Published in final edited form as:

Neurocrit Care. 2020 October ; 33(2): 449–457. doi:10.1007/s12028-019-00904-3.

EEG Correlates of Language Function in Traumatic Disorders of Consciousness

Camille Chatelle, PhD^{a,b,*}, Eric S. Rosenthal, MD^{b,c,d,*}, Yelena G. Bodien, PhD^{b,e}, Camille A. Spencer-Salmon, BSc^d, Joseph T. Giacino, PhD^{e,**}, Brian L. Edlow, MD^{b,f,**}

^aGIGA Consciousness, Coma Science Group, University of Liège, Liège, Belgium

^bCenter for Neurotechnology and Neurorecovery, Massachusetts General Hospital, Boston, MA

^cDepartment of Neurology, Massachusetts General Hospital, Harvard Medical School, Boston, MA

^dClinical Data Animation Center, Massachusetts General Hospital, Boston, MA

^eDepartment of Physical Medicine and Rehabilitation, Spaulding Rehabilitation Hospital, Harvard Medical School, Boston, MA

^fAthinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Harvard Medical School, Boston, MA

Abstract

Background/Objective: Behavioral examinations may fail to detect language function in patients with severe traumatic brain injury (TBI) due to confounds such as intubation. We investigated whether resting and stimulus-evoked EEG methods detect the presence of language function in patients with severe TBI.

Methods: Four EEG measures were assessed: 1) resting background (applying Forgacs' criteria), 2) reactivity to speech, 3) background and reactivity (applying Synek's criteria); and 4) an automated support vector machine (classifier for speech versus rest). Cohen's kappa measured agreement between the four EEG measures and evidence of language function on a behavioral

Corresponding Author: Dr. Camille Chatelle, GIGA-Consciousness – University of Liège, Avenue de l'Hôpital, 11, 4000 Liège, camillechatelle@gmail.com.

Author contributions:

C Chatelle: Substantial contributions to conception and design; Acquisition of data; Analysis and interpretation of data; Drafting the article; Final approval of the version to be published

ES Rosenthal: Substantial contributions to conception and design; Analysis and interpretation of data; Drafting the article; Final approval of the version to be published

YG Bodien: Acquisition of data; Critical Review of Manuscript; Final approval of the version to be published

CA Spencer-Salmon: Acquisition and analysis of data; Critical Review of Manuscript; Final approval of the version to be published

JT Giacino: Substantial contributions to conception and design; Critical Review of Manuscript; Final approval of the version to be published.

BL Edlow: Substantial contributions to conception and design; Acquisition of data; Analysis and interpretation of data; Drafting the article; Final approval of the version to be published.

* co-first authors

** co-senior authors

Publisher's Disclaimer: This manuscript has not been published or under consideration elsewhere. This manuscript complies with all instructions to authors. All authors have met requirements for authorship, as outlined below, and approved the final manuscript. Institutional Review Board approval was obtained prior to study initiation and informed consent was waived. STROBE reporting checklist was utilized as appropriate and included at the end of the manuscript.

(Coma Recovery Scale-Revised [CRS-R]) and composite (CRS-R or functional MRI) reference standard. Sensitivity and specificity of each EEG measure was calculated against the reference standards.

Results: We enrolled 17 adult patients with severe TBI (mean±SD age 27.0±7.0 years; median [range] 11.5 [2-1173] days post-injury) and 16 healthy subjects (age 28.5±7.8 years). The classifier, followed by Forgacs' criteria for resting background, demonstrated the highest agreement with the behavioral reference standard. Only Synek's criteria for background and reactivity showed significant agreement with the composite reference standard. The classifier and resting background showed balanced sensitivity and specificity for behavioral (sensitivity=84.6% and 80.8%; specificity=57.1% for both) and composite reference standards (sensitivity=79.3% and 75.9%, specificity=50% for both).

Conclusions: Methods applying an automated classifier, resting background, or resting background with reactivity may identify severe TBI patients with preserved language function. Automated classifier methods may enable unbiased and efficient assessment of larger populations or serial timepoints, while qualitative visual methods may be practical in community settings.

Keywords

Electroencephalography; consciousness disorder; language; traumatic brain injury; intensive care unit

Introduction

Detecting language function in patients with disorders of consciousness (DoC) is a critical challenge in the intensive care unit (ICU) and in long-term rehabilitation facilities. Behavioral evidence of language function, such as verbal expression and command-following, may influence decision-making about goals of care and affect access to rehabilitative care¹⁻⁵. However, behavioral assessments of language function are complicated by motor impairment, ventilation, sedation, fluctuating arousal, and ambiguous or rapidly habituating responses⁶. The misdiagnosis rate for detecting language function is unknown, although multiple studies have reported a 40% misdiagnosis rate for detecting signs of consciousness among patients clinically diagnosed as unconscious⁷⁻⁹. Objective, sensitive assessment tools are needed to detect signs of language function, a key component of consciousness.

