Language assessment in awake brain surgery: The Portuguese adaptation of the Dutch Linguistic Intraoperative Protocol (DuLIP)

Joana Alves[1], Mafalda Cardoso[2], Mariana Morgado[3], Elke De Witte[4], Djaina Satoer[4], Andreia Hall[5] and Luis M. T. Jesus[5]

[1]Hospital das Forças Armadas – Pólo Porto, Porto, Portugal

[2]Unidade de Cuidados Continuados Solar Billadonnes, Penacova, Portugal

[3]Campus Neurológico Sénior, Torres Vedras, Portugal

[4]Department of Neurosurgery, Erasmus MC – University Medical Center Rotterdam, Rotterdam, The Netherlands

[5]University of Aveiro, Aveiro, Portugal

joanapalves@ua.pt; mafaldacardoso@ua.pt; mariana.morgado@ua.pt; elkedewitte@gmail.com; d.satoer@erasmusmc.nl; andreia.hall@ua.pt; lmtj@ua.pt

Awake brain surgery, combined with neurophysiological evaluation and intraoperative mapping, is one of the preferential lines of treatment when approaching low-grade gliomas. Speech and language assessment is used while applying Direct Electrical Stimulation (DES) and during the resection of a lesion/tumour, as it allows to establish related eloquent areas and optimise the extent of the resection and avoid impairments. Patients need to be assessed pre, intra and post-surgery, but in under resourced countries such as Portugal, there are still no standardised and validated tools to conduct this type of evaluation. To address this need, the tasks of the Dutch Linguistic Intraoperative Protocol (DuLIP) were adapted to European Portuguese, and the resulting materials were standardised for a group of 144 Portuguese participants. For each task, the impact of age, gender and schooling were measured. The resulting Portuguese version of the DuLIP (DuLIP-EP) consists of 17 tasks, including phonological, syntactic, semantic, naming and articulatory tests. No significant differences were found between male and female participants. However, schooling influenced phonological and syntactic fluency, object naming and verb generation. Schooling and age had a significant impact on semantic fluency and reading with semantic odd word out tasks. This is the first contribution to the standardisation of a tool that can be used during an awake brain surgery in Portugal, which includes a new *phonological odd word out* task that is not currently available in the Dutch version.

Keywords: language; linguistics; phonology; syntax; semantics; awake brain surgery; low-grade gliomas.

Introduction

The brain is organised in a distributed complex network underpinning sensorimotor, visuospatial, language, cognitive and emotional functions (Duffau, 2018). According to Friederici (2011), different brain regions, not only in the left hemisphere, but also in the right hemisphere have been identified to support language functions. Therefore, any lesion or growing process on these areas, like a brain tumour, can cause impairments in many different domains of language or even in motor speech.

Current guidelines to perform brain tumours removal, particularly Low-grade gliomas (LGGs), a type of brain tumour that grows slowly causing progressive lesions (Duffau, 2005), have established awake brain surgery as a standardised procedure to maximise resection and postoperative gains.

Direct Electrical Stimulation (DES) is considered the gold-standard surgical procedure during awake brain surgery. It consists of a biphasic electrical current that can mimic genuine temporary lesions and it allows the mapping of the brain, which in turn enables the identification of cortical and subcortical areas and pathways that are fundamental for speech and language. This procedure is applied by the neurosurgeon while linguistic tasks are performed, to assess the risk/benefit of different resection strategies before the actual procedure is performed. It has been shown to have a positive impact on the survival time and quality of life of the patients (Duffau, 2018), as it enables a greater remission of the tumour areas with fewer deficits associated to eloquent areas. It is, therefore referenced as the best method to maximise the "onco-functional balance" as it seeks to prevent the patient from losing a certain level of functionality (Bizzi, 2009; Coello et al., 2013; Darder & Lopez, 2012; Pereira et al., 2009; Rofes et al., 2017).

The assessment performed by a Speech and Language Therapist (SLT) in such a surgical context, has an important role in ensuring the preservation of the domains

associated with eloquent areas. This role includes implementing and interpreting the results of the pre-, intra-, and post-operative tasks, and the implementation of the rehabilitation process that follows (Geemen et al., 2014).

In most countries, language assessment in awake brain surgery is mainly conducted with the use of informal tools created by SLTs. Portugal is not an exception, so the language assessment of adults is mainly conducted with the use of informal tools created by the SLTs for personal use. However, the tests in use are quite limited in their evaluation of language skills during the intraoperative cortical mapping (Połczyńska, 2009).

In these surgeries, several technical-contextual specificities are considered and the evaluation must be performed accordingly, i.e., the spontaneous production of speech is assessed and pre-selected language tasks are performed by the patient, allowing the identification of the brain areas that might be affected during the resection (Bello et al., 2007).

In surgery, the patient can perform several sensorimotor, visuospatial, language, cognitive or emotional tasks simultaneously with the temporary disruption of brain structures using electrostimulation. Considering the previous (pre-surgical) assessment of the clinical case, the team should select the tests to be used during surgery. Specifically, to the language tests it is important to encompass the various domains of comprehension and expression of language, that should be designed for the conditions experienced in the operating room. Thus, the time for presentation and response to stimuli (the 4s of DES), the body positioning and the fact that the head is fixed in the same position for hours, the limited of space for equipment and the stress and discomfort of the patient are important aspects to consider. In addition, it is known that low-grade gliomas are characterised by slow growth, and this can lead to neuroplasticity processes. This fact, associated with the

fact that users have less marked deficits when compared with users with acquired aphasia after the occurrence of a stroke (Desmurget et al., 2007; Geemen et al., 2014), highlights the need to use a sensitive test that detects small variations.

To address this need for standardised tools, the Dutch Linguistic Intraoperative Protocol (DuLIP) was created (De Witte et al., 2015). Originally developed in the Netherlands, it includes specific and sensitive tasks from the areas of phonology, semantics, syntax and articulation. It was designed to be applied in the operating room and to detect small variations that might go unnoticed in a regular language assessment test. Since it includes tasks to collect data pre-, intra-, and post-surgery, it is quite an extensive protocol and therefore specific tasks have to be selected before the surgery based on the characteristics of the tumour (localisation and its functional implications) and the patients (De Witte et al., 2015).

Thus, this study aims to culturally adapt the DuLIP to European Portuguese (EP), standardise the DuLIP for the Portuguese population, support the role of the SLT in the intraoperative context and contributing to the improvement of patients' quality of life, as it is standardised according to the cultural habits and knowledge of the Portuguese population and allows for a more individualised intervention.

Method

The study comprised the linguistic-cultural adaptation of the DuLIP and a contribution towards the standardisation of this test battery in Portugal. To accomplish these goals, data were quantitatively and qualitatively analysed through descriptive statistics. All ethical procedures were followed. This study was approved by the Ethics Committee of the Research Unit of Health Sciences at the School of Nursing in Coimbra (UISCISA reference 535/01-2019, Coimbra, Portugal).

