tasks for stimulation-based language mapping are mainly limited to overt production or complex comprehension paradigms. This, however, precludes brain tumor patients with severe expressive language impairments from stimulation-based language mapping. To account for this, we developed a novel comprehension-based language mapping task suitable for patients with severe expressive aphasia to be used for pre- as well as intraoperative mapping of the remaining language function.

4.1. *Feasibility and utility of the single-word auditory nTMS comprehension task*

The auditory single-word comprehension test, CompreTAP, was feasible in all patients and healthy controls. All healthy controls performed this task without difficulty and within the time limits, hence no item needed to be excluded retrospectively. This highlights the simplicity of the task which was necessary to allow language comprehension mapping under the time-restricted conditions in our severely aphasic patient cohort. Whereas language production was severely impaired in all patients, the receptive language comprehension skills of our aphasic patient cohort were preserved to varying extends prior to surgery. Especially case 1 and 6 presented with pre-existing moderate language comprehension deficits. Yet, they were able to identify sufficient items correctly and reproducibly during the two baseline trials to be used during mapping. In contrast, none of the patients could produce sufficient object naming items for classic naming-based approaches. This underlines the usefulness of this novel short single-word task as this comprehension

Table 5 – Overview of classified errors by each rater and Inter-rater agreement coefficient.

Comprehension error category	Patients Number of errors		Cohen's Kappa Patients (K)	Controls Number of errors		Cohen's Kappa Controls (K)
	Rater 1	Rater 2		Rater 1	Rater 2	
No response	47	32	.6***	5	1	.3***
Searching behavior	54	62	.6***	145	121	.6***
Selection of wrong target item	39	40	.7***	2	2	1.00***
Hesitation/delayed response	6	35	<.1	54	58	.3***

p-value: *<.05, **<.01, ***<.001.

Table 6 – Descriptives of individual response times across patients and healthy subjects.

ID	Patients RT in sec Mean ± SD (range)	Patients n hesitations	ID	Controls RT in sec Mean ± SD (range)	Controls n hesitations
P1	2.0 ± .4 (1.1–3.5)	4	C1	1.2 ± .2 (.1–1.8)	3
P2	1.7 ± .3 (1.0–3.0)	4	C2	1.3 ± .3 (.7–2.4)	5
P3	2.4 ± .4 (1.5–3.5)	2	C3	1.5 ± .3 (1.1–3.4)	3
P4	2.0 ± .3 (1.4–3.2)	5	C4	1.2 ± .3 (.7–3.2)	6
P5	2.4 ± .5 (1.3–4.1)	5	C5	1.0 ± .2 (.6–2.4)	4
P6	2.1 ± .6 (1.2–3.9)	5	C6	1.2 ± .2 (.9–2.0)	5
			C7	1.4 ± .3 (.8–2.2)	5
			C8	1.3 ± .3 (.8–2.1)	4
			C9	1.5 ± .3 (1.0–2.7)	7
			C10	1.2 ± .2 (.9–2.1)	6
			C11	1.4 ± .3 (.8–2.6)	5
			C12	1.3 ± .2 (.9–2.3)	3
			C13	1.4 ± .3 (.8–2.4)	8
			C14	1.2 ± .2 (.8–2.3)	2
			C15	1.7 ± .3 (1.0–3.4)	3

The table provides an overview of the mean, standard deviation (SD) and range of individual reaction times (RT) in seconds (sec) and the number (n) of hesitations identified as delayed response times exceeding the individual mean of a subject by more than two standard deviations.

setup was suitable even for patients with severe expressive aphasia.

Moreover, our results provide first support for the utility of this task for the preservation of residual language function if the results are used in combination with function-based tractography to guide tumor removal. Only in case 6, a transient worsening of aphasia symptoms was reported following comprehension-based nTMS guided surgery, which is often described following overt production tasks as well. This patient presented with the most severe comprehension deficit of the investigated cohort even prior to surgery. Only 23 out of the 62 target items presented could be used during stimulation which may limit the validity of mapping results as this reduced number of items may not be an adequate representation of the residual language comprehension abilities. Still, even this limited representation provided valuable insights into the patient's functional language comprehension network and may have supported the preservation of residual comprehension skills as no permanent worsening of symptoms arose. Moreover, since the task in use has not been evaluated prior to this study, adequate cut-off criteria need yet to be established. For case 1, 3, 4, and 5, however, for whom comprehension-based mapping results were used intraoperatively, no worsening of language deficits was identified even if the resected tumors were moderately (case 5) to highly language-eloquent (case 1, 3 and 4). Still, further studies assessing the benefit of this task for postoperative outcome in a larger cohort are needed.

4.2. Test construction and theoretical considerations

In this auditory single-word comprehension task, subjects are asked to choose the auditorily presented target item out of a set of four picture stimuli by button press. Due to methodological differences this setup is not directly comparable to other comprehension mapping paradigms. Few studies so far have tested specific intraoperative comprehension setups requiring only a pointing-based response from the patient and in line with our task no overt responses (Roux et al., 2015). Still, most intraoperative comprehension setups are based on tasks in which a patient must produce an overt response (De Witte et al., 2015; Martin-Monzon et al., 2022). A single study piloted a preoperative mapping based on sentence comprehension and picture matching in children with nTMS without overt responses (Rejno-Habte Selassie et al., 2020). In these non-overt comprehension stimulation-based approaches, visually presented association tests or complex auditory comprehension tasks were implemented, such as adaptions of classic “Token Test” (De Renzi & Vignolo, 1962) or sentence comprehension tests. The former association test based on visually presented images of objects fails to examine additional auditory comprehension levels such as acoustic-phonological categorization and lexical access processes (DeWitt & Rauschecker, 2012; Friederici, 2012; Okada et al., 2010). At the same time, tasks utilizing comprehension of sentences or connected speech involve additional syntactic, semantic and prosodic processing steps next to the aforementioned acoustic-phonological and lexical ones (Friederici, 2002) which in turn increase the difficulty of the task. Whilst sentence comprehension seems to recruit widespread frontal and temporal language areas, neuroimaging, stimulation and lesion studies indicate the involvement of partly similar areas in single-word comprehension (Bornkessel et al., 2005; DeWitt & Rauschecker, 2016; Herrmann et al., 2009; Mesulam et al., 2019; Roux et al., 2015; Zaccarella & Friederici, 2017). Our results show that this single-word-based comprehension mapping allowed to identify extensive cortical comprehension sites. Consequently, more complex auditory setups on sentence level might not necessarily be beneficial for identifying wide-spread comprehension-relevant sites across the entire left hemisphere. What is even more, the time-restricted presentation mode during stimulation-based language mappings limits the possibilities of introducing more complex and consequently more lengthy comprehension paradigms. Still, direct comparisons between sentence-based and single-word nTMS tasks would be needed to answer this question.

More and more comprehensive language testing batteries tailored to patients needs and individual lesion characteristics as well as locations are employed during awake stimulation-based language mapping (De Witte et al., 2015; Fernandez Coello et al., 2013; Martin-Monzon et al., 2022; Rofes et al., 2015). For the same reason, new tasks for preoperative stimulation-based language mapping are developed (Hauck et al., 2015; Ohlerth et al., 2020). Whilst object naming remains the method of choice and is one of the most frequently utilized tools pre- and intraoperatively, benefits of multiple-task approaches were verified (Ohlerth, Bastiaanse, Nickels, et al., 2021). Up to date, however, especially these preoperative tasks are based on assessing the language production network. Thus, CompreTAP might not only be suitable for patients with severe expressive aphasia but may also add valuable insights into the localization of language functions in context of multi-task approaches irrespective of aphasia severity.

4.3. Comprehension errors and reliability of this comprehension task

As opposed to classic nTMS language protocols, this new comprehension-based setup is based on button press. This, however, alters the process of error evaluation drastically and, therefore, deserves further validation: By comparing the analysis agreement of two highly experienced specialists, we showed that this new task has a substantial inter-rater reliability for patient and control data.