In the ICU, the most commonly used behavioral assessment tool is the Glasgow Coma Scale (GCS)¹⁰, but the GCS provides limited sensitivity for detecting language, particularly in intubated patients. The Full Outline of Unresponsiveness (FOUR)¹¹ circumvents this limitation by integrating assessment of command-following using the visual modality. Currently, the behavioral standard for assessing language and consciousness in patients with DoC is the Coma Recovery Scale-Revised (CRS-R), which is more time-consuming to perform but provides enhanced detection of language function. However, serial CRS-R assessments may be required to detect language function and consciousness¹², and purposeful behaviors may still go undetected^{13,14}.

To address these limitations, EEG and fMRI methods have been developed over the past decade to identify patients with covert command-following¹⁵⁻¹⁷, a condition called cognitive motor dissociation¹⁸. These methods typically involve instructing the patient to execute a motor-imagery task, such as imagining opening and closing one's hand^{13,19-21}, or a motor command, such as squeezing one's hand²². The precise incidence of cognitive motor dissociation in the ICU and in long-term rehabilitation facilities is unknown, but emerging evidence suggests that 15- 20% of patients who are unresponsive or minimally responsive at the bedside are able to perform an EEG- or fMRI-based task^{8,13,14,20,22}. Concurrent with the development of task-based EEG and fMRI methods to detect *active* brain responses, stimulus-based methods have been developed to detect *passive* brain responses that do not require a volitional response by the patient but that suggest the possibility of preserved language function^{20,23-25}.

In order to translate advanced diagnostic techniques into clinical practice, it is necessary to evaluate the relative diagnostic utility, feasibility and generalizability of these techniques, each of which may complement the CRS-R behavioral assessment^{13,14,26-29}. From a data acquisition standpoint, EEG methods offer advantages over fMRI with respect to access, expense, and the opportunity to perform serial assessments to monitor recovery or evaluate the effectiveness of therapies. From a data processing and analysis standpoint, automated tools offer advantages because brain responses are analyzed without variability, bias, or the need for interpretation by clinical neurophysiologists²⁸. Yet recent evidence also suggests that qualitative visual assessment of a resting-state EEG may provide diagnostically relevant information about correlates of language function without the need for stimulus administration or automated analysis tools²⁷. Given that few studies have tested resting-state versus stimulus-based or automated versus qualitative EEG methods, it remains unknown which method provides the greatest diagnostic sensitivity for detecting clinically meaningful information about a patient's cognitive abilities.

In this prospective observational study, we tested resting-state and stimulus-based EEG methods, and both automated and qualitative analytic techniques, to determine which method is most reliable at identifying patients with traumatic DoC who retain language function. The presence of language function was based on two reference standards: a behavioral reference standard (i.e. command-following, intelligible verbalization, or intentional communication on the CRS-R), as well as a composite reference standard consisting of either behavioral evidence of language function on the CRS-R or fMRI evidence of covert command-following^{13,19,20,23,30}. The composite reference standard was chosen to overcome the limitations of the behavioral reference standard (e.g., cognitive and physical impairment) and account for covert language abilities that may be undetected on the CRS-R. Importantly, evidence of language function with neuroimaging has been demonstrated to correlate with EEG responses to language²³. We tested the agreement between four EEG measures (three qualitatively performed by a clinical neurophysiologist and one determined by an automated classifier) and the two reference standards. We also calculated the sensitivity and specificity of each method for identifying subjects with language function by the behavioral and composite reference standards. Finally, we explored the utility of combining EEG measures using logistic regression.

Methods

Participants

Patients with traumatic DoC and healthy subjects were prospectively screened and enrolled between 2012 and 2014 according to a protocol approved by the Institutional Review Board. Patients P1-P14 were enrolled consecutively in the ICU³¹, and patients P15-17 were enrolled subsequently in 2014-2015 to expand the cohort. Written informed consent was obtained from patients' surrogates and healthy subjects. Inclusion criteria for DoC patients were age 18 to 65 years, traumatic brain injury (TBI) with initial Glasgow Coma Scale score of 3-8, and no eye opening for at least 24 hours post-injury. Exclusion criteria were life expectancy less than six months, prior severe brain injury or neurodegenerative disease, and lack of English fluency prior to the injury (language paradigms were administered in English). Healthy subjects had no history of neurological, psychiatric, cardiovascular, pulmonary, renal or endocrinological disease.

EEG Index Tests

Critically ill patients were assessed in the ICU, and all other patients were assessed in their inpatient hospital rooms or in the EEG laboratory. EEG data were acquired as previously described²⁰, using a 19-electrode clinical EEG XLTEK system (Natus Medical Inc.; Pleasanton, CA) at a 200- or 256-Hz sampling rate and recorded at rest (10 minutes). A 10-20 montage of EEG electrodes was applied by EEG technologists. The language paradigm consisted of 12 alternating 24-second blocks of rest and speech and 5 minutes of continuous speech. Auditory stimuli were administered via a portable speaker system (Scosche Inc., Oxnard, CA) placed on a table next to the subject's bed. EEG was acquired in the absence of sedation in 76.5% of the patient cohort (see Table S2).