Adapting the DuLIP to European Portuguese

When translating and adapting language assessment tools, researchers face various challenges (Fyndanis et al., 2017): The cultural differences between languages; differences in the access to databases with information on linguistic attributes; differences specific to each domain of the language.

Therefore, much more than a literal translation was necessary to culturally and linguistically adapt the DuLIP according to the theoretical constructs that were at the core of this assessment tool. Nevertheless, following DuLIP's original authors' (De Witte et al., 2015) suggestions, the adaptation process began with the literal translation of the stimuli to Portuguese by three authors of this paper (Dutch and English versions were provided by the original authors). The materials resulting from this phase were discussed by all the authors of this paper and DuLIP's original authors, in expert meetings, to gather consensus for a pilot version. Since the psycholinguistic characteristics at the basis of creation of the original stimuli were not observed in some the Portuguese language items resulting from the initial literal translation, these materials were then individually adjusted and adapted according to those same properties.

Specific criteria such as frequency (data extracted from the Reference Corpus of Contemporary Portuguese (CLUL, 2013)), imageability, age of acquisition, prevalence,

word class, word complexity levels, number of phonemes, number of syllables and syllabic structure, time and verbal mode, syntactic constituents order, number of words and sentence type were contemplated. Additionally, there was an attempt to avoid stimuli with emotional connotation and to keep the original items (resulting from the translation into Portuguese), whenever possible.

As in the original DuLIP, black and white drawings were chosen with the same image size and simple line borders. The Snodgrass & Vanderwart (1980) image database was used, since Ventura (2003) had previously validated these images for the Portuguese population. The stimuli were presented in a white background Microsoft PowerPoint slide show, using a laptop computer.

The Portuguese version of the DuLIP (DuLIP-EP) encompasses one object naming task, seven semantic tasks (semantic odd word out, semantic picture out, semantic association, sentence completions - closed and broad context, semantic fluency, semantic sentence judgment), four phonological tasks (repetition words, phonological odd word out, phonological fluency and phonological sentence judgment), four syntax tasks (verb generation, syntactic fluency, syntactic sentence judgment I & II), and one articulatory task. Examples of some adapted tasks can be found in Figure 1.

[Figure 1 near here]

Regarding the phonological domain, the word repetition task contemplates six groups that vary by the presence or absence of consonant cluster and phonemic similarities, which results in a total of six complexity levels of the words. One phonological odd word out task was included in the EP version, despite its absence on the original study (De Witte et al., 2015). However, in order to include this task, it was decided that the words would only be presented as auditory stimuli. Concerning the phonological sentence judgment task, the pseudowords originally used did not conform

to Portuguese phonology, e.g., the syllabic structure and the position of the phonemes in the syllable, so they were substituted for others that did. For the phonological fluency task, phonemes /p, m, ɾ/ were chosen for the EP version (Cavaco et al., 2013).

The verbal diadochokinesis task was used to evaluate the articulatory domain, including the production of anterior sounds, which involve the orbicular musculature of the lips, the tongue tip and the back of the tongue, so there was no need to adapt this task to EP.

The use of synonyms in the naming task was analysed using the *hit rate*: If a unique synonym was systematically used rather than the original literal translation in more than 80 % of the collected data for a specific image (using the same cut-off / threshold for the exclusion of items), the word related to that image was changed for that synonym, in an attempt to have culturally more frequent and significant words for the general population. In DuLIP-EP, only object naming is considered, since verbs are assumed to be more complex to name and no EP validated action images are currently available.

Standardising the DuLIP-EP

Participants and Data Collection

This normative study involved 144 healthy adult volunteers. The participants had to be over 18 years old, with no upper age limit. The following additional inclusion criteria were defined: EP as the mother tongue; no record of cardiovascular, neurological, psychiatric or speech and language disorders; no record of substance abuse; normal vision and hearing; no use of sleep induction, psychopharmaceutical or neuroleptic drugs; a score over 24/30 on the Mini-Mental State Examination (MMSE). In the current study the Portuguese version of the Addenbrooke's Cognitive Examination III (Peixoto et al.,

2013) open access measurement tool was used instead of MMSE, but the results were converted to MMSE scores (Matías-Guiu et al., 2018) in order to facilitate comparisons with the original study (De Witte et al., 2015).

Participants were informed about the objectives of the study and gave written consent. The selection of the participants was based on their availability to take the test (convenience sample), and data were collected between April and August of 2019.

Procedures and Assessment Tools

Test administration, registration and correction of the items were discussed by the five authors of this paper. Each test had an average duration of 90 minutes and included the administration of the following assessment tools: Clinical history standardised form; Addenbrooke's Cognitive Examination III; DuLIP-EP. Data were collected by three speech and language therapists who followed specific and previously defined instructions. Example items were included in all tasks of DuLIP-EP to ensure that all participants understood the proposed tasks. The tasks were applied in the same order.

The first DuLIP-EP task to be presented was the *word repetition* task. In this task the person providing the stimuli read the word with good diction and projection while stressing the tonic syllable. Lip reading was avoided. Participants were asked to immediately repeat two and three-syllable words, e.g., <papel> / <*paper*> and <barulho> / <*noise*>, respectively. Phonologically or semantically related words were not sequentially presented to avoid perseveration. Afterwards, the *phonological odd word out* task was presented. This task is not currently available in the Dutch version of DuLIP because the original authors determined that it should be refined to better assess the function targeted. However, for the EP version, it was decided that the stimuli would only be presented auditorily and that this task would be kept in DuLIP-EP. Accordingly, the participants listened to a list of four words where one of the words did not rime with the others. The assessor read all the words with the same prosody, without emphasising the

intruder. This required some training before administering the task. The participant was then asked to identify which of the words did not rime with the others. An example of this task can be found in Figure 1 (A).

Then, *reading with semantic odd word out* task was presented. With this task it is possible to evaluate the semantic judgement via the lexical visual input route, as well as reading, verbal semantic knowledge and semantic processing. After that, the *naming with semantic odd picture out* task was presented. The purpose of the task is to evaluate non-verbal semantic judgment and naming. Just as in semantic odd word out task, semantic knowledge and processing are evaluated but, instead of written words, pictures are used. In both of these tasks, participants should see a sequence of three stimuli (picture or word) in a PowerPoint presentation and identify the one that was not correlated with the other two. Some examples of both tasks, reading and naming odd word out, are given in Figure 1 (F) and (E).

Then, the *semantic association* task was presented to analyse the lexical-semantic processing and also for assessing reading, semantic knowledge and producing words according to a specified concept. Participants should read two words and add a third semantically linked to the previous ones, e.g., <autocarro> / <*bus*> and <carro> / <*car*> can be related to <comboio> / <*train*>. The next task to be applied was the *semantic sentence completion I (closed context)*, in which the participant should complete a sentence in a closed context that was semantically correct. The aim was to evaluate speech production and reading, language dynamics and production of semantically and syntactically appropriate speech. It also allowed the evaluation of the auditory input route, semantic awareness, knowledge and comprehension, e.g., <A mulher… lê> / <*The woman… reads*>.