Both raters classified the error pattern across healthy subjects and patients into four error categories: no responses, selection of a wrong target item, searching behavior, and hesitant or delayed responses. All of these errors can potentially be caused by stimulation-induced interference on the ability to visually identify figures and to select the target item via a button press based on adequately matched colors. Still, whilst in classic naming-based approaches the same potential of disrupting the ability to identify figures adequately exists, ample research has shown that the results of these mappings can be used for preserving language function (Hendrix et al., 2017; Ille et al., 2021; Ille, Sollmann, et al., 2016; Natalizi et al., 2022; Raffa et al., 2019). The preliminary results of the present study additionally provide first indications that the comprehension-based setup may support the preservation of residual language function as no new or only a single transient worsening of aphasic symptoms manifested post-surgery. Thus, the majority of results generated seem to reflect disrupted higher-order language processes. Whilst no responses are one of the most frequent errors during naming-based stimulation approaches, underlying mechanisms causing this error type are not well understood (Corina et al., 2010). In production-based approaches, the origin of no response errors is difficult to disentangle since they may result from blocked speech motor planning as well as word finding difficulties. Our comprehension-based approach can circumvent this issue by using the ipsilateral hand to press the corresponding button to minimize stimulation-effects on the hand-motor response. Thus, this setup may allow an even clearer interpretation of disrupted language processing manifesting in no response errors compared to overt naming tasks. At the same time, like production-based nTMS, hesitations elicited by comprehension-based nTMS may reflect acoustic-phonetic, conceptual or lexico-semantic retrieval difficulties. Since the time-restricted presentation mode of nTMS did not allow to include any semantically or phonologically related distractor items, these processes cannot be clearly disambiguated. Similarly, the selection of a wrong item and searching behavior may indicate either conceptual or lexico-semantic word finding difficulties or a breakdown of the acoustic-phonetic comprehension stage, and can, therefore not be clearly attributed to just one of these comprehension processes.

For specific error categories, searching behavior and selection of wrong target item could be assigned with a high inter-rater-reliability. Whilst additionally no responses were assigned with high concordance across the two raters for patients, the agreement was only fair for controls. This may be attributable to the low overall occurrence of clear no response errors within the present healthy cohort. Moreover, during no responses, hesitant response behavior may simultaneously be observable. This may explain the attribution of errors to different categories across the raters since rater 2 classified 60.0% of no response errors assigned by rater 1 as hesitations. At the same time, the overall agreement for differentiating stimulation-induced disruptions from no errors irrespective of the type of error category assigned was substantial. Thus, whilst error types may not be attributable to just a single category, classifying deviations in response behavior induced by nTMS was shown to be highly reliable. Still, hesitations seem to be the most uncertain error category as the agreement was limited for controls and non-significant for patients. This is in line with naming-based approaches, in which hesitations errors remain the most subjective error category. Up to date, no programs for objective individual response time analysis are readily available for the present setup without employing third-party programs (Schramm et al., 2020). However, these are due to their high time-extensiveness not feasible for clinical applications. As this study shows, a manual analysis within a third-party program takes on average nearly an hour in healthy subjects and even longer in the patient cohort (mean = 79.3 min). While the decision whether a response was delayed or hesitant is a highly subjective and less accurate one, these errors can still result from stimulation-induced disruptions of the language network. Therefore, they are frequently considered for analysis (Krieg et al., 2016; Ohlerth, Bastiaanse, Negwer, et al., 2021).

For this reason, reaction time analyses, which are standard procedures in psycholinguistic experiments, were performed, since they offer a more objective identification of delays in response time induced by nTMS. Whilst this analysis showed only a slight, yet significant agreement with the analysis of rater 1 for patients and controls, just over a fifth of hesitations identified by the rater were attributed on the basis of a seemingly delayed button press. The largest proportion of hesitations were identified based on halting or indecisive hand movements showing a clear hesitant, yet not necessarily substantially delayed response behavior. Thus, by reducing the dimension of the analysis format to audio tracks more objective measurements of reactions times became possible whilst reluctant hand motions as identifiable within the video recordings could not support the identification of nTMS-induced hesitant errors anymore. Still, approximately half of the seemingly delayed reaction times assigned by rater 1 were concordant with the objective reaction time analysis. Hence, the latter analysis may substantially increase reliability for differentiating delayed responses. However, no rule of thumb or definite cut-off criteria for which delays constitute a clear hesitant response induced by nTMS are consistently described across studies. It is well established that response times not only vary considerably between subjects (Jodzio et al., 2023), but may also change significantly throughout the stimulation exam of a single subject. Based on this, Sollmann et al. (2017) suggested to identify hesitations during nTMS-based naming tasks as delays of at least 200 msec compared to preceding or subsequent items named. Thus, a separate analysis of all hesitations identified by rater 1 and the respectively preceding items was performed. Here, a high concordance across patients and controls was verified for the 15 items identified solely on the basis of delayed button press behavior by rater 1 and the item-subset specific reaction time analysis. Hence, to account for intra-subject variations throughout the stimulation examination, subset-specific cut-off criteria would be required. Consequently, by employing the objective, intra-subject specific identification of delayed responses next to the video-based identification of hesitant hand motions, searching behavior, no responses and selection of a wrong target item, reliability and reproducibility of mapping results may be substantially supported. Still, to make this clinically applicable, a system-integrated approach is necessary as this additional analysis within a third-party program considerably increases analysis duration.

4.4. *Cortical language comprehension relevant sites*

Whilst intraoperative mapping is determined by the cortex area exposed during a craniotomy, the non-invasive nature of nTMS allows to create individual, large-scale maps of language-relevant cortical areas and to examine functional reorganization (Ille et al., 2019; Krieg et al., 2013). Our results highlight the heterogeneity of cortical language areas seemingly involved in comprehension. The areas with the highest error rates were widely distributed across patients and controls. This, furthermore, emphasizes the necessity of individual localization of language function prior to surgery to allow individually guided surgical planning and resection of tumors located in language areas. Still, some common patterns across patients and particularly the larger control group were identifiable: All patients had high error rates and the group of 15 healthy controls had moderate to high error rates within temporal regions. Separate analyses of different error types showed that this pattern occurred particularly for searching behavior. In patient case 1, 4, 5 and 6 as well as across the 15 controls, high error rates were found for areas that are typically described as the classic “Wernicke's” area that is the middle and posterior superior temporal gyrus (Binder, 2017). Across the whole patient group stimulation over these cortical sites elicited no responses and searching behavior, whilst only very few no response errors were found within the controls.

Moreover, five of the patient cases (P1, P2, P3, P5 and P6) and nearly half of the control cases (see Appendix Table B) had additionally moderate to high error rates in the posterior MTG. This cortical region was shown to be a critical language comprehension hub since its connectivity profile of the comprehension network was considerably high especially in comparison to other cortical comprehension areas (Turken & Dronkers, 2011). As stimulation of this area resulted in low to moderate error rates for no response, searching behavior and selection of wrong target item across patients and for searching behavior across healthy participants – all comprehension error types which were identifiable with high reliability – the present results underline the important role of posterior MTG in auditory single-word comprehension. Half of the patients (P2, P4, P6) and one third of the healthy subjects (see Appendix Table B) had additional high error rates in anterior STG. As error type analyses revealed this was mainly driven by moderate error rates for searching behavior in patients and controls. Although this area is frequently associated with word comprehension, partial removal of this section during surgery does not typically cause persistent language impairments (DeWitt & Rauschecker, 2013). However, our findings do not only indicate involvement of these rather classic temporal comprehension regions but also point towards an extensive involvement of frontal and parietal ones across patients and controls, in line with naming-based nTMS language mapping results (Krieg et al., 2016). For instance, 83.3% of the patient cases presented with high error rates within the opercular IFG, and 66.7% of patients had at least a moderate error rate in the triangular IFG. Moreover, apart from three cases, all healthy subjects showed moderate to high error rates within trIFG or opIFG. Whereas the IFG was originally attributed with language production, more recent findings increasingly corroborate its involvement in language production and comprehension, particularly a dissociation of phonological processing in the opercular and semantic processing in the anterior IFG (Gough et al., 2005; Klaus & Hartwigsen, 2019).

4.5. *Surgical perspective*

There is a considerable number of glioma patients with severe impairment of language capabilities but who are still able to communicate and have an independent life. Thus, we need to treat these patients – and surgery is still the most powerful therapeutic option today – but we also need to preserve their limited but useful language capabilities. Hence, the presented mapping workflow and setup helps us to identify the underlying network to preserve this residual function. Some years ago, we still operated those patients awake but realized that their language abilities are worse after craniotomy then the days before and mapping was almost impossible. Having a methodology at hand which allows patients being mapped in a calm, relaxing atmosphere with video-recorded evaluation and a lot of time to tailor tests and setup, helps us to produce a much better visualization of the underlying language network.