Three qualitative methods requiring visual analysis of EEG data were performed by a board-certified neurophysiologist (E.S.R), who was blinded to the behavioral and fMRI results. While the classifier results in the current study were previously reported, the Synek, Forgacs, and reactivity analyses are new.

1. *Forgacs' criteria* for resting EEG background: EEG features observed during ten minutes of rest were rated semi-quantitatively according to American Clinical Neurophysiology Society Critical Care EEG nomenclature and criteria³², as proposed by Forgacs and colleagues²⁷. Briefly, the resting EEG was classified into four categories based on the degree of abnormalities observed: "normal" (posterior dominant rhythm of 8–12Hz, amplitude difference not more than 50% between hemispheres, expected anteroposterior gradient [gradual increase in frequency and decrease in amplitude from posterior to frontal areas, with dominant beta activity over the frontal cortices], and no focal or hemispheric slowing); "mildly abnormal" (asymmetric or mildly slowed posterior dominant rhythm [not lower than 7 Hz], anteroposterior gradient not well organized and/or mild degree of focal or hemispheric slowing [slowing into the theta-range]); "moderately abnormal" (dominance of theta [4–7Hz] posterior rhythms and/or presence of moderate degree of focal or hemispheric slowing [slowing mostly in the theta-range with occasional delta-range slowing]) or "severely abnormal"

(delta [$< 4\text{Hz}$] waves over most of the brain areas). For statistical analyses, the score was dichotomized as “normal” or “mildly normal” (positive) versus “moderately abnormal” or “severely abnormal” (negative).

2. *EEG reactivity* was rated using data acquired during presentation of speech. We analyzed two pairs of 24-s blocks at rest (immediately prior to the speech stimulus) and during the speech stimulus. If reactivity was present in either pair of blocks, then the patient was graded as having reactivity. EEG reactivity was defined as change in amplitude or frequency in response to stimulation, based on American Clinical Neurophysiology Society Critical Care EEG nomenclature and criteria³².
3. *Synek's criteria for EEG background and reactivity* was used to rate EEG background activity and its reactivity³³, as implemented by Bagnato et al.³⁴: 1) dominant alpha activity with some scattered theta activity; 2) dominant reactive theta activity; 3) dominant widespread non-reactive delta activity or small amplitude, diffuse irregular delta activity; 4) burst suppression pattern, epileptiform discharges, non-reactive alpha or theta coma patterns; or 5) isoelectric activity. For statistical analyses, the score was dichotomized as positive (1 to 2) or negative (3 to 5).

A fourth method used an automated approach.

4. The automated classifier assessed EEG responses to spoken language during 12 alternating blocks of rest and speech. Following artifact rejection, the power spectral density of the voltage activity (i.e., absolute power) recorded at each electrode was estimated then averaged within four frequency bands [delta (1–3 Hz), theta (4–7 Hz), alpha (8–13 Hz), beta (14–30 Hz)], resulting in a matrix (76 features $\times$ 576 s) that was used for classifier analysis. A support vector machine with a linear kernel was used to classify the data with a 20-fold cross-validation method. This procedure was performed for each patient independently. An individual classification approach was used to account for the high inter-trial and inter-individual variability often observed in this population with heterogeneous brain lesions³⁵. We defined a positive response when $p < 0.05$ upon performance of a permutation test ($n=500$), as previously described²⁰.

Reference Standards

1. Behavioral reference standard: The CRS-R was performed immediately before each EEG to assess for behavioral evidence of language function, as indicated by signs of language perception, comprehension, and/or expression. Patients were classified as having “behavioral signs of language” if the CRS-R-derived behavioral diagnosis was minimally conscious state (MCS) plus (i.e. command following, intentional communication, or intelligible verbalization³⁶) or emerged from MCS and in post-traumatic confusional state³⁷. Alternatively, patients were classified as having “no behavioral signs of language” if their CRS-R-derived behavioral diagnosis was coma, vegetative state/unresponsive wakefulness syndrome, or MCS minus. The CRS-R was performed by a neurologist, who was

blinded to the fMRI and EEG data (B.L.E.). Evaluation of the CRS-R was performed in the absence of sedation. Healthy subjects demonstrated behavioral evidence of language function when providing informed consent.