After that, the *verb generation* task from the syntactic domains was applied. In this task, 50 singular nouns were presented so that the participant could relate them to a verb, evoking it. The answers considered correct were those that were closely linked to the verb presented. Incorrect answers, that is not related to the verb, were excluded. For example, for the item shown in Figure I (D), <barco> / <*boat*>, one possible correct answer is <navegar> / <*to sail*>, and one incorrect answer could be <comer> / <*to eat*>.

A *verbal diadochokinesis* task was also used to evaluate the articulatory domain. This task aimed to evaluate articulatory performance, i.e., the planning, coordination and execution abilities of the motor speech system by measuring the ability to repeat a segment of speech at high rate. It is one of the recommended tasks to evaluate neurological disorders (Devadiga & Bhat, 2012): Alternating motion rate (AMR) (e.g., PAPAPA) and sequential motion rate (SMR) (e.g., PATAKA) are the two traditional tests of oral diadochokinesis used to assess motor speech production (Pierce et al., 2013).

Then the *naming* task was applied. Despite being related to the semantics' domain, naming skills were assessed and analysed separately through a task of object naming, given its clinical importance and transversally. In this task, 100 simple black and white line drawings were presented and the participant named them, within a time frame of 4s. Each image was accompanied by an inducing sentence <Isto é…> (<This is....>), as shown in Figure 1 (C).

The final tasks, had no time frame. These included the *phonological sentence judgment* task that was then presented (example in Figure 1 (B)), aiming to assess phonological awareness by asking the participants to point out which of the auditorily presented sentences was phonologically correct and incorrect (incorrect sentences contained pseudowords). This task also measures phonological decoding and verbal short-term memory (De Witte et al., 2015). *Phonological fluency* was assessed with a task

that involved the production (in one minute) of as many words beginning with /p/, /m/ or /ʀ/ (Cavaco et al., 2013) as participants could, after being instructed to avoid repetitions. Participants could say words belonging to any part of speech, e.g., proper nouns, common nouns, verbs and adjectives. For example, for the phoneme /p/, the answers could be <Pedro>, <pau> /<stick>, <presentear>/<to give> or <potente>/<potent>. This task assessed the ability to evoke phonemes and required several cognitive domains such as lexical and phonological memory, self-monitoring and cognitive flexibility (Baldo et al., 2006; Satoer et al., 2013).

Subsequently, the judgment of semantically anomalous and correct sentences (*semantic sentence judgment*) task was presented. The goal of this task was to discriminate between semantically correct and incorrect sentences (50 items). It allowed the evaluation of the auditory input route, semantic awareness, knowledge and comprehension and was used to assess temporal cortical and subcortical areas (Bello et al., 2007; Bertani et al., 2009; Pereira et al., 2009; Santini et al., 2012). For example, the item <O gato está a vestir roupa> (<The cat is putting on clothes>).

Then, the *semantic fluency* task was used to appraise word generation and semantic knowledge, through the production of the maximum number of animal and job names within a minute (for each category). The assessment of the semantic domain also included a *semantic sentence completion (broad context)* task, with similar goals to the closed context task, which induced more complete and open answers. For the item <A rapariga começou…> (<The girl began…), one possible answer is "<a estudar> (<to study>).

After that, the *Syntactic Judgment I & II* tasks were applied. These tasks aimed to assess the ability of discriminating between correct and incorrect sentences, in 25 items per task, considering different types of errors. The Syntactic Judgment I task included

incorrect word order in a sentence (e.g., <O chá bebe a mulher> / <*The tea drinks the woman*>), and the Syntactic Judgment II task was based on errors in nouns and verbs conjugations and lack of preposition (e.g., <Estou cansado de sentado> / <*I am tired of sat*>). All of the sentences in previous tasks (Syntactic Judgment I & II) varied in type and voice. These tasks involve frontal areas of the cortex. The application of DuLIP-EP ended with a *fluency syntactic task*, that aimed to induce the production of a maximum quantity of verbs possible within a minute.

The questions were answered orally and written down by the assessor. If the participants showed any signs of fatigue during the application of DuLIP-EP, the test was interrupted and postponed to a later date, if necessary. To score the tasks in the DuLIP-EP, a point was given for each correct answer that respects the time limit. One repetition was allowed. When the assessor had any doubts about how to score a test, the working group (three STLs) discussed until a consensus was reached. Self-corrections were allowed and considered correct as long as they respected the time limit.

In the DuLIP-EP, the tasks are split into two groups: Time-limited and not time-limited. The stimulus provided (DES) should not take more than 4s, preventing epileptic seizures (Duffau, 2016), so the language tasks should be performed within this time frame. In Table 1 tasks are grouped according to their time limit (4s) or no time limit.

[Table 1 near here]

Data analysis

Data obtained from the 144 participants were initially registered in a Microsoft Excel 2019 database and subsequently analysed with IBM SPSS Statistics 24. The stimuli adaptation hit rate (number of correct answers / number of total answers) was calculated per task and per item. Following the premise that items from widely used

neuropsychological batteries have been considered to discriminate brain-damaged patients from normal controls when hit rates (in normal controls) are higher than 90% (Robert et al., 2007), items in DuLIP-EP were only kept if the hit rate was above 90%. This was also the hit rate used in the original study (De Witte et al., 2015), with one exception: In the object naming task, items were changed if less than 80% of the participants correctly named the presented image.

Descriptive statistics were used to characterise the sample and summarise the data based on different measurements. The sample was demographically characterised according to data available for the different regions of Portugal (INE, 2015). The handedness of the participants was also registered. Moreover, two gender groups (male and female), two age groups (≤ 54 years of age and > 54 years of age) and two years of schooling groups (≤ 12 years of schooling and > 12 years of schooling) were considered for the analyses of the results.

Eurostat data (EUROSTAT, 2020), shows that the average life expectancy in Portugal and Netherlands is the same (82 years of age), so the same age threshold (54 years of age) as in the original study (De Witte et al., 2015) was used to form the two age groups, allowing for direct comparison between the Dutch and Portuguese studies. Since compulsory education in Portugal comprises 12 years of education, the 12 year threshold was used to form the two education groups, which is also the same as in the original study (De Witte et al., 2015).

The dependent variables included the number of correct answers per task. The average number of words produced per minute were calculated for the fluency tasks. The verbal diadochokinesis tasks were used to determine the average number of seconds taken to pronounce the monosyllabic words. Graphic and statistical methods (Shapiro-Wilks test) were applied to explore the normality of the data distribution. Considering that the

majority of the variables studied did not follow the normal distribution, the non-parametric Mann-Whitney U test was used to measure the effects of age, years of schooling and gender on ordinal test scores.

Since there was a large number of variables under study, in an atempt to reduce the statistical error of the tests, the p-value was corrected using the Sidak correction method. A p-value of 0.0010 was used as the criterion of significance (18×3 for age, gender and years of schooling).