4.6. *Limitations and perspectives*

This study is the first to present a comprehension-based nTMS language mapping paradigm for brain tumor patients with severe expressive aphasia. While the results of this case series are promising, all implications are based on a relatively small patient sample size. Inclusion of larger cohort sizes may allow for a more differential error pattern analysis in patients and in-depth group-wise comparisons of cortical error pattern distribution as well as of the subcortical language comprehension network. To identify strong and generalizable differences between patients and controls, a larger sample size needs to be recruited and both groups, moreover, need to be matched in age. In addition, while all six patients included in the present study presented with severe expressive deficits, their language comprehension was additionally impaired to varying extends. Even if the number of items included varied across patients, all patients were able to select the correct target item reliably and reproducibly prior to stimulation. Still, since the presence of severe aphasia may limit the reliability of stimulation-based mappings (Schwarzer et al., 2018), this may have impacted the present mapping results in case 1, 3 and 6. Hence, subsequent studies with a large sample size of patients without or light receptive deficits are warranted. This may substantially advance the understanding of single-word comprehension and functional reorganization in brain tumor patients with severe aphasia. Moreover, conducting thorough qualitative and quantitative analyses of the cortical and subcortical components involved in the four identified different error categories may support the delineation of different comprehension processes in lesioned patient populations.

Furthermore, subsequent studies may compare the functional cortical and subcortical language comprehension network between patients with semantic and phonologic comprehension deficits. A wealth of imaging, stimulation, neurologic and neurodegenerative lesion studies indicate a relevance of anterior for phonological and posterior language network components for semantic processes during comprehension and production as opposed to the classic language models building on Broca's and Wernicke's original work (Butler et al., 2014; Ingram et al., 2020; Klaus & Hartwigsen, 2019; Mesulam et al., 2015; Mirman et al., 2015; Tremblay & Dick, 2016). The present task required patients to select an auditorily presented target item out of a set of four picture stimuli via the push of a button matched in color to the background of the correct target item while stimulation is applied. Since task demands can impact the expression of semantic deficits (Jefferies & Lambon Ralph, 2006), this task may potentially impact the mapping reliability in semantically impaired patients. Moreover, this task demand may explain some of the cortical sites identified as comprehension relevant. For instance, Lambon Ralph et al. (2017) linked inferior frontal and posterior middle temporal activation to executively demanding semantic processes.

In addition, since patient 1 is ambidextrous, his mapping results may not necessarily be directly comparable to right-handed patients and subjects. While a left-hemispheric dominance for language in right-handed people is widely established, in left-handers and to a lesser extent in ambidextrous people, a higher possibility of right-hemispheric language dominance has been reported (Isaacs et al., 2006; Knecht et al., 2000). Performing a bihemipsheric comprehension mapping may additionally provide a more comprehensive picture of the healthy and the impaired comprehension processes across subjects as even in right-handed subjects a right-hemispheric involvement during comprehension has been corroborated (Gajardo-Vidal et al., 2018).

Furthermore, whereas we did not observe any impact of the noise of the stimulation application on the ability of controls and patients to hear the auditory target item, the present setup does not allow to delineate whether next to complex language comprehension processes additional lower-order hearing processes are disrupted by stimulation. However, across all healthy participants and patients heterogeneous left-hemispheric areas were linked to language comprehension many of which are thought to comprise cortical language-relevant ones. Still, a control task testing non-verbal auditory comprehension may allow to delineate these lower-order hearing and complex language comprehension processes.

Moreover, this study did not evaluate the feasibility and utility of this new comprehension-based stimulation language mapping in context of intraoperative awake language monitoring, the gold standard for preservation of language function in language-eloquent brain tumors. Whilst the timing of the task may already fit the time-restricted intraoperative conditions, slight adaptions of screen and button setup might be needed. As the results of the present study already demonstrated, language comprehension disruptions through stimulation are identifiable with a high inter-rater reliability. Since naming-based approaches elicit similar error patterns with nTMS as with DES (Corina et al., 2010; Lioumis et al., 2012; Talacchi et al., 2013), it is expected that comprehension errors under DES resemble the error pattern under nTMS. Consequently, the task itself may easily be transferable into the operating room and may allow for an instant identification of cortical language comprehension sites. This may enable awake surgeries in patients with severe expressive aphasia whose language impairment thus far precluded them from pre- and intraoperative naming-based language mappings, respectively, and may substantially support the preservation of residual language function.

If, furthermore, both stimulation-based language mapping methods were employed, the concordance of identifying relevant and non-relevant cortical language comprehension sites can be evaluated. Thus far, nTMS-based language mappings are known for their high sensitivity and negative predictive values compared to the gold standard, particularly in comparison to other preoperative functional neuroimaging methods (Ille et al., 2015; Picht et al., 2013; Tarapore et al., 2013). Hence, the largest limitation of nTMS-based language mappings remains the poor specificity which results in a high reliance on negative mapping results. Still, the present study did not only show a high inter-rater reliability for the differentiation of comprehension positive and negative sites but also the induction of multiple errors for the stimulation of the same cortical sites. This in combination with preserved residual abilities supports the reliability of the current mapping approach. However, to confirm that these sites are indeed functionally relevant, direct comparisons with intraoperative stimulation mapping, the gold standard for the preservation of functionality (Duffau, 2015), would be required. While it is not yet the standard of care, nTMS is increasingly integrated into the preoperative workflow across many centers especially in cases for whom DES-based mappings are not an option (Ille, Sollmann, et al., 2016; Raffa et al., 2022).

At the same time, due to methodological and technical limitations, contemporary white matter imaging techniques thus far inaccurately represent the anatomical network (Catani et al., 2013). The results of Maier-Hein et al. (2017) show that DTI-based tractographies are limited by reconstructing a large proportion of non-existing, anatomically non-valid tracts while the reconstruction contained approximately 90% of anatomical valid connections. Conversely, preclinical imaging studies show a high rate of anatomically valid tracts missed by DTI tractography (Aydogan et al., 2018; Grisot et al., 2021). Hence, the potential of false positive or negative subcortical reconstructions may result in the preservation functionally non-relevant or the resection of relevant subcortical tracts impacting life expectancy or a patient's quality of life (Brown et al., 2016; Hervey-Jumper & Berger, 2016). The most direct way to investigate the functional role of the subcortical language network in neurosurgical patients remains subcortical stimulation during awake surgeries (Duffau, 2015) which is not impacted by the methodological and technical limitations of the preoperative techniques presented within the current study. Nevertheless, DTI offers a unique, in-vivo and non-invasive way to investigate subcortical connections. Multiple studies show that by using nTMS-based cortical sites to derive DTI-based tractographies of the functional language network, the preservation of functionality and preoperative risk stratification can be supported already non-invasively prior to a resection (Giampiccolo et al., 2020; Ille et al., 2018; Raffa et al., 2019; Sollmann, Zhang, Fratini, et al., 2020).

5. Conclusion

This study tested the feasibility and reliability of a new non-overt language-comprehension task for pre- as well as intraoperative language mapping. The task was feasible and its analysis highly reliable for patient and control data. The present setup not only allowed a language mapping in patients with severe expressive aphasia thus far precluded from classic overt language-production based mapping, but also enabled to preserve residual language function if the results were employed in combination with function-based tractography for surgical planning and resection.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors, it was completely financed by institutional grants from the Department of Neurosurgery of the Klinikum rechts der Isar, Technical University of Munich.

CRediT author statement

Leonie Kram, MSc: Conceptualization; Methodology; Validation; Formal analysis; Investigation; Data curation; Writing – original draft. Ann-Katrin Ohlerth, PhD: Conceptualization; Methodology; Validation; Formal analysis; Writing – review & editing. Sebastian Ille, MD: Writing – review & editing; Project administration. Bernhard Meyer, MD: Writing – review & editing; Project administration. Sandro M. Krieg, MD, MBA: Conceptualization; Methodology; Validation; Resources; Data curation; Writing – review & editing; Supervision; Project administration.

Data availability statement

All data required to reproduce the reported analyses appear in the figures of this article, the error rates for individual subjects and per group are provided within the Appendix. The data is stored in an institutional repository and not publicly available due to hospital legislation and medical ethical objections. All data presented in this study are available upon reasonable request, access will be granted to named individuals in accordance with ethical procedures governing the reuse of sensitive data, i.e., if the ethical committee approves and if the data sharing agreement is signed by both a demanding and providing party. Readers seeking access to the data are advised to contact the corresponding author, Prof. Dr. med. S.M. Krieg.