2. Composite reference standard: The composite reference standard included data from both behavioral and neuroimaging assessments (i.e. CRS-R and task-based fMRI). The fMRI task involved motor imagery (right hand squeeze), which has been previously used in patients with acute and chronic DoC^{13,19-21}. This paradigm consisted of alternating 24-second blocks of rest and motor imagery, during which instructions were repeated at 6-s intervals (e.g. ‘keep squeezing’ or ‘keep resting’). Suprathreshold activity in a bilateral supplementary motor area and premotor cortex region was considered as evidence of command following, and hence language function, as previously described²⁰. fMRI was performed in the absence of sedation for 47.1% of patients with DoC (see Table A.1). Either behavioral evidence of language function on the CRS-R or fMRI evidence of covert command-following coded positive for command-following^{14,15,20,23,27}.

Statistical analyses

We calculated Cohen’s kappa to test agreement between each of the four EEG methods and the behavioral and composite reference standards. We calculated sensitivity (i.e. true positive rate; proportion of patients who retain language function correctly identified by the EEG method), and specificity (i.e. true negative rate; proportion of patients who do not retain language function and are correctly identified by the EEG method), for each reference standard in three populations: DoC patients, healthy subjects, and the combined cohort. Specificity could not be assessed for healthy subjects except when grouped with DoC patients because all healthy subjects had intact language function. We do not report an accuracy metric because it is highly influenced by sensitivity when the prevalence of a behavior is greater than 50%. Therefore, as noted in the discussion, “accuracy” is a misleading descriptor of the validity of each EEG method in this study³⁸.

We also performed a logistic regression to explore whether combinations of EEG measures identify patients with language function based on the behavioral or the composite reference standard. We optimized the model using stepwise binary logistic regression to investigate the additional information provided by combining EEG methods. We first performed univariate Fisher’s exact tests (binary variables). Variables with univariate $p \leq 0.2$ met criteria for inclusion in the multivariate model. Hosmer and Lemeshow goodness of fit tests were performed to assess the stability of the model. Results were considered significant at $p < 0.05$.

Results

Patient Demographics and Clinical Characteristics

We enrolled 16 healthy subjects (4F, age 28.5 ± 7.8 years) and 17 patients with traumatic DoC (4F, age 27.0 ± 7.0 years; 97.5 ± 282.4 days post injury, 13 right-handed [Table 1, Table S1]). Clinical and imaging data regarding each patient’s burden of structural brain injury are reported in Table S1. Briefly, on acute MRI, focal contusions were present in varying regions in 13 out of 17 patients. Eight patients had contusions in regions that may contribute

to language function (i.e., superior temporal gyri and inferior frontal gyri). At the time of EEG, 10 of 17 (58.8%) DoC patients demonstrated language function based on the behavioral reference standard, (MCS plus, n=5 and post-traumatic confusional state, n=5), whereas 7 patients demonstrated no behavioral signs of language (coma, n=2; unresponsive wakefulness syndrome, n=2, MCS minus, n=3).

We acquired fMRI data for all but one patient (in whom agitation prevented data acquisition) and all healthy subjects; however, fMRI data could not be analyzed in one patient due to artifact from a ventriculoperitoneal shunt. Seven out of 15 (46.7%) DoC patients and 11/16 (68.7%) healthy subjects demonstrated evidence of command-following on task-based fMRI. Thirteen of 17 (76.5%) DoC patients were positive for the composite reference standard, demonstrating either behavioral or fMRI evidence of language function. The mean percentage of EEG time-series data discarded due to artifact for the classifier analysis was 2.7+/-4.2% for patients and 1.2%+/-1.7% for healthy subjects.

Agreement Between EEG Measures and Reference Standards

Agreement with the behavioral reference standard was highest for the automated classifier ($k=0.40$; fair agreement; $p=0.02$), followed by Forgacs' criteria for resting background ($k=0.34$; fair agreement; $p=0.046$). Synek's criteria and EEG reactivity to language did not demonstrate significant agreement with the behavioral diagnosis. Agreement with the composite reference standard was significant only for Synek's criteria for background and reactivity ($k=0.37$; fair agreement $p=0.01$).

Sensitivity and Specificity of EEG Measures

Results for each EEG test detecting the behavioral reference standard are reported in Table 2. For the combined cohort of patients and healthy subjects, the classifier and Forgacs' criteria showed balanced sensitivity (84.6% and 80.8%, respectively) and specificity (57.1% for both). Sensitivity for identifying patients with language function was 100% for Synek's criteria for background and reactivity and 92.3% for reactivity itself, but specificity was only 14.3% for both measures due to their high false positive rates.

Results for each EEG test for detecting the composite reference standard are reported in Table 3. For the combined cohort of patients and healthy subjects, the classifier and Forgacs' criteria also showed balanced sensitivity (79.3% and 75.9%, respectively) and specificity (50% for both). Sensitivity was 100% for Synek's criteria and 93.1% for reactivity, but specificity was only 25% for both measures due to their high false positive rates.