Results

Sample characterisation

The average age of the sample was 36.81±14.86 [range 18 to 89] years old. Two age groups were established: 18 to 54 years old (n=121) and 55 years and older (n=23). Both male (n=62) and female (n=82) participants were included. The average years of schooling were 15.36±4.14 [range 4 to 24]. Participants who integrated the >12 years of schooling group (n=132) outnumbered those from the ≤12 years of schooling group (n=32). According to the MMSE scores, none of the participants were cognitively impaired (28.64±0.97 points). Most of the participants were from the Centre (n=94) and North (n=35) of Portugal. The remaining were from the Lisbon Metropolitan Area (n=13) and the Autonomous Region of Madeira (n=2). The demographic data of the participants are detailed in Table 2.

[Table 2 near here]

Adaptation of phonological tasks

During the adaptation of the word repetition task 222/300 (74%) of the items were changed, since the literal translation did not meet the linguistic criteria of the original DuLIP. These changes were either new insertions or transfers from one level to another. A total of 140 words (47%) were inserted and 82 (27%) were moved from their original group and integrated into another complexity group. The remaining 78 words (26%) resulted from literal translations and were not moved into another complexity level.

This initial translation and adaption of DuLIP was evaluated using the hit rate, calculated on a per item basis, from the results of the application of DuLIP-EP to the 144 study participants.

The hit rate for this task was 99.63%: Only the word <limiar> had a hit rate (78.47%) which was lower than the cut-off score (90%). After this item was eliminated, the hit rate was recalculated and a score of 99.70% was achieved for the word repetition task. The phonological odd word out task included 60 monosyllabic words distributed across 15 groups. The items inside each group shared the same syllabic structure. The hit rate initially calculated for this task was 96.13%. After removing item 8, which had a hit rate of 88.89%, a final score of 96.69% was reached. This item consists of the sequence '<chá-pó-lá-pá> ([ˈʃaˈpɔˈlaˈpa]), translation of <tea-dust-there-shovel>.

The hit rate for the phonological judgment task was 99.75%. None of the items scored less than 90%, so no sentences were deleted. In the verbal diadochokinesis task, 5 items exceeded the 4-second mark.

Adaptation of naming tasks

As explained previously, an 80% hit rate of correct naming item inclusion/retention criteria was considered exclusively for these tasks, which resulted in the exclusion of two images: <boneca> (<doll>) (67.36%), which was frequently named as <menina> (<girl>); <patim> (<roller skate>) (77.08%), confounded with various objects. The final naming task was therefore composed of 98 objects with a final hit rate of 97.40%.

Adaption of syntactic tasks

In the Syntactic Judgment I task, one item had to be excluded (item 13), since the hit rate of this item was only 85.42%, whereas in the Syntactic Judgment II task, no change had to be made. For the same reason, six words (nouns) had to be excluded from the Verb Generation Task. An example is <mente> (<mind>) that was frequently related to its homonym that is connected to a different verb (<to lie>). Hit rates, in this case, were between 80.56% and 89.59%. As a result, the final version of the Verb Generation Task

was composed of 44 items (the hit rate of this task was 96%), and the Syntactic Judgment I & II included 24 and 25 items, with a final hit rate per task of 98% and 99%, respectively.

Adaptation of semantic tasks

The semantic tasks that produced errors were: Reading with Semantic Odd Word Out; Naming with Semantic Odd Picture Out; Semantic Association; Semantic Judgment; Semantic Sentence Completion.

In the Reading with Semantic Odd Word Out task, five items were excluded since their hit rates were between 83.33% and 88.89%. Two items were eliminated from the Naming with Semantic Odd Picture Out task because their accuracy rate was lower than 90%. Eleven items were removed from the Semantic Association task, having revealed hit rates as low as 61.11%. Four sentences in the Semantic Judgment task had hit rates between 77.78% and 89.58% and were therefore suppressed. No items had to be eradicated from both Semantic Sentence Completion tasks.

The DuLIP-EP final version resulting from this adaptation stage, included 17 tasks, with a hit rate for each task and the items that had to be eliminated after the initial translation and adaptation, shown in Table 3.

[Table 3 near here]

Standardisation

Results were statically analysed, focusing on each variable (gender, age and years of schooling) and using Sidak's correction (p=0.0010) as in the original study (De Witte et al., 2015), in order to understand their practical impact in the performance of different tasks. It is important to consider that the sample of the present study (N = 144) is not

equally distributed among the 3 variables contemplated. A table with the p-values for each task and variable can be found in Appendix I.

No significant differences were found between male and female participants of this study, both in terms of average age ($p = 0.269$) and average years of schooling ($p = 0.293$). In general, participants obtained high scores, meaning that the average score is not far behind the maximum score.

As in the original study (De Witte et al., 2015), percentiles were calculated in order to define a cut-off score to distinguish patients that are considered to be *pathologically impaired* (percentile 2) and from those that are *clinically impaired* (percentile 7). These cut-off values were empirically defined and based on prior clinical practice, experience and methodologies (Palmer et al., 1998). Normative data resulting from this study is shown in tables 4 to 6, for four groups (Table 4) in tasks influenced by age and years of schooling, for two groups in tasks influenced by the years of schooling (Table 5), and for one group with all the participants (Table 6) in the tasks with no significant differences in terms of gender, age and years of schooling.

There was a significant difference in scores considering the younger and older less educated groups, and between older/less educated and younger/highly educated, for the *reading with semantic odd word out* task. The group with highest performance was the younger and highly educated, and the group with lowest performance was the oldest and less educated, as shown in Table 4. Mean value across the different groups was 19.02, where the minimum score obtained was 5 and the maximum 20, and the percentiles 2 and 7 were 11.80 and 16.00, respectively.

[Table 4 near here]

The *semantic fluency* task was also influenced both by years of schooling and age. Considering the four groups shown in Table 4, the mean score was 20.90 evoked words per minute, the minimum was 9.5 and the maximum 32.5.

Considering *verb generation* performance, only years of schooling influenced the scores and there was a clear high performance (shown in Table 5) among the population who had more than 12 years of schooling (highly educated), achieving scores near the highest score possible in this task (a total of 46 points / correct items). Participants with fewer years of schooling achieved scores around 40 correct items (less than 87% of the maximum score). The global mean score for this task was 44.29.

[Table 5 near here]

Years of schooling also had an influence on *syntactic fluency*, evidenced by the larger number of verbs evoked per minute among the highly educated population, as shown in Table 5. Only this group obtained scores between 35 and 40. The mean score across both groups was 22.4 words per minute, the minimum 5 and maximum 40.

Similarly, the *naming task* was only influenced by the years of schooling and as expected, high performance was found among the highly educated group as shown in Table 5. However, some of the less educated participants achieved the same amount of correctly named images, near the maximum score obtained (98). The global mean score for this task were 95.43.

Table 6 shows the normative data for all the other tasks with non-significant difference for gender, age and years of schooling.

[Table 6 near here]

Discussion

During this study, the tasks of DuLIP were translated and adapted for the Portuguese population and then applied to a group of normal participants in order to obtain normative data. This new version of DuLIP was based on the same constructs as the original one (De Witte et al., 2015). Therefore, it can also be used in pre, intra and post-operative surgeries with a patient with LGG, in awake brain surgery in Portugal.