Legal copyright restrictions do not permit us to publicly archive or share the full set of stimuli used in this experiment in a trusted digital repository. All items used stem from the "verb and noun test for peri-operative testing (Van-POP)", the copyright of the pictures belongs to the Rijksuniversiteit Groningen, the Netherlands. Readers seeking access to the stimuli are advised to contact the author, Dr. A.-K. Ohlerth [Ann-Katrin.Ohlerth@mpi.nl; ann.katrin.ohlerth@gmail.com]. Stimuli will be released if the declaration of usage agreement is signed, and all points of the agreement are followed closely. This agreement can be received upon request.

Code availability

The data set for reproducing the inter-rater analysis together with the respective R analysis code is available within the supplementary data.

Declaration of competing interest

B.M. received honoraria, consulting fees, and research grants from Medtronic (Meerbusch, Germany), Icotec ag (Altstätten, Switzerland), and Relievant Medsystemy Inc. (Sunnyvale, CA, USA), honoraria and research grants from Ulrich Medical (Ulm, Germany), honoraria and consulting fees from Spineart Deutschland GmbH (Frankfurt, Germany) and DePuy Synthes (West Chester, PA, USA), royalties from Spineart Deutschland GmbH (Frankfurt, Germany). S.M.K. is consultant for Ulrich medical (Ulm, Germany) and Brainlab AG (Munich, Germany) and received honoraria from Nexstim Plc (Helsinki, Finland), Spineart Deutschland GmbH (Frankfurt, Germany), Medtronic (Meerbusch, Germany) and Carl Zeiss Meditec (Oberkochen, Germany). BM received research grants and is consultant for Brainlab AG (Munich, Germany). SI is consultant for Brainlab AG (Munich, Germany). None of the authors state any conflict of interest.

Acknowledgments

We would like to express our gratitude to Prof. Simon N Jacob (Technical University of Munich) for his valuable advice on the design of the comprehension-testing setup as well as his constructive comments on the initial version of the manuscript. We are grateful to all patients and subjects who participated in this study.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cortex.2023.09.023.

Appendix A. Stimulation-induced error rates (number or errors/number of stimulations applied) across all error types of patients and controls for each parcellated cortical region.

CPS region specific stimulation-induced error rates across all error types for patients.

Cortical area	Patient						Total[a]
	P1	P2	P3	P4	P5	P6	
anG	.25 (3/12)	.14 (2/14)	.25 (3/12)	.07 (1/14)	.17 (2/12)	.42 (5/12)	.22 (16/76)
aSMG	.00 (0/6)	.17 (1/6)	.33 (2/6)	.17 (1/6)	.25 (2/8)	.00 (0/6)	.15 (6/38)
aSTG	.00 (0/3)	.33 (1/3)	.00 (0/2)	.33 (1/3)	.00 (0/1)	.33 (1/3)	.17 (3/15)
dPoG	.00 (0/3)	.00 (0/3)	.00 (0/3)	.00 (0/3)	.00 (0/3)	.67 (2/3)	.11 (2/18)
dPrG	.00 (0/3)	.33 (1/3)	.00 (0/3)	.25 (1/4)	.00 (0/3)	.00 (0/3)	.10 (2/19)
mMFG	.15 (2/13)	.15 (3/20)	.35 (6/17)	.00 (0/18)	.25 (5/20)	.06 (1/18)	.16 (17/106)
mMTG	.17 (1/6)	.17 (1/6)	.00 (0/6)	.25 (2/8)	.00 (0/6)	.17 (1/6)	.13 (5/38)
mPoG	.17 (1/6)	.33 (2/6)	.00 (0/6)	.17 (1/6)	.00 (0/6)	.33 (2/6)	.17 (6/36)
mPrG	.13 (1/6)	.00 (0/4)	.17 (1/6)	.17 (1/6)	.00 (0/3)	.50 (3/6)	.16 (6/31)
mSFG	/ (0/0)	.29 (2/7)	.20 (2/10)	.00 (0/9)	.25 (1/4)	.00 (0/9)	.15 (5/39)
mSTG	.17 (1/6)	.00 (0/3)	.00 (0/6)	.00 (0/6)	.25 (2/8)	.67 (4/6)	.18 (7/35)
opIFG	.17 (1/6)	.33 (2/6)	.33 (2/6)	.33 (2/6)	.43 (3/7)	.33 (2/6)	.32 (12/37)
pMFG	.00 (0/6)	.20 (1/5)	.00 (0/6)	.14 (1/7)	.00 (0/1)	.17 (1/6)	.08 (3/31)
pMTG	.22 (2/9)	.17 (2/12)	.30 (3/10)	.00 (0/6)	.50 (5/10)	.33 (3/9)	.25 (15/56)
pSFG	.00 (0/3)	.00 (0/2)	.00 (0/3)	.00 (0/3)	.50 (1/2)	.00 (0/3)	.08 (1/16)
pSMG	.17 (1/6)	.00 (0/6)	.33 (2/6)	.17 (1/6)	.50 (3/6)	.33 (2/6)	.25 (9/36)
pSTG	.33 (1/3)	.00 (0/3)	.00 (0/3)	.33 (1/3)	.20 (1/5)	.67 (2/3)	.26 (5/20)
SPL	.33 (2/6)	.33 (2/6)	.00 (0/6)	.00 (0/6)	.00 (0/6)	.00 (0/6)	.11 (4/36)
trIFG	.17 (1/6)	.29 (2/7)	.00 (0/8)	.00 (0/6)	.13 (1/8)	.11 (1/9)	.11 (5/44)
vPoG	.17 (1/6)	.33 (2/6)	.50 (3/6)	.33 (2/6)	.00 (0/3)	.17 (1/6)	.25 (9/33)
vPrG	.00 (0/6)	.33 (2/6)	.00 (0/6)	.17 (1/6)	.00 (0/3)	.67 (4/6)	.20 (7/33)

[a] The total error rate for each cortically parcellated area was calculated as the mean of error rates for each cortical area across patients, thus, slight deviations from calculations on the basis of the absolute numbers may be present.

CPS region specific stimulation-induced error rates across all error types for controls (C1 - C9).

Cortical area	Controls								
	C1	C2	C3	C4	C5	C6	C7	C8	C9
anG	.00	.00	.00	.08	.17	.17	.08	.00	.00
aSMG	.00	.00	.20	.17	.00	.00	.00	.17	.00
aSTG	.67	.00	.33	.00	.00	.00	.00	.00	.33
dPoG	.00	.00	.00	.00	.00	.00	.00	.33	.33
dPrG	.33	.00	.00	.33	.33	.00	.00	.00	.00
mMFG	.17	.07	.00	.06	.22	.06	.17	.11	.00
mMTG	.00	.00	.00	.17	.00	.17	.33	.17	.00
mPoG	.17	.00	.00	.00	.33	.33	.00	.00	.00
mPrG	.00	.00	.17	.17	.17	.00	.17	.17	.00
mSFG	.11	.17	.00	.00	.33	.33	.00	.00	.00
mSTG	.00	.00	.00	.00	.00	.50	.17	.00	.17
opIFG	.17	.00	.00	.17	.17	.17	.11	.33	.00
pMFG	.17	.00	.17	.00	.00	.17	.17	.17	.17
pMTG	.11	.00	.00	.00	.44	.11	.00	.17	.00
pSFG	.00	.00	.00	.00	.00	.00	.67	.33	.00
pSMG	.17	.17	.00	.00	.00	.00	.17	.00	.17
pSTG	.00	.00	.00	.00	.00	.00	.67	.00	.33
SPL	.00	.17	.00	.17	.17	.00	.00	.00	.00
trIFG	.11	.00	.11	.22	.11	.00	.17	.11	.00
vPoG	.00	.00	.17	.00	.17	.17	.00	.11	.00
vPrG	.17	.00	.17	.17	.33	.17	.00	.00	.00

CPS region specific stimulation-induced error rates across all error types for controls (C10 - C15).