Predicting language function using a multi-modal EEG Approach

Based on the univariate Fisher's exact tests for the combined patient and healthy subject cohort, the classifier ($p=0.04$) and Forgacs score ($p=0.07$) met criteria for inclusion in the regression model with the behavioral reference standard. The addition of the classifier led to a model prediction of behavioral evidence of language function of 81.8% ($R^2=0.23$; $p=.02$). In the regression for predicting the composite reference standard ($p=0.12$), Synek's criteria met the criteria for inclusion, but the results of Hosmer-Lemeshow test ($\chi^2(0)=0$, $p < 0.05$) indicated a poor logistic regression model fit in the setting of a high positive rate of

language function on the composite measure. Thus, model determination for the composite reference standard was rejected.

Discussion

We aimed to determine whether resting-state (i.e. Forgacs' criteria) or stimulus-based (i.e. classifier, Synek's criteria, or reactivity) EEG methods agree with behavioral (CRS-R) and composite (CRS-R and/or motor imagery fMRI) evidence of language function in patients with traumatic DoC and healthy subjects. The use of a composite reference standard was chosen to overcome the limitations of the behavioral reference standard and account for covert language abilities evident on fMRI that may be undetected by the CRS-R²³. We found that the automated classifier ($k=0.4$) and Forgacs score ($k=0.34$) showed the highest agreement with the behavioral reference standard for language function, while only Synek's criteria ($k=0.37$) showed agreement with the composite reference standard. The agreement was only fair for each measure, suggesting that these EEG methods should be considered as an adjunct to the behavioral examination and that there is a need for ongoing refinement of these methods prior to translation to clinical practice. Importantly, the classifier and Forgacs' criteria for resting-state EEG also provided the most balanced sensitivity and specificity for identifying patients and healthy subjects with language function. Our exploratory regression analysis did not demonstrate a benefit in combining EEG methods to identify patients with language function.

Collectively, our findings suggest that multiple EEG methods might be used as adjunctive tools to identify language function in patients with traumatic DoC. For patients who lack behavioral evidence of language function, EEG indications of language function could be the first signs of covert or emerging higher levels of consciousness. Identifying these signs early in recovery, when decisions regarding treatment intensity and life-sustaining therapies are made, is critical for accurate diagnosis and prognosis. Traditionally, detection of covert consciousness has been made based on performance of cognitively mediated tasks during fMRI or EEG. However, given that covert behaviors often go undetected even in healthy subjects and patients with behavioral evidence of language function (e.g. due to fluctuations in wakefulness or high cognitive task demands^{20,39}), it is necessary to develop methods of detecting cortical function that do not require volitional performance of goal-directed tasks. Although the stimulus-based and resting-state EEG methods in this study cannot definitively determine whether a patient is conscious, they may represent the precursors to emerging consciousness. Indeed, Braimann et al. recently showed that passive responses to spoken language on EEG (i.e. the natural speech envelope) correlate with active fMRI task performance²³.

This study provides preliminary guidance for clinical centers in the application of EEG methods for detection of cortical function in patients with DoC. Our findings suggest that both stimulus-based and resting-state EEG may be used as adjunctive tools for diagnosis. In addition, visual inspection and automated classifier approaches may both be appropriate. The quantitative automated classifier method may be suitable for centers that do not have access to a neurophysiologist, and it provides the advantage of limiting examiner bias and reducing workload when assessing numerous patients or timepoints. This automated method

also provides the potential to perform repeated measures to track recovery or assess responses to interventions²⁸. Conversely, the qualitative visual assessment methods (i.e. Forgacs's criteria) may be suitable for centers that have neurophysiologists but do not have the capacity to perform automated classifier analysis or when fewer data samples are being assessed. Ultimately, the implementation of such tools in centers that care for patients with traumatic DoC will depend upon careful consideration of each center's capabilities. As the methodology here involved review by a clinical neurophysiologist, visual inspection of EEG responsiveness performed in a community clinical setting may require a neurologist with expertise in neurophysiology.

Our findings should be interpreted in the context of several limitations. While the sample size here was comparable to prior studies^{13,19,27,30}, a larger sample might have enabled cross-validation. Of note, there was a well-balanced range of behavioral diagnoses across the spectrum of DoC in our patient sample, and the inclusion of a healthy cohort allowed for optimized assessment of sensitivity in a population with definitive evidence of intact language function. However, when the composite measure was used as the reference standard for language function, nearly all patients demonstrated language function. Therefore, the logistic regression model that combined EEG measures to identify patients with language function could not be fit. Our focus on patients with TBI may explain the high frequency of patients identified as having language function based on the composite reference standard, as previous studies have also shown that covert consciousness is more likely to be observed in patients with traumatic rather than non-traumatic brain injuries^{14,19,24,40}. Larger sample sizes with a greater variability in overt and covert language function (including different etiologies) will be needed to determine whether the classifier and Forgacs' criteria identify patients with language function when a composite behavioral and fMRI measure of language function is used. Of note, the aim of the study was to determine whether EEG responses could identify patients with behavioral evidence of language function, and not to use EEG to determine whether patients could understand language. Therefore, we chose to test resting and stimulus-based EEG measures that are either commonly used in clinical settings (e.g., Forgacs, reactivity, Synek) or that provided a standardized automating scoring approach (e.g., classifier). We did not use EEG to directly determine whether patients could understand language, which would have required a control condition to differentiate between responses to speech versus sound.