During the adaptation process, an effort was made to keep the maximum possible number of original DuLIP items, using their literal translation, but still all of these items had to conform to specific psycholinguistic properties. Accordingly, some items had to be eliminated or replaced due to linguistic and cultural incompatibilities. All tasks and their items were discussed in an expert group that integrated the original authors of the DuLIP, and only after consensus was found did the items integrate the protocol.

The DuLIP-EP was administered to 144 participants and due to the minimum hit rate criteria of 90% in all tasks except in the object naming task with a hit rate of 80%, some items had to be removed and so the total number of items for each task decreased because participants did not respond within the 4 s time limit or answered incorrectly. Therefore, it was not be possible to make direct comparisons with all the results of the original study (De Witte et al., 2015), but some of the effects observed then, were also present in the current study: For example, there was not a significant difference between male and female participants. These gender differences have been shown to be negligible both during cognitive development (Ardila et al., 2011) and in adulthood (De Witte et al., 2015), and the current Portuguese results seem to corroborate this idea.

The *fluency* (*phonological*, *syntactic* and *semantic*), *reading with semantic odd word out*, *verb generation* and *object naming* tasks were influenced by years of schooling. Participants with more years of schooling showed better results, coinciding with the

original study (De Witte et al., 2015). This is also corroborated by several authors in different studies (Ganguli et al., 2010). Similar phonological fluency and semantic tasks conclusions can found in the literature (Cavaco et al., 2013), namely the effects of age and gender are less pronounced than the effects of years of schooling.

The tasks of *reading with semantic odd word out* and *semantic fluency* were influenced by age as well, so four groups were created intersecting the younger/older population with high/low educated. The group with < 54 years old participants and > 12 years of schooling obtained the best results, both in the semantic fluency task and in the reading with semantic odd word out task. In the original study (De Witte et al., 2015) the same four groups were created, and the results were similar: There was a significant difference when comparing the older/less educated population with the younger/highly educated sample. The older participants are, the lower the test score is and the lower the education level is the lower the scores are (De Witte et al., 2015; Papatzalas et al., 2020).

Cut-offs were calculated and analysed for the four existing groups. Since our sample is smaller than the original, the values of percentile 2 and 7 are not very accurate. Nevertheless, we have opted to present the table with the groups created and the respective values and percentiles for illustration purposes.

Participants found it more difficult to answer correctly or within the time limit during the *semantic association task*, with 11 items eliminated. In the original study, this was also the most difficult task possibly because in order to produce the correct answer it is necessary to ensure intact reading, semantic knowledge, language initiation, memory, attention and executive function (De Witte et al., 2015). A small number of items also had to be excluded from the *verb generation*, *object naming*, *syntactic judgment I*, *semantic judgment*, *repetition of words* and *phonological odd word out* tasks. Nevertheless, the mean scores obtained in each task are close to the maximum score

possible, which means that, globally, a good performance was achieved and few doubts, hesitation and ambiguities were perceived regarding the items included. The phonological odd word out task was included in DuLIP-EP, however the stimuli were adapted so as to be provided auditorily.

The use of synonyms on the *naming* task was frequent and mostly related to the participants' origin (most from northern Portugal). Yet, one item had to be changed considering the amount of times that was used instead of the original translation (>80%) (<dedo> / <finger> instead of <polegar> /<thumb>).

The need for this study arises from the lack of validated instruments that can be used in awake brain surgeries. As shown by Spena et al. (2017), in Portugal, awake brain surgeries are used on a small scale compared with the other 19 countries presented. Some data regarding the Portuguese practice in Garcia de Orta's Hospital in Almada is presented, such as the inexistence of a SLT during surgery. According to the same authors (Spena et al., 2017), the presence of a SLT in this surgical context is only considered in four cities: Madrid, Nice, Poitiers and Paris. It is known that there are other hospitals in Portugal where this kind of surgery is performed, however, only Anaesthesiology studies have been presented (Oliveira et al., 2011; Silva et al., 2014).

Conclusions

Better knowledge about the history of LGG, their behaviour and clinical outcomes, as well as the techniques available to map and resect this particular brain tumour, contributes to our understanding of how to control this disease aiming to improve quality of life in these patients. A validated assessment test battery is an essential tool to map the language functions in the brain, in order to preserve them in post-operatory time.

The main goal of this study was to linguistically and culturally adapt the DuLIP, as well as to contribute to its standardisation in Portugal. The test battery has a wide range of speech and language levels: Articulation, phonology, semantic and syntactic. Accordingly, stimuli were carefully created based on literal translations and in compliance with the psycholinguistic properties of the original instrument. This test battery was applied to a sample of 144 healthy Portuguese participants. Data were analysed according to age, gender and years of schooling of the participants. The results show that the tasks have a high hit rate and their mean scores are close to the maximum possible.

The DuLIP-EP is a test battery that can be used to assess language and articulation skills in pre-, intra- and postoperative moments in EP-speaking patients. This test battery has yet to be validated. To the author's best knowledge, no other language evaluation protocol for use during awake brain surgery has yet been adapted or validated for EP. Therefore, this is an innovative study in Portugal, where it will fill an existing gap in clinical practice and contribute to improving the quality of life of patients.

Limitations and future work

The internal consistency, test-retest reliability, criterion validity and construct validity of DuLIP-EP have yet to be studied, however, this was not the aim of this paper. In order to validate DuLIP-EP it is necessary to increase the number of items per task, since it was

not possible to compare the results obtained with the original study, due to these differences.

In the future it is important to increase the number of participants in order to be able to use more powerful statistical tests. Another limitation is that, the four groups created, are very heterogenous in terms of the number of participants, so it would be important to ensure balanced groups in future work.

Since previous studies of similar assessment instruments (designed for awake brain surgery) have been conducted with patients, the comparison with a normal population such as ours is quite difficult. It is vital to use the DuLIP-EP protocol with clinical cases, to test its application and comprehend the patient's performance.

As the main outcome of this study, there is now a Portuguese version of DuLIP which could be effortlessly adapted to clinical practice in that country. The translation and adaptation procedures, and the analysis task of performance developed during this study can be easily adopted by other teams that will adapt similar assessment instruments.

Acknowledgements

The authors would like to thank Evy Brink. This research was funded by the Foundation for Science and Technology (FCT), in the context of the projects UIDB/00127/2020 and and UIDB/04106/2020.