Cortical area	Controls						Total
	C10	C11	C12	C13	C14	C15	
anG	.08	.08	.00	.17	.00	.08	.06
aSMG	.00	.17	.00	.17	.00	.33	.08
aSTG	.00	.33	.00	.00	.33	.00	.13
dPoG	.33	.00	.00	.00	.00	.00	.07
dPrG	.33	.33	.00	.00	.00	.00	.11
mMFG	.06	.17	.11	.06	.11	.00	.09
mMTG	.00	.50	.17	.67	.50	.17	.19
mPoG	.00	.17	.17	.00	.33	.00	.10
mPrG	.00	.17	.00	.00	.17	.00	.08
mSFG	.00	.00	.00	.00	.11	.00	.07

(continued)

Cortical area	Controls						Total
	C10	C11	C12	C13	C14	C15	
mSTG	.17	.17	.00	.17	.00	.17	.10
opIFG	.17	.33	.17	.33	.00	.00	.14
pMFG	.33	.17	.33	.00	.00	.17	.13
pMTG	.00	.17	.00	.00	.67	.17	.12
pSFG	.00	.00	.00	.00	.00	.00	.07
pSMG	.00	.17	.00	.17	.17	.00	.08
pSTG	.00	.67	.33	.00	.33	.33	.18
SPL	.00	.17	.00	.00	.00	.17	.06
trIFG	.00	.44	.11	.33	.11	.00	.12
vPoG	.11	.22	.11	.11	.00	.22	.09
vPrG	.17	.17	.00	.17	.00	.33	.12

Appendix B. Error type specific error rates across patient- and control-group

Cortical area	Group	Error type[a]			
		NR	SB	SW	H
anG	Patient	.10	.08	.04	.00
	Control	.00	.04	.00	.02
aSMG	Patient	.03	.05	.05	.03
	Control	.01	.03	.00	.04
aSTG	Patient	.06	.11	.00	.00
	Control	.00	.09	.00	.04
dPoG	Patient	.06	.00	.06	.00
	Control	.00	.04	.00	.02
dPrG	Patient	.00	.06	.04	.00
	Control	.02	.07	.02	.00
mMFG	Patient	.06	.05	.04	.01
	Control	.00	.06	.00	.03
mMTG	Patient	.03	.08	.02	.00
	Control	.01	.16	.00	.02
mPoG	Patient	.03	.08	.06	.00
	Control	.00	.07	.00	.03
mPrG	Patient	.00	.08	.03	.03
	Control	.01	.03	.00	.03
mSFG	Patient	.05	.06	.06	.00
	Control	.00	.05	.00	.03
mSTG	Patient	.10	.06	.02	.00
	Control	.00	.09	.00	.01
opIFG	Patient	.03	.19	.10	.00
	Control	.00	.10	.00	.04
pMFG	Patient	.00	.08	.00	.00
	Control	.00	.08	.01	.04
pMTG	Patient	.05	.08	.10	.02
	Control	.00	.10	.00	.02
pSFG	Patient	.08	.00	.00	.00
	Control	.00	.04	.00	.02
pSMG	Patient	.14	.00	.08	.03
	Control	.00	.04	.00	.03
pSTG	Patient	.11	.11	.03	.00
	Control	.00	.11	.00	.07
SPL	Patient	.03	.06	.03	.00
	Control	.00	.03	.00	.02
trIFG	Patient	.02	.03	.05	.03
	Control	.00	.10	.00	.02
vPoG	Patient	.11	.08	.00	.06
	Control	.00	.08	.00	.02
vPrG	Patient	.05	.11	.03	.00
	Control	.00	.10	.00	.02

[a] NR = no response, SB = searching behavior, SW = selection of wrong target, H = hesitant responses.

REFERENCES

Alarcon, G., Bird Pedersen, M., Juarez-Torrejon, N., Martin-Lopez, D., Ughratdar, I., Selway, R. P., & Valentin, A. (2019). The single word auditory comprehension (SWAC) test: A simple method to identify receptive language areas with electrical stimulation. *Epilepsy & Behavior, 90*, 266–272. https://doi.org/10.1016/j.yebeh.2018.10.022

Aydogan, D. B., Jacobs, R., Dulawa, S., Thompson, S. L., Francois, M. C., Toga, A. W., Dong, H., Knowles, J. A., & Shi, Y. (2018). When tractography meets tracer injections: A systematic study of trends and variation sources of diffusion-based connectivity. *Brain Structure & Function, 223*(6), 2841–2858. https://doi.org/10.1007/s00429-018-1663-8

Bährend, I., Muench, M. R., Schneider, H., Moshourab, R., Dreyer, F. R., Vajkoczy, P., Picht, T., & Faust, K. (2020). Incidence and linguistic quality of speech errors: A comparison of preoperative transcranial magnetic stimulation and intraoperative direct cortex stimulation. *Journal of Neurosurgery, 134*(5), 1409–1418. https://doi.org/10.3171/2020.3.JNS193085

Bangdiwala, S. I. (1988). *The agreement chart*. The University of North Carolina.

Bangdiwala, S. I., & Shankar, V. (2013). The agreement chart. *BMC Medical Research Methodology, 13*, 97. https://doi.org/10.1186/1471-2288-13-97

Bello, L., Gallucci, M., Fava, M., Carrabba, G., Giussani, C., Acerbi, F., … Gaini, S. M. (2007). Intraoperative subcortical language tract mapping guides surgical removal of gliomas involving speech areas. *Neurosurgery, 60*(1), 67–82. https://doi.org/10.1227/01.NEU.0000249206.58601.DE

Binder, J. R. (2017). Current controversies on Wernicke's area and its role in language. *Current Neurology and Neuroscience Reports, 17*(8), 58. https://doi.org/10.1007/s11910-017-0764-8

Boersma, P., & Weenink, D. (2023). *Praat: Doing phonetics by computer [computer program]*. version 6.3.04. Retrieved 24 January 2023 from http://www.praat.org/.

Bornkessel-Schlesewsky, I., Staub, A., & Schlesewsky, M. (2016). The timecourse of sentence processing in the brain. In G. Hickok, & S. L. Small (Eds.), *Neurobiology of language* (pp. 607–620). https://doi.org/10.1016/b978-0-12-407794-2.00049-3

Bornkessel, I., Zysset, S., Friederici, A. D., von Cramon, D. Y., & Schlesewsky, M. (2005). Who did what to whom? The neural basis of argument hierarchies during language comprehension. *NeuroImage, 26*(1), 221–233. https://doi.org/10.1016/j.neuroimage.2005.01.032

Brown, T. J., Brennan, M. C., Li, M., Church, E. W., Brandmeir, N. J., Rakszawski, K. L., Patel, A. S., Rizk, E. B., Suki, D., Sawaya, R., & Glantz, M. (2016). Association of the extent of resection with survival in glioblastoma: A systematic review and meta-