Additionally, clinical confounders such as fluctuations in arousal and sedation may have caused inconsistent responses between the behavioral, EEG and fMRI assessments⁴¹. Given that prior CRS-R studies indicate that repeated assessments are necessary to optimize detection of purposeful behaviors¹², it is possible that serial CRS-R may have improved behavioral detection of preserved language and altered the sensitivity/specificity results. Repeated EEG assessments may similarly be necessary to optimize detection of covert cognition²². To explore this issue in future studies, behavioral and EEG assessment should both be conducted serially in the ICU²². Whenever clinically feasible, we mitigated the impact of sedation by interrupting anesthetic infusions before the behavioral, EEG, and fMRI assessments. However, the potential contribution of sedation on the findings cannot be ruled out. Prior studies have shown that cortical function may be observed even in sedated patients^{22,42-44}, potentially due to individual differences in metabolism of sedating agents.

Although several EEG metrics had statistically significant agreement with the behavioral and composite reference standards, the overall strength of the agreement was fair for each of the significant measures. Thus, prior to clinical application, more methodological development and/or larger studies will be needed to improve the sensitivity, specificity, and agreement rate. Additional EEG methods such as the natural speech envelope²³ or N400⁴⁵ should also be investigated. Additional EEG methods such as the natural speech envelope or N400, which have been investigated in the chronic setting and have shown modest accuracy for identifying patients who demonstrate behavioral evidence of language function, should be evaluated in the ICU. It is currently unknown whether these methods will provide complementary, superior, or redundant information when combined with the techniques described here⁴⁵⁻⁵⁰.

Conclusion

Our data suggest that both automated and neurophysiologist assessment of stimulus-based and resting-state EEG methods can aid in detection of patients with traumatic DoC who retain language function. These EEG measures of cognitive function may provide quantitative, reproducible, and longitudinal proxies of language function in patients with traumatic DoC. Additional longitudinal studies are needed to determine whether early stimulus-evoked and resting-state EEG responses in the ICU predict long-term language recovery.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Study funding

The authors thank the nursing staffs of the Massachusetts General Hospital Neurosciences ICU, Multidisciplinary ICU, and the Surgical ICU. We also thank Joseph Cohen and the EEG technologists, as well as Kellie Cahill and the MRI technologists for their assistance with data acquisition. We are grateful to the patients and families in this study for their participation and support.

The study was funded by the NIH National Institute of Neurological Disorders and Stroke (K23NS094538, K23NS105950, R21NS109627, RF1NS115268), NIH Director's Office (DP2HD101400), American Academy of Neurology/American Brain Foundation, James S. McDonnell Foundation, Rappaport Foundation, Tiny Blue Dot Foundation, The Belgian American Educational Foundation, Wallonie Bruxelles International Excellence Scholarship Program and the European Commission (H2020-MSCA-IF-2016-ADOC-752686). The authors report no disclosures.

References

1. Luaute J, Maucourt-Boulch D, Tell L, et al. Long-term outcomes of chronic minimally conscious and vegetative states. *Neurology* 2010;75(3):246–52. [PubMed: 20554940]
2. Erler KS, Whiteneck GG, Juengst SB, et al. Predicting the Trajectory of Participation After Traumatic Brain Injury: A Longitudinal Analysis. *J Head Trauma Rehabil* 2018;33(4):257–65. [PubMed: 29385019]
3. Whyte J, Gosseries O, Chervoneva I, et al. Predictors of short-term outcome in brain-injured patients with disorders of consciousness. *Prog Brain Res* 2009;177(4):257–65.