Table 1: Intraoperative linguistic tasks from DuLIP-EP

Timing of assessment(s)	Linguistic level	Task
During DES (in 4 s)	Phonology	- Word repetition - 3-syllable words without CC and PS - 3-syllable words with CC without PS - 2-syllable words without CC and PS - 2-syllable words with CC and without PS - 3-syllable words without CC and with PS - 3-syllable words with CC and PS
	Semantics	- Reading with semantic odd word out* and - Naming with semantic odd picture out* - Semantic association* - Sentence completion (closed context)*
	Syntax	- Verb generation*
	Articulation	- Verbal diadochokinesis*
	Naming	- Object naming*
Not during DES (without time limit)	Phonology	- Phonological sentence judgment - Phonological odd word out - Phonological fluency
	Semantics	- Semantic sentence judgment - Semantic fluency - Sentence completion (broad context)*
	Syntax	- Syntactic sentence judgment I & II - Syntactic fluency

*presented as a Microsoft PowerPoint 2019 slide show; CC = Consonant cluster; PS = Phonemic similarities

Table 2: Demographic characteristics of the participants

Demographics	Average	St. Dev.	Range
Age (y)	36.81	14.86	18-89
Schooling (y)	15.36	4.14	4-24
MMSE	28.64	0.97	25-30

Demographics	Groups	N. of participants	Percentage
Gender	M	62	43.10%
	F	82	56.90%
Age (y)	18-54 y	121	84.03%
	+55 y	23	15.97%
Schooling (y)	≤ 12	32	22.22%
	>12	112	77.78%
Region	North	35	24.31%
	Centre	94	65.28%
	Lisbon metropolitan area	13	9.03%
	Madeira islands	2	1.4%
Handedness	L	3	2.08%
	A	1	0.67%
	R	140	97.22%

MMSE = Mini-Mental State Examination; M = male; F = female; y = years; L = left-handed; A = ambidexter; R = right-handed

Table 3: Adaptation results

Linguistic level	Tasks	Hit rate (initial-final)	Items eliminated (%)
Phonology	Word repetition	99.63% - 99.70%	<Limiar> (78.47%)
	Phonological odd word out	96.13% - 96.69%	Item 8 (88.89%)
	Phonological judgment	99.75%	----------
	Phonological fluency	----------	----------
Articulatory	Verbal diadochokinesis	----------	----------
Semantics	Reading with semantic odd word out	93.55% - 95.31%	Item 5 (83.33%) Item 6 (88.89%) Item 12 (86.80%) Item17 (87.50%) Item 19 (86.11%)
	Naming with semantic picture out	96.87% - 98.02%	Item 20 (85.42%) Item 21 (81.94%)
	Semantic association	87.01% - 95.50%	Item 5 (86.81%) Item 10 (75.00%) Item 11 (81.94%) Item 12 (61.11%) Item 13 (62.50%) Item 16 (87.50%) Item 17 (84.72%) Item 18 (81.94%) Item 21 (68.96%) Item 22 (69.44%) Item 23 (79.17%)
	Sentence completion (Closed context)	98.08%	----------
	Sentence completion (Broad context)	99.67%	----------
	Semantic judgment	97.58% - 98.72%	Sentence 5 (85.42%) Sentence 14 (77.78%) Sentence 23 (89.58) Sentence 34 (85.42%)
	Semantic fluency	----------	----------
	Syntactic fluency	----------	

Table 3 (continued): Adaptation results

Linguistic level	Tasks	Hit rate (initial-final)	Items eliminated (%)
Naming		96.91% - 97.40%	<Boneca> (67.36%) <Patim> (77.08%)
Syntax	Syntactic judgment I	98.18% - 98.74%	1 sentence (84.72)
	Syntactic judgment II	98.08%	----------
	Verb generation	94.90% - 96.12%	<Sopa> (87.50%) <Hotel> (83.33%) <Fato de banho > (88.19%) <Idioma> (89.59%) <Mente> (80.56%) <Gato> (87.5%)
	Syntactic fluency	----------	

Table 4: Normative data for the tasks influenced by age and years of schooling

	Younger and highly educated		Younger and less educated		Older and less educated		Older and highly educated	
Reading with semantic odd word out	Mean	19.57	Mean	17.90	Mean	13.60	Mean	19.38
	Median	20.00	Median	18.50	Median	13.50	Median	19.00
	St. Dev.	0.95	St. Dev.	2.38	St. Dev.	4.19	St. Dev.	0.65
	Percentile 2	15.48	Percentile 2	12.00	Percentile 2	5.00	Percentile 2	18.00
	Percentile 7	18.00	Percentile 7	12.00	Percentile 7	5.00	Percentile 7	18.00
Semantic fluency	Mean	21.73	Mean	20.90	Mean	14.35	Mean	19.08
	Median	22.00	Median	19.75	Median	14	Median	16.50
	St. Dev.	4.41	St. Dev.	4.16	St. Dev.	2.77	St. Dev.	4.23
	Percentile 2	11.86	Percentile 2	16.00	Percentile 2	9.50	Percentile 2	11.50
	Percentile 7	15.50	Percentile 7	16.00	Percentile 7	9.50	Percentile 7	11.50

Table 5: Normative data for the tasks influenced by years of schooling

Task	Highly educated		Less educated	
Phonological fluency	Mean	18.22	Mean	14.20
	Median	17.67	Median	15.00
	St. Dev.	4.96	St. Dev.	3.89
	Percentile 2	10.00	Percentile 2	7.00
	Percentile 7	11.33	Percentile 7	7.63
Verb generation	Mean	44.89	Mean	40.60
	Median	46.00	Median	42.50
	St. Dev.	2.19	St. Dev.	5.51
	Percentile 2	35.50	Percentile 2	27.00
	Percentile 7	41.00	Percentile 7	28.41
Syntactic fluency	Mean	23.43	Mean	16.30
	Median	23.00	Median	15.50
	St. Dev.	7.03	St. Dev.	7.53
	Percentile 2	7.50	Percentile 2	6.00
	Percentile 7	13.75	Percentile 7	6.47
Naming	Mean	96.01	Mean	91.85
	Median	97.00	Median	93.00
	St. Dev.	2.09	St. Dev.	5.42
	Percentile 2	91.00	Percentile 2	81.00
	Percentile 7	92.75	Percentile 7	81.47

Table 6: Normative data for the tasks with no significant differences for gender, age and years of schooling

Phonological Intruder		Repetition of words		Phonological analysis		Semantics odd picture out		Semantic association	
Mean	13.50	Mean	298.13	Mean	29.94	Mean	22.63	Mean	13.40
Median	14.00	Median	299.00	Median	30.00	Median	23.00	Median	14.00
St. Dev.	1.31	St. Dev.	1.15	St. Dev.	0.26	St. Dev.	0.91	St. Dev.	1.08
Percentile 2	8.80	Percentile 2	295.00	Percentile 2	29.00	Percentile 2	18.90	Percentile 2	10.70
Percentile 7	12.00	Percentile 7	296.00	Percentile 7	30.00	Percentile 7	21.00	Percentile 7	12.00
Semantic completion (close)		Semantic completion (broad)		Semantic analysis		Syntactic judgment I		Syntactic judgment II	
Mean	24.51	Mean	24.92	Mean	45.42	Mean	23.72	Mean	24.51
Median	25.00	Median	25.00	Median	45.00	Median	24.00	Median	25.00
St. Dev.	0.95	St. Dev.	0.30	St. Dev.	0.99	St. Dev.	0.63	St. Dev.	0.81
Percentile 2	21.90	Percentile 2	24.00	Percentile 2	42.90	Percentile 2	21.90	Percentile 2	22.00
Percentile 7	23.00	Percentile 7	24.00	Percentile 7	44.00	Percentile 7	22.15	Percentile 7	23.00