analysis. *JAMA Oncology, 2*(11), 1460–1469. https://doi.org/10.1001/jamaoncol.2016.1373
Butler, R. A., Lambon Ralph, M. A., & Woollams, A. M. (2014). Capturing multidimensionality in stroke aphasia: Mapping principal behavioural components to neural structures. *Brain, 137*(12), 3248–3266. https://doi.org/10.1093/brain/awu286
Catani, M., Thiebaut de Schotten, M., Slater, D., & Dell'Acqua, F. (2013). Connectomic approaches before the connectome. *NeuroImage, 80*, 2–13. https://doi.org/10.1016/j.neuroimage.2013.05.109
Chang, E. F., Raygor, K. P., & Berger, M. S. (2015). Contemporary model of language organization: An overview for neurosurgeons. *Journal of Neurosurgery, 122*(2), 250–261. https://doi.org/10.3171/2014.10.JNS132647
Corina, D. P., Gibson, E. K., Martin, R., Poliakov, A., Brinkley, J., & Ojemann, G. A. (2005). Dissociation of action and object naming: Evidence from cortical stimulation mapping. *Human Brain Mapping, 24*(1), 1–10. https://doi.org/10.1002/hbm.20063
Corina, D. P., Loudermilk, B. C., Detwiler, L., Martin, R. F., Brinkley, J. F., & Ojemann, G. (2010). Analysis of naming errors during cortical stimulation mapping: Implications for models of language representation. *Brain and Language, 115*(2), 101–112. https://doi.org/10.1016/j.bandl.2010.04.001
De Renzi, E., & Vignolo, L. A. (1962). The token test: A sensitive test to detect receptive disturbances in aphasics. *Brain, 85*, 665–678. https://doi.org/10.1093/brain/85.4.665
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
De Witte, E., Satoer, D., Robert, E., Colle, H., Verheyen, S., Visch-Brink, E., & Marien, P. (2015). The Dutch linguistic intraoperative protocol: A valid linguistic approach to awake brain surgery. *Brain and Language, 140*, 35–48. https://doi.org/10.1016/j.bandl.2014.10.011
DeWitt, I., & Rauschecker, J. P. (2012). Phoneme and word recognition in the auditory ventral stream. *Proceedings of the National Academy of Sciences of the United States of America, 109*(8), E505–E514. https://doi.org/10.1073/pnas.1113427109
DeWitt, I., & Rauschecker, J. P. (2013). Wernicke's area revisited: Parallel streams and word processing. *Brain and Language, 127*(2), 181–191. https://doi.org/10.1016/j.bandl.2013.09.014
DeWitt, I., & Rauschecker, J. P. (2016). Convergent evidence for the causal involvement of anterior superior temporal gyrus in auditory single-word comprehension. *Cortex, 77*, 164–166. https://doi.org/10.1016/j.cortex.2015.08.016
Duffau, H. (2015). Stimulation mapping of white matter tracts to study brain functional connectivity. *Nature Reviews Neurology, 11*(5), 255–265. https://doi.org/10.1038/nrneurol.2015.51
Duffau, H., & Mandonnet, E. (2013). The "onco-functional balance" in surgery for diffuse low-grade glioma: Integrating the extent of resection with quality of life. *Acta Neurochirurgica, 155*(6), 951–957. https://doi.org/10.1007/s00701-013-1653-9
Eckstein, K., & Friederici, A. D. (2006). It's early: Event-related potential evidence for initial interaction of syntax and prosody in speech comprehension. *Journal of Cognitive Neuroscience, 18*(10), 1696–1711. https://doi.org/10.1162/jocn.2006.18.10.1696
Faulkner, J. W., Wilshire, C. E., Parker, A. J., & Cunningham, K. (2017). An evaluation of language in brain tumor patients using a new cognitively motivated testing protocol. *Neuropsychology, 31*(6), 648–665. https://doi.org/10.1037/neu0000374
Fernandez Coello, A., Moritz-Gasser, S., Martino, J., Martinoni, M., Matsuda, R., & Duffau, H. (2013). Selection of intraoperative tasks for awake mapping based on relationships between tumor location and functional networks. *Journal of Neurosurgery, 119*(6), 1380–1394. https://doi.org/10.3171/2013.6.JNS122470
Fridriksson, J., Fillmore, P., Guo, D., & Rorden, C. (2015). Chronic Broca's aphasia is caused by damage to Broca's and Wernicke's areas. *Cerebral Cortex, 25*(12), 4689–4696. https://doi.org/10.1093/cercor/bhu152
Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences, 6*(2), 78–84.
Friederici, A. D. (2011). The brain basis of language processing: From structure to function. *Physiological Reviews, 91*(4), 1357–1392. https://doi.org/10.1152/physrev.00006.2011
Friederici, A. D. (2012). The cortical language circuit: From auditory perception to sentence comprehension. *Trends in Cognitive Sciences, 16*(5), 262–268. https://doi.org/10.1016/j.tics.2012.04.001
Friederici, A. D. (2017). *Language in our brain: The origins of a uniquely human capacity*. MIT Press.
Gajardo-Vidal, A., Lorca-Puls, D. L., Hope, T. M. H., Parker Jones, O., Seghier, M. L., Prejawa, S., Crinion, J. T., Leff, A. P., Green, D. W., & Price, C. J. (2018). How right hemisphere damage after stroke can impair speech comprehension. *Brain, 141*(12), 3389–3404. https://doi.org/10.1093/brain/awy270
Gamer, M., Lemon, J., & Singh, I. F. P. (2019). *irr: Various coefficients of interrater reliability and agreement*. R package version 0.84.1 https://CRAN.R-project.org/package=irr.
Gatignol, P., Capelle, L., Le Bihan, R., & Duffau, H. (2004). Double dissociation between picture naming and comprehension: An electrostimulation study. *NeuroReport, 15*(1), 191–195. https://doi.org/10.1097/01.wnr.0000099474.09597.6a
Getz, L. M., & Toscano, J. C. (2021). The time-course of speech perception revealed by temporally-sensitive neural measures. *Wiley Interdisciplinary Reviews: Cognitive Science, 12*(2), e1541. https://doi.org/10.1002/wcs.1541
Giampiccolo, D., Howells, H., Bährend, I., Schneider, H., Raffa, G., Rosenstock, T., Vergani, F., Vajkoczy, P., & Picht, T. (2020). Preoperative transcranial magnetic stimulation for picture naming is reliable in mapping segments of the arcuate fasciculus. *Brain Communications, 2*(2), 1–15. https://doi.org/10.1093/braincomms/fcaa158
Gogos, A. J., Young, J. S., Morshed, R. A., Hervey-Jumper, S. L., & Berger, M. S. (2020). Awake glioma surgery: Technical evolution and nuances. *Journal of Neuro-oncology, 147*(3), 515–524. https://doi.org/10.1007/s11060-020-03482-z
Gough, P. M., Nobre, A. C., & Devlin, J. T. (2005). Dissociating linguistic processes in the left inferior frontal cortex with transcranial magnetic stimulation. *The Journal of Neuroscience: the Official Journal of the Society for Neuroscience, 25*(35), 8010–8016. https://doi.org/10.1523/JNEUROSCI.2307-05.2005
Grisot, G., Haber, S. N., & Yendiki, A. (2021). Diffusion MRI and anatomic tracing in the same brain reveal common failure modes of tractography. *NeuroImage, 239*, 118300. https://doi.org/10.1016/j.neuroimage.2021.118300
Haddad, A. F., Young, J. S., Berger, M. S., & Tarapore, P. E. (2021). Preoperative applications of navigated transcranial magnetic stimulation. *Frontiers in Neurology, 11*, 628903. https://doi.org/10.3389/fneur.2020.628903
Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science, 304*(5669), 438–441. https://doi.org/10.1126/science.1095455
Hauck, T., Tanigawa, N., Probst, M., Wohlschlaeger, A., Ille, S., Sollmann, N., Maurer, S., Zimmer, C., Ringel, F., Meyer, B., & Krieg, S. M. (2015). Task type affects location of language-positive cortical regions by repetitive navigated transcranial magnetic stimulation mapping. *PLoS One, 10*(4), Article e0125298. https://doi.org/10.1371/journal.pone.0125298

Hendrix, P., Senger, S., Simgen, A., Griessenauer, C. J., & Oertel, J. (2017). Preoperative rTMS language mapping in speech-eloquent brain lesions resected under general anesthesia: A pair-matched cohort study. *World Neurosurgery, 100*, 425–433. https://doi.org/10.1016/j.wneu.2017.01.041

Herrmann, B., Maess, B., Hasting, A. S., & Friederici, A. D. (2009). Localization of the syntactic mismatch negativity in the temporal cortex: An MEG study. *NeuroImage, 48*(3), 590–600. https://doi.org/10.1016/j.neuroimage.2009.06.082

Hervey-Jumper, S. L., & Berger, M. S. (2016). Maximizing safe resection of low- and high-grade glioma. *Journal of Neuro-oncology, 130*(2), 269–282. https://doi.org/10.1007/s11060-016-2110-4

Huber, W., Poeck, K., & Springer, L. (1983). *Aachener Aphasietest (AAT)*. Hogrefe.

IJzerman-Korevaar, M., Snijders, T. J., de Graeff, A., Teunissen, S., & de Vos, F. Y. F. (2018). Prevalence of symptoms in glioma patients throughout the disease trajectory: A systematic review. *Journal of Neuro-oncology, 140*(3), 485–496. https://doi.org/10.1007/s11060-018-03015-9

Ille, S., Engel, L., Albers, L., Schroeder, A., Kelm, A., Meyer, B., & Krieg, S. M. (2019). Functional reorganization of cortical language function in glioma patients – A preliminary study. *Frontiers in Oncology, 9*, 446. https://doi.org/10.3389/fonc.2019.00446

Ille, S., Engel, L., Kelm, A., Meyer, B., & Krieg, S. M. (2018). Language-eloquent white matter pathway tractography and the course of language function in glioma patients. *Frontiers in Oncology, 8*, 572. https://doi.org/10.3389/fonc.2018.00572

Ille, S., Kulchytska, N., Sollmann, N., Wittig, R., Beurskens, E., Butenschoen, V. M., Ringel, F., Vajkoczy, P., Meyer, B., Picht, T., & Krieg, S. M. (2016a). Hemispheric language dominance measured by repetitive navigated transcranial magnetic stimulation and postoperative course of language function in brain tumor patients. *Neuropsychologia, 91*, 50–60. https://doi.org/10.1016/j.neuropsychologia.2016.07.025

Ille, S., Schroeder, A., Albers, L., Kelm, A., Droese, D., Meyer, B., & Krieg, S. M. (2021). Non-invasive mapping for effective preoperative guidance to approach highly language-eloquent gliomas—A large scale comparative cohort study using a new classification for language eloquence. *Cancers, 13*(2), 207. https://doi.org/10.3390/cancers13020207