4. Giacino JT, Kalmar K. The Vegetative and Minimally Conscious States: A Comparison of Clinical Features and Functional Outcome. *J Head Trauma Rehabil* [Internet] 1997;12(4). Available from: https://journals.lww.com/headtraumarehab/Fulltext/1997/08000/The_Vegetative_and_Minimally_Conscious_States__A.5.aspx
5. Bodien YG, Carlowicz CA, Chatelle C, Giacino JT. Sensitivity and specificity of the coma recovery scale-revised total score in detection of conscious awareness. *Arch Phys Med Rehabil* 2016;97(3).
6. Lotteries Gill-Thwaites H., loopholes and luck: misdiagnosis in the vegetative state patient. *Brain Inj* 2006;20(13–14):1321–8. [PubMed: 17378223]
7. Schnakers C, Vanhaudenhuyse A, Giacino J, et al. Diagnostic accuracy of the vegetative and minimally conscious state: clinical consensus versus standardized neurobehavioral assessment. *BMC Neurol* 2009;9:35. [PubMed: 19622138]
8. Stender J, Gosseries O, Bruno M-A, et al. Diagnostic precision of PET imaging and functional MRI in disorders of consciousness: A clinical validation study. *Lancet* 2014;384(9942).
9. Andrews K, Murphy L, Munday R, Littlewood C. Misdiagnosis of the vegetative state: retrospective study in a rehabilitation unit. *BMJ* 1996;313(7048):13–6. [PubMed: 8664760]
10. Teasdale G Jennett B. Assessment of coma and impaired consciousness. *Lancet Neurol* 1974;
11. Wijdicks EFM, Bamlet WR, Maramattom B V, Manno EM, McClelland RL. Validation of a new coma scale: The FOUR score. *Ann Neurol* 2005;58(4):585–93. [PubMed: 16178024]
12. Wannez S, Heine L, Thonnard M, Gosseries O, Laureys S. The repetition of behavioral assessments in diagnosis of disorders of consciousness. *Ann Neurol* 2017;81(6):883–9. [PubMed: 28543735]
13. Cruse D, Chennu S, Chatelle C, et al. Bedside detection of awareness in the vegetative state: A cohort study. *Lancet* 2011;378(9809).
14. Monti MM, Vanhaudenhuyse A, Coleman MR, et al. Willful modulation of brain activity in disorders of consciousness. *N Engl J Med* 2010;362(7):579–89. [PubMed: 20130250]
15. Owen AM, Coleman MR, Boly M, Davis MH, Laureys S, Pickard JD. Detecting awareness in the vegetative state. *Science* 2006;313(5792):1402. [PubMed: 16959998]
16. Giacino JT, Katz DI, Schiff ND, et al. Comprehensive systematic review update summary: Disorders of consciousness: Report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology; the American Congress of Rehabilitation Medicine; and the. *Neurology* 2018;91(10):450–60. [PubMed: 30089618]
17. Rohaut B, Eliseyev A, Claassen J. Uncovering Consciousness in Unresponsive ICU Patients: Technical, Medical and Ethical Considerations. *Crit Care* 2019;
18. Schiff ND. Cognitive Motor Dissociation Following Severe Brain Injuries. *JAMA Neurol* 2015;72(12):1413–5. [PubMed: 26502348]
19. Cruse D, Chennu S, Chatelle C, et al. Relationship between etiology and covert cognition in the minimally conscious state. *Neurology* 2012;78(11).
20. Edlow BL, Chatelle C, Spencer CA, et al. Early detection of consciousness in patients with acute severe traumatic brain injury. *Brain* 2017;140(9).
21. Bodien YG, Giacino JT, Edlow BL. Functional MRI Motor Imagery Tasks to Detect Command Following in Traumatic Disorders of Consciousness. *Front Neurol* 2017;8:688. [PubMed: 29326648]
22. Claassen J, Doyle K, Matory A, et al. Detection of brain activation in unresponsive patients with acute brain injury. *N Engl J Med* 2019;
23. Braiman C, Fridman EA, Conte MM, et al. Cortical Response to the Natural Speech Envelope Correlates with Neuroimaging Evidence of Cognition in Severe Brain Injury. *Curr Biol* 2018;28(23):3833–3839.e3. [PubMed: 30471997]
24. Coleman MR, Davis MH, Rodd JM, et al. Towards the routine use of brain imaging to aid the clinical diagnosis of disorders of consciousness. *Brain* 2009;132(Pt 9):2541–52. [PubMed: 19710182]
25. Gosseries O, Zasler ND, Laureys S. Recent advances in disorders of consciousness: focus on the diagnosis. *Brain Inj* 2014;28(9):1141–50. [PubMed: 25099018]