<Mel> (<honey>) **<Fel>** (<gall>) **<Gel>** (<gel>) **<Mil>** (<a thousand>) **A**	**<O gaco corre para a Lirdo>** <A tempestade estragou as vindimas> **<O cromel faz um jagole>** <A roupa está pendurada na corda> **B**	Isto é… **C** 
Barco **D**	 **E**	**<Livro>** (<book>) **<Mar>** (<sea>) **<Revista>** (<magazine>) **<Jornal>** (<newspaper>) **F**

Figure 1. Examples of adapted tasks. (A) Phonological odd word out. The target is <mil> because it is the only word that does not rhyme with the others. (B) Phonological sentence judgment. The objective is to validate the sentences: The first and the third sentences are wrong and the second and fourth sentences as correct. (C) Naming. The target is <bicicleta> (<bike>). (D). Verb generation. One possible target would be <navegar>, because it's a verb related to the noun although more answers (synonyms such as <velejar>) are correct. (E) Naming with semantic odd picture out. The target is <telefone> (<telephone>) because it's not semantically related to the other two images. (F) Reading with semantic odd word out. The target is <mar> (<sea>) because it's not semantically related to the other three words.

References

Ardila, A., Rosselli, M., Matute, E., & Inozemtseva, O. (2011). Gender Differences in Cognitive Development. *Developmental Psychology*, *47*(4), 984–990. https://doi.org/10.1037/a0023819

Baldo, J., Schwartz, S., Wilkins, D., & Dronkers, N. F. (2006). Role of frontal versus temporal cortex in verbal fluency as revealed by voxel-based lesion symptom mapping. *Journal of the International Neuropsychological Society*, *12*(6), 896–900.

Bello, L., Gallucci, M., Marica, F., Carrabba, G., Giussani, C., Acerbi, F., Baratta, P., Songa, V., Conte, V., Branca, V., Stocchetti, N., Papagno, C., & Gaini, S. M. (2007). Intraoperative Subcortical Language Tract Mapping Guide Surgical Removal of Gliomas Involving Speech Areas. *Neurosurgery*, *60*(1), 67–82.

Bertani, G., Fava, E., Casaceli, G., Carrabba, G., Casarotti, A., Papagno, C., Castellano, A., Falini, A., Gaini, S. M., & Bello, L. (2009). Intraoperative mapping and monitoring of brain functions for the resection of low-grade gliomas: Technical considerations. *Neurosurgical Focus*, *27*(4), E4.

Bizzi, A. (2009). Presurgical Mapping of Verbal Language in Brain Tumors with Functional MR Imaging and MR Tractography. *Neuroimaging Clinics of North America*, *19*(4), 573–596. https://doi.org/10.1016/j.nic.2009.08.010

Cavaco, S., Gonçalves, A., Pinto, C., Almeida, E., Gomes, F., Moreira, I., Fernandes, J., & Pinto, A. T. (2013). Semantic Fluency and Phonemic Fluency : Regression-based Norms for the Portuguese Population. *Archives of Clinical Neuropsychology, 28*(3), 262–271. https://doi.org/10.1093/arclin/act001

CLUL. (2013). *Corpus de Referência do Português Contemporâneo (CRPC)*. Centro de Linguística da Universidade de Lisboa (CLUL). http://alfclul.clul.ul.pt/CQPweb/

Coello, A. F., 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.

Darder, J. M. G., & Lopez, P. G. (2012). Management of Brain Tumors in Eloquent Areas. In A. L. Abujamra (Ed.), *Diagnostic Techniques and Surgical Management of Brain Tumors*, 455–492. InTech. https://doi.org/10.5772/21796

De Witte, E., Satoer, D., Robert, E., Colle, H., Verheyen, S., Visch-Brink, E., & Mariën, P. (2015). The Dutch Linguistic Intraoperative Protocol: A valid linguistic approach to awake brain surgery. *Brain and Language*, *140*(1), 35–48.

Desmurget, M., Bonnetblanc, F., & Duffau, H. (2007). Contrasting acute and slow-growing lesions: A new door to brain plasticity. *Brain*, *130*(4), 898–914.

Devadiga, D. N., & Bhat, J. S. (2012). Oral Diadokokinetic Rate - An Insight into Speech Motor Control. *International Journal of Advanced Speech and Hearing Research*, *1*(1), 10–14. http://medical.cloud-journals.com/index.php/IJASHR/article/view/Med-04

Duffau, H. (2016). Stimulation Mapping of Myelinated Tracts in Awake Patients. *Brain Plasticity*, *2*(1), 99–113.

Duffau, H. (2018). The error of Broca: from the traditional localizationist concept to a connectomal anatomy of human brain. *Journal of Chemical Neuroanatomy*, *89*, 73–81.

EUROSTAT. (2020). *EUROSTAT: Your key to European statistics*. European Commission. https://ec.europa.eu/eurostat/

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

Fyndanis, V., Lind, M., Varlokosta, S., Kambanaros, M., Soroli, E., Ceder, K., Grohmann, K. K., Rofes, A., Simonsen, H. G., Bjekić, J., Gavarró, A., Kuvač Kraljević, J., Martínez-Ferreiro, S., Munarriz, A., Pourquie, M., Vuksanović, J., Zakariás, L., & Howard, D. (2017). Cross-linguistic adaptations of The Comprehensive Aphasia Test: Challenges and solutions. *Clinical Linguistics and Phonetics*, *31*(7–9), 697–710.

Ganguli, M., Snitz, B. E., Lee, C. W., Vanderbilt, J., Saxton, J. A., & Chang, C. C. H. (2010). Age and education effects and norms on a cognitive test battery from a population-based cohort: The Monongahela-Youghiogheny Healthy Aging Team. *Aging and Mental Health*, *14*(1), 100–107. https://doi.org/10.1080/13607860903071014

Geemen, K. Van, Herbet, G., Moritz-gasser, S., & Duffau, H. (2014). *Limited Plastic Potential of the Left Ventral Premotor Cortex in Speech Articulation : Evidence From Intraoperative Awake Mapping in Glioma Patients*. *1596*, 1587–1596.