Ille, S., Sollmann, N., Butenschoen, V. M., Meyer, B., Ringel, F., & Krieg, S. M. (2016b). Resection of highly language-eloquent brain lesions based purely on rTMS language mapping without awake surgery. *Acta Neurochirurgica, 158*(12), 2265–2275. https://doi.org/10.1007/s00701-016-2968-0

Ille, S., Sollmann, N., Hauck, T., Maurer, S., Tanigawa, N., Obermueller, T., Negwer, C., Droese, D., Zimmer, C., Meyer, B., Ringel, F., & Krieg, S. M. (2015). Combined noninvasive language mapping by navigated transcranial magnetic stimulation and functional MRI and its comparison with direct cortical stimulation. *Journal of Neurosurgery, 123*(1), 212–225. https://doi.org/10.3171/2014.9.JNS14929

Ingram, R. U., Halai, A. D., Pobric, G., Sajjadi, S., Patterson, K., & Lambon Ralph, M. A. (2020). Graded, multidimensional intra- and intergroup variations in primary progressive aphasia and post-stroke aphasia. *Brain, 143*(10), 3121–3135. https://doi.org/10.1093/brain/awaa245

Isaacs, K. L., Barr, W. B., Nelson, P. K., & Devinsky, O. (2006). Degree of handedness and cerebral dominance. *Neurology, 66*(12), 1855–1858. https://doi.org/10.1212/01.wnl.0000219623.28769.74

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

Jodzio, A., Piai, V., Verhagen, L., Cameron, I., & Indefrey, P. (2023). Validity of chronometric TMS for probing the time-course of word production: A modified replication. *Cerebral Cortex*. https://doi.org/10.1093/cercor/bhad081

Kertesz, A. (2007). *WAB-R: Western aphasia battery-revised*. The Psychological Corporation.

Klaus, J., & Hartwigsen, G. (2019). Dissociating semantic and phonological contributions of the left inferior frontal gyrus to language production. *Human Brain Mapping, 40*(11), 3279–3287. https://doi.org/10.1002/hbm.24597

Knecht, S., Dräger, B., Deppe, M., Bobe, L., Lohmann, H., Flöel, A., Ringelstein, E.-B., & Henningsen, H. (2000). Handedness and hemispheric language dominance in healthy humans. *Brain, 123*, 2512–2518. https://doi.org/10.1093/brain/123.12.2512

Krieg, S. M., Lioumis, P., Mäkelä, J. P., Wilenius, J., Karhu, J., Hannula, H., Savolainen, P., Lucas, C. W., Seidel, K., Laakso, A., Islam, M., Vaalto, S., Lehtinen, H., Vitikainen, A. M., Tarapore, P. E., & Picht, T. (2017). Protocol for motor and language mapping by navigated TMS in patients and healthy volunteers; workshop report. *Acta Neurochirurgica, 159*(7), 1187–1195. https://doi.org/10.1007/s00701-017-3187-z

Krieg, S. M., Sollmann, N., Hauck, T., Ille, S., Foerschler, A., Meyer, B., & Ringel, F. (2013). Functional language shift to the right hemisphere in patients with languagelanguage-eloquent brain tumors. *PLoS One, 8*(9), Article e75403. https://doi.org/10.1371/journal.pone.0075403

Krieg, S. M., Sollmann, N., Tanigawa, N., Foerschler, A., Meyer, B., & Ringel, F. (2016). Cortical distribution of speech and language errors investigated by visual object naming and navigated transcranial magnetic stimulation. *Brain Structure & Function, 221*(4), 2259–2286. https://doi.org/10.1007/s00429-015-1042-7

Lambon Ralph, M. A., Jefferies, E., Patterson, K., & Rogers, T. T. (2017). The neural and computational bases of semantic cognition. *Nature Reviews. Neuroscience, 18*(1), 42–55. https://doi.org/10.1038/nrn.2016.150

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159–174. https://doi.org/10.2307/2529310

Lioumis, P., Zhdanov, A., Mäkelä, N., Lehtinen, H., Wilenius, J., Neuvonen, T., Hannula, H., Deletis, V., Picht, T., & Mäkelä, J. P. (2012). A novel approach for documenting naming errors induced by navigated transcranial magnetic stimulation. *Journal of Neuroscience Methods, 204*(2), 349–354. https://doi.org/10.1016/j.jneumeth.2011.11.003

Maier-Hein, K. H., Neher, P. F., Houde, J. C., Cote, M. A., Garyfallidis, E., Zhong, J., Chamberland, M., Yeh, F. C., Lin, Y. C., Ji, Q., Reddick, W. E., Glass, J. O., Chen, D. Q., Feng, Y., Gao, C., Wu, Y., Ma, J., He, R., Li, Q., … Descoteaux, M. (2017). The challenge of mapping the human connectome based on diffusion tractography. *Nature Communications, 8*(1), 1349. https://doi.org/10.1038/s41467-017-01285-x

Mandonnet, E., Winkler, P. A., & Duffau, H. (2010). Direct electrical stimulation as an input gate into brain functional networks: Principles, advantages and limitations. *Acta Neurochirurgica, 152*(2), 185–193. https://doi.org/10.1007/s00701-009-0469-0

Martin-Monzon, I., Rivero Ballagas, Y., & Arias-Sanchez, S. (2022). Language mapping: A systematic review of protocols that evaluate linguistic functions in awake surgery. *Applied Neuropsychology Adult, 29*(4), 845–854. https://doi.org/10.1080/23279095.2020.1776287

Mesulam, M. M., Rader, B. M., Sridhar, J., Nelson, M. J., Hyun, J., Rademaker, A., Geula, C., Bigio, E. H., Thompson, C. K., Gefen, T. D., Weintraub, S., & Rogalski, E. J. (2019). Word comprehension in temporal cortex and Wernicke area: A PPA perspective. *Neurology, 92*(3), e224–e233. https://doi.org/10.1212/WNL.0000000000006788

Mesulam, M. M., Thompson, C. K., Weintraub, S., & Rogalski, E. J. (2015). The Wernicke conundrum and the anatomy of language comprehension in primary progressive aphasia. *Brain, 138*(8), 2423–2437. https://doi.org/10.1093/brain/awv154

Mirman, D., Chen, Q., Zhang, Y., Wang, Z., Faseyitan, O. K., Coslett, H. B., & Schwartz, M. F. (2015). Neural organization of spoken language revealed by lesion-symptom mapping. *Nature Communications, 6*, 6762. https://doi.org/10.1038/ncomms7762

Natalizi, F., Piras, F., Vecchio, D., Spalletta, G., & Piras, F. (2022). Preoperative navigated transcranial magnetic stimulation: New insight for brain tumor-related language mapping. *Journal of Personalized Medicine, 12*(10), 1589. https://doi.org/10.3390/jpm12101589

Negwer, C., Ille, S., Hauck, T., Sollmann, N., Maurer, S., Kirschke, J. S., Ringel, F., Meyer, B., & Krieg, S. M. (2017). Visualization of subcortical language pathways by diffusion tensor imaging fiber tracking based on rTMS language mapping. *Brain Imaging and Behavior, 11*(3), 899–914. https://doi.org/10.1007/s11682-016-9563-0

Ohlerth, A.-K., Bastiaanse, R., Negwer, C., Sollmann, N., Schramm, S., Schröder, A., & Krieg, S. M. (2021a). Bihemispheric navigated transcranial magnetic stimulation mapping for action naming compared to object naming in sentence context. *Brain Sciences, 11*(9), 1190. https://doi.org/10.3390/brainsci11091190

Ohlerth, A.-K., Bastiaanse, R., Nickels, L., Neu, B., Zhang, W., Ille, S., Sollmann, N., & Krieg, S. M. (2021b). Dual-task nTMS mapping to visualize the cortico-subcortical language network and capture postoperative outcome – A patient series in Neurosurgery. *Frontiers in Oncology, 11*, Article 788122. https://doi.org/10.3389/fonc.2021.788122

Ohlerth, A. K., Valentin, A., Vergani, F., Ashkan, K., & Bastiaanse, R. (2020). The verb and noun test for peri-operative testing (VAN-POP): standardized language tests for navigated transcranial magnetic stimulation and direct electrical stimulation. *Acta Neurochirurgica, 162*(2), 397–406. https://doi.org/10.1007/s00701-019-04159-x

Okada, K., Rong, F., Venezia, J., Matchin, W., Hsieh, I. H., Saberi, K., Serences, J. T., & Hickok, G. (2010). Hierarchical organization of human auditory cortex: Evidence from acoustic invariance in the response to intelligible speech. *Cerebral Cortex, 20*(10), 2486–2495. https://doi.org/10.1093/cercor/bhp318

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychology, 9*(1), 97–113. https://doi.org/10.1016/0028-3932(71)90067-4

Ottenhausen, M., Krieg, S. M., Meyer, B., & Ringel, F. (2015). Functional preoperative and intraoperative mapping and monitoring: Increasing safety and efficacy in glioma surgery. *Neurosurgical Focus, 38*(1), E3. https://doi.org/10.3171/2014.10.FOCUS14611

Picht, T., Krieg, S. M., Sollmann, N., Rösler, J., Niraula, B., Neuvonen, T., Savolainen, P., Lioumis, P., Mäkelä, J. P., Deletis, V., Meyer, B., Vajkoczy, P., & Ringel, F. (2013). A comparison of language mapping by preoperative navigated transcranial magnetic stimulation and direct cortical stimulation during awake surgery. *Neurosurgery, 72*(5), 808–819. https://doi.org/10.1227/NEU.0b013e3182889e01

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing https://www.R-project.org/.

Raffa, G., Marzano, G., Curcio, A., Espahbodinea, S., Germanò, A., & Angileri, F. F. (2022). Personalized surgery of brain tumors in language areas: The role of preoperative brain mapping in patients not eligible for awake surgery. *Neurosurgical Focus,* 53(6). https://doi.org/10.3171/2022.9.Focus22415

Raffa, G., Quattropani, M. C., & Germanò, A. (2019). When imaging meets neurophysiology: The value of navigated transcranial magnetic stimulation for preoperative neurophysiological mapping prior to brain tumor surgery. *Neurosurgical Focus, 47*(6), E10. https://doi.org/10.3171/2019.9.FOCUS19640

Rejno-Habte Selassie, G., Pegenius, G., Karlsson, T., Viggedal, G., Hallbook, T., & Elam, M. (2020). Cortical mapping of receptive language processing in children using navigated transcranial magnetic stimulation. *Epilepsy & Behavior, 103*(Pt A), Article 106836. https://doi.org/10.1016/j.yebeh.2019.106836

Rofes, A., & Miceli, G. (2014). Language mapping with verbs and sentences in awake surgery: A review. *Neuropsychology Review, 24*(2), 185–199. https://doi.org/10.1007/s11065-014-9258-5

Rofes, A., Spena, G., Miozzo, A., Fontanella, M. M., & Miceli, G. (2015). Advantages and disadvantages of intraoperative language tasks in awake surgery: a three-task approach for prefrontal tumors. *Journal of Neurosurgical Sciences, 59*, 337–349.

Roux, F. E., Miskin, K., Durand, J. B., Sacko, O., Rehault, E., Tanova, R., & Demonet, J. F. (2015). Electrostimulation mapping of comprehension of auditory and visual words. *Cortex, 71*, 398–408. https://doi.org/10.1016/j.cortex.2015.07.001

Schramm, S., Tanigawa, N., Tussis, L., Meyer, B., Sollmann, N., & Krieg, S. M. (2020). Capturing multiple interaction effects in L1 and L2 object-naming reaction times in healthy bilinguals: A mixed-effects multiple regression analysis. *BMC Neuroscience, 21*(1), 3. https://doi.org/10.1186/s12868-020-0549-x

Schwarzer, V., Bährend, I., Rosenstock, T., Dreyer, F. R., Vajkoczy, P., & Picht, T. (2018). Aphasia and cognitive impairment decrease the reliability of rnTMS language mapping. *Acta Neurochirurgica, 160*(2), 343–356. https://doi.org/10.1007/s00701-017-3397-4

Sollmann, N., Hauck, T., Hapfelmeier, A., Meyer, B., Ringel, F., & Krieg, S. M. (2013). Intra- and interobserver variability of language mapping by navigated transcranial magnetic brain stimulation. *BMC Neuroscience, 14*, 150. https://doi.org/10.1186/1471-2202-14-150

Sollmann, N., Ille, S., Boeckh-Behrens, T., Ringel, F., Meyer, B., & Krieg, S. M. (2016). Mapping of cortical language function by functional magnetic resonance imaging and repetitive navigated transcranial magnetic stimulation in 40 healthy subjects. *Acta Neurochirurgica, 158*(7), 1303–1316. https://doi.org/10.1007/s00701-016-2819-z

Sollmann, N., Ille, S., Negwer, C., Boeckh-Behrens, T., Ringel, F., Meyer, B., & Krieg, S. M. (2017). Cortical time course of object naming investigated by repetitive navigated transcranial magnetic stimulation. *Brain Imaging and Behavior, 11*(4), 1192–1206. https://doi.org/10.1007/s11682-016-9574-x

Sollmann, N., Kelm, A., Ille, S., Schroder, A., Zimmer, C., Ringel, F., Meyer, B., & Krieg, S. M. (2018). Setup presentation and clinical outcome analysis of treating highly language-eloquent gliomas via preoperative navigated transcranial magnetic stimulation and tractography. *Neurosurgical Focus, 44*(6), E2. https://doi.org/10.3171/2018.3.FOCUS1838

Sollmann, N., Zhang, H., Fratini, A., Wildschuetz, N., Ille, S., Schroder, A., Zimmer, C., Meyer, B., & Krieg, S. M. (2020a). Risk assessment by presurgical tractography using navigated TMS maps in patients with highly motor- or language-eloquent brain tumors. *Cancers, 12*(5), 1264. https://doi.org/10.3390/cancers12051264

Sollmann, N., Zhang, H., Schramm, S., Ille, S., Negwer, C., Kreiser, K., Meyer, B., & Krieg, S. M. (2020b). Function-specific tractography of language pathways based on nTMS mapping in patients with supratentorial lesions. *Clinical Neuroradiology, 30*(1), 123–135. https://doi.org/10.1007/s00062-018-0749-2

Szelényi, A., Bello, L., Duffau, H., Fava, E., Feigl, G. C., Galanda, M., Neuloh, G., Signorelli, F., Sala, F., & Workgroup for

Intraoperative Management in Low-Grade Glioma Surgery within the European Low-Grade Glioma Network. (2010). Intraoperative electrical stimulation in awake craniotomy: Methodological aspects of current practice. *Neurosurgical Focus, 28*, E7. https://doi.org/10.3171/2009.12.FOCUS09237

Talacchi, A., Santini, B., Casartelli, M., Monti, A., Capasso, R., & Miceli, G. (2013). Awake surgery between art and science. Part II: Language and cognitive mapping. *Functional Neurology, 28*(3), 223–239. https://doi.org/10.11138/FNeur/2013.28.3.223

Tarapore, P. E., Findlay, A. M., Honma, S. M., Mizuiri, D., Houde, J. F., Berger, M. S., & Nagarajan, S. S. (2013). Language mapping with navigated repetitive TMS: Proof of technique and validation. *NeuroImage, 82*, 260–272. https://doi.org/10.1016/j.neuroimage.2013.05.018

Tarapore, P. E., Picht, T., Bulubas, L., Shin, Y., Kulchytska, N., Meyer, B., Berger, M. S., Nagarajan, S. S., & Krieg, S. M. (2016). Safety and tolerability of navigated TMS for preoperative mapping in neurosurgical patients. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology, 127*(3), 1895–1900. https://doi.org/10.1016/j.clinph.2015.11.042

Tremblay, P., & Dick, A. S. (2016). Broca and Wernicke are dead, or moving past the classic model of language neurobiology. *Brain and Language, 162*, 60–71. https://doi.org/10.1016/j.bandl.2016.08.004

Tremblay, P., Dick, A. S., & Small, S. L. (2011). New insights into the neurobiology of language from functional brain imaging. In H. Duffau (Ed.), *Brain mapping: From neural basis of cognition to surgical applications* (pp. 131–143). Vienna, Austria: SpringerWienNewYork.

Turken, A. U., & Dronkers, N. F. (2011). The neural architecture of the language comprehension network: Converging evidence from lesion and connectivity analyses. *Frontiers in Systems Neuroscience, 5*, 1. https://doi.org/10.3389/fnsys.2011.00001

Zaccarella, E., & Friederici, A. D. (2017). The neurobiological nature of syntactic hierarchies. *Neuroscience and Biobehavioral Reviews, 81*(Pt B), 205–212. https://doi.org/10.1016/j.neubiorev.2016.07.038

Zulko. (2020). *MoviePy (python module for video editing)* (vol. accessed on 19.05.23) https://github.com/Zulko/moviepy.