INE (2015). As Novas Unidades Territoriais Para Fins Estatísticos. Instituto Nacional de Estatística (INE): Lisbon. ISBN: 978-989-25-0341-7

Matías-Guiu, J. A., Pytel, V., Cortés-Martínez, A., Valles-Salgado, M., Rognoni, T., Moreno-Ramos, T., & Matías-Guiu, J. (2018). Conversion between Addenbrooke's Cognitive Examination III and Mini-Mental State Examination. *International Psychogeriatrics*, *30*(8), 1227–1233. https://doi.org/10.1017/S104161021700268X

Oliveira, C., Gomes, C., Cerqueira, I., & Pereira, M. J. (2011). *1ª Anestesia para Craniotomia em doente acordado no Hospital de Braga*. Poster presented at the Congresso da Sociedade Portuguesa de Anestesiologia, 24-26 March 2011, Cascais, Portugal. http://hdl.handle.net/10400.26/1887

Palmer, B., Boone, K., Lesser, I., & Wohl, M. (1998). Base rates of "impaired"

neuropsychological test performance among healthy older adults. *Archives of Clinical Neuropsychology*, *13*(6), 503–511. https://doi.org/10.1016/S0887-6177(97)00037-1

Papatzalas, C., Fountas, K., Brotis, A., Kapsalaki, E., & Papathanasiou, I. (2020). The Greek linguistic assessment for awake brain surgery: development process and normative data. *Clinical Linguistics & Phonetics*. https://doi.org/10.1080/02699206.2020.1792997

Peixoto, B., Beata, E., & Pimentel, P. (2013). *Addenbrooke ' s Cognitive Examination* (pp. 1–6). CESPU-IUCS, Centro Hospitalar do Alto Minho, Centro Hospitalar de Trás-os-Montes e Alto Douro.

Pereira, L. C. M., Oliveira, K. M., L'Abbate, G. L., Sugai, R., Ferreira, J. A., & Da Motta, L. A. (2009). Outcome of fully awake craniotomy for lesions near the eloquent cortex: Analysis of a prospective surgical series of 79 supratentorial primary brain tumors with long follow-up. *Acta Neurochirurgica*, *151*(10), 1215–1230.

Pierce, J. E., Cotton, S., & Perry, A. (2013). Alternating and sequential motion rates in older adults. *International Journal of Language and Communication Disorders*, *48*(3), 257–264. https://doi.org/10.1111/1460-6984.12001

Połczyńska, M. (2009). New Tests for Language Mapping with Intraoperative Electrical Stimulation of the Brain to Preserve Language in Individuals with Tumors and Epilepsy: A Preliminary Follow-Up Study. *Poznań Studies in Contemporary Linguistics*, *45*(2), 261–279.

Robert, S., Saul, R., & Arnold, D. P. (2007). Neuropsychology for Psychologists, Health Care Professionals, and Attorneys (3rd ed.). CRC Press.

Rofes, A., Mandonnet, E., Godden, J., Baron, M. H., Colle, H., Darlix, A., de Aguiar,

V., Duffau, H., Herbet, G., Klein, M., Lubrano, V., Martino, J., Mathew, R., Miceli, G., Moritz-Gasser, S., Pallud, J., Papagno, C., Rech, F., Robert, E., … Wager, M. (2017). Survey on current cognitive practices within the European Low-Grade Glioma Network: towards a European assessment protocol. *Acta Neurochirurgica*, *159*(7), 1167–1178.

Santini, B., Talacchi, A., Squintani, G., Casagrande, F., Capasso, R., & Miceli, G. (2012). Cognitive outcome after awake surgery for tumors in language areas. *Journal of Neuro-Oncology*, *108 (2)*, 319–326.

Satoer, D., Kloet, A., Vincent, A., Dirven, C., Visch-brink, E., Satoer, D., Kloet, A., Vincent, A., Dirven, C., & Visch-brink, E. (2013). Dynamic aphasia following low-grade glioma surgery near the supplementary motor area : A selective spontaneous speech deficit. *Neurocase: The Neural Basis of Cognition*, *00*(00), 1–13. https://doi.org/10.1080/13554794.2013.841954

Silva, A. C., Mira, A., & Ferreira, C. (2014). Dexmedetomidina em craniotomia com doente acordado. *Revista Da Sociedade Portuguesa de Anestesiologia*, 23 (1), 14–16. https://doi.org/10.25751/rspa.3679

Snodgrass, J. D., & Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, *6 (2)*, 174-215. https://doi.org/10.1037/0278-7393.6.2.174

Spena, G., Schucht, P., Seidel, K., Rutten, G. J., Freyschlag, C. F., D'Agata, F., & Almairac, F. (2017). Brain tumors in eloquent areas: a European multicenter survey of intraoperative mapping techniques, intraoperative seizures occurrence, and antiepileptic drug prophylaxis. *Neurosurgical Review*, *40 (2)*, 287–298.

Ventura, P. (2003). Normas para figuras do corpus de Snodgrass e Vanderwart (1980).

Laboratório de Psicologia, *1 (1)*, 5–19. https://doi.org/10.1037/0278-7393.6.2.174

Linguistic Area	Task	Parameters	Gender	Age	Years of study
Phonology	repetition words	U	1845.50	1304.50	1468.50
		Z	-3.089	-0.522	-1.709
		p-value	0.002	0.602	.087
	phonological odd word out	U	2382.00	1209.50	1381.00
		Z	-0.845	-1.299	-2.585
		p-value	0.398	0.194	.010
	phonological fluency	U	1844.5	964.00	1025.00
		Z	-2.815	-2.332	-3.687
		p-value	0.005	0.02	.000*
	phonol. sentence judgment	U	2469.50	1329.00	1689.50
		Z	-0.785	-0.915	-1.322
		p-value	0.432	0.360	.186
Articulation	Verbal diadochokinesis	U	2196.00	1059.50	1767.50
		Z	-1.505	-1.811	-0.118
		p-value	0.132	0.070	.906
Semantics	Reading with semantic odd word out	U	2378.500	691.000	951.500
		Z	-.760	-4.400	-4.653
		p-value	.447	.000	.000*
	Naming with semantic odd picture out	U	2464.000	995.500	1308.500
		Z	-.444	-3.049	-3.281
		p-value	.657	.002	.001
	Semantic association	U	2042.500	1122.000	2042.500;
		Z	-2.363	-1.837	-2.363
		p-value	.018	.066	.018
	Semantic sentence completion (closed context)	U	2073.500	1122.000	1372.500
		Z	-2.363	-1.837	-2.520
		p-value	.018	.066	.012
	Semantic sentence completion (broad context)	U	2454.000	1298.000	1605.000
		Z	-.772	-1.108	-1.953
		p-value	.440	.268	.051
	Semantic fluency means	U	1916.500	628.500	849.500
		Z	-2.526	-4.164	-4.533
		p-value	.012	.000*	.000*
	Semantic sentence judgment	U	1909.000	1128.500	1411.500
		Z	-2.947	-1.655	-2.110
		p-value	.003	.098	.035
Syntax	syntactic judg. I	U	2519.00	1353.00	1656.00
		Z	-0.150	-0.300	-0.935
		p-value	0.881	0.746	.350
	syntactic judg. II	U	2265.00	1333.50	1381.50
		Z	-1.351	-0.382	-2.384
		p-value	0.177	0.702	.017
	verb generation	U	1903.00	1002.00	783.50
		Z	-2.769	-2.281	-5.205
		p-value	0.006	0.023	.000*
	syntactic fluency	U	1969.50	863.00	884.00
		Z	-2.312	-2.885	-4.368
		p-value	0.021	0.004	.000*
Naming	object naming	U	2267.50	865.00	932.00
		Z	-1.129	-2.926	-4.212
		p-value	0.259	0.00343	.000*

Appendix 1: Table with the p-values per category and task. * p <.001