26. Stender J, Kupers R, Rodell A, et al. Quantitative rates of brain glucose metabolism distinguish minimally conscious from vegetative state patients. *J Cereb Blood Flow Metab* 2015;35(1).
27. Forgacs PB, Conte MM, Fridman EA, Voss HU, Victor JD, Schiff ND. Preservation of electroencephalographic organization in patients with impaired consciousness and imaging-based evidence of command-following. *Ann Neurol* 2014;76(6):869–79. [PubMed: 25270034]
28. Engemann DA, Raimondo F, King J-R, et al. Robust EEG-based cross-site and cross-protocol classification of states of consciousness. *Brain* 2018;141(11):3179–92. [PubMed: 30285102]
29. Stender J, Mortensen KN, Thibaut A, et al. The Minimal Energetic Requirement of Sustained Awareness after Brain Injury. *Curr Biol* 2016;26(11):1494–9. [PubMed: 27238279]
30. Curley WH, Forgacs PB, Voss HU, Conte MM, Schiff ND. Characterization of EEG signals revealing covert cognition in the injured brain. *Brain* 2018;141(5):1404–21. [PubMed: 29562312]
31. Edlow BL, Chatelle C, Spencer CA, et al. Early detection of consciousness in patients with acute severe traumatic brain injury. *Brain* 2017;140(1):23–31.
32. Guideline 7: Guidelines for writing EEG reports. *J Clin Neurophysiol* 2006;23(2):118–21. [PubMed: 16612228]
33. Synek VM. Prognostically important EEG coma patterns in diffuse anoxic and traumatic encephalopathies in adults. *J Clin Neurophysiol* 1988;5(2):161–74. [PubMed: 3074973]
34. Bagnato S, Boccagni C, Prestandrea C, Sant'Angelo A, Castiglione A, Galardi G. Prognostic value of standard EEG in traumatic and non-traumatic disorders of consciousness following coma. *Clin Neurophysiol* 2010;121(3):274–80. [PubMed: 20005157]
35. Noirhomme Q, Brecheisen R, Lesenfants D, Antonopoulos G, Laureys S. “Look at my classifier’s result”: Disentangling unresponsive from (minimally) conscious patients. *Neuroimage* 2017;145.
36. Schnakers C, Edlow BL, Chatelle C, Giacino JT. Minimally Conscious State. In: *The Neurology of Consciousness*. 2016.
37. Stuss DT, Binns MA, Carruth FG, et al. The acute period of recovery from traumatic brain injury: posttraumatic amnesia or posttraumatic confusional state? *J Neurosurg* 1999;90(4):635–43. [PubMed: 10193606]
38. Alberg AJ, Park JW, Hager BW, Brock MV, Diener-West M. The use of “overall accuracy” to evaluate the validity of screening or diagnostic tests. *J Gen Intern Med* 2004;
39. Cruse D, Chennu S, Chatelle C, et al. Bedside detection of awareness in the vegetative state: a cohort study. *Lancet* 2011;378(9809):2088–94. [PubMed: 22078855]
40. Schnakers C, Perrin F, Schabus M, et al. Voluntary brain processing in disorders of consciousness. *Neurology* 2008;71(20):1614–20. [PubMed: 19001251]
41. Piarulli A, Bergamasco M, Thibaut A, Cologan V, Gosseries O, Laureys S. EEG ultradian rhythmicity differences in disorders of consciousness during wakefulness. *J Neurol* 2016;263(9):1746–60. [PubMed: 27294259]
42. Kirsch M, Guldenmund P, Ali Bahri M, et al. Sedation of patients with disorders of consciousness during neuroimaging: Effects on resting state functional brain connectivity. In: *Anesthesia and Analgesia*. 2017.
43. Threlkeld ZD, Bodien YG, Rosenthal ES, et al. Functional networks reemerge during recovery of consciousness after acute severe traumatic brain injury. *Cortex* 2018;106:299–308. [PubMed: 29871771]
44. Bodien YG, Threlkeld ZD, Edlow BL. Default mode network dynamics in covert consciousness. *Cortex*. 2019;
45. Steppacher I, Eickhoff S, Jordanov T, Kaps M, Witzke W, Kissler J. N400 predicts recovery from disorders of consciousness. *Ann Neurol* 2013;73(5):594–602. [PubMed: 23443907]
46. Balconi M, Arangio R, Guarnerio C. Disorders of consciousness and N400 ERP measures in response to a semantic task. *J Neuropsychiatry Clin Neurosci* 2013;
47. Cruse D, Beukema S, Chennu S, Malins JG, Owen AM, McRae K. The reliability of the N400 in single subjects: Implications for patients with disorders of consciousness. *NeuroImage Clin* 2014;
48. Erlbeck H, Kübler A, Kotchoubey B, Veser S. Task instructions modulate the attentional mode affecting the auditory MMN and the semantic N400. *Front Hum Neurosci* 2014;

49. Rohaut B, Faugeras F, Chausson N, et al. Probing ERP correlates of verbal semantic processing in patients with impaired consciousness. *Neuropsychologia* 2015;66:279–92. [PubMed: 25447058]
50. Beukema S, Gonzalez-Lara LE, Finoia P, et al. A hierarchy of event-related potential markers of auditory processing in disorders of consciousness. *NeuroImage Clin* 2016;

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1.
Demographics, behavioral, fMRI and EEG findings.

F: female; Legend: M: male;; MCS+: minimally conscious state with language; MCS–: minimally conscious state without language; PTCS: post-traumatic confusional state; VS/UWS: vegetative state/unresponsive wakefulness syndrome; +: positive result; –: negative result. * significant at $p < .05$.

Patient ID	Age/Sex	Days Since Injury	Behavioral Diagnosis	Behavioral reference	Composite reference	Classifier	Forgacs	Reactivity	Synek
P1	27/M	17	PTCS	+	+	+	–	+	+
P2	21/M	2	MCS–	–	–	+	–	–	–
P3	19/F	4	Coma	–	–	–	+	+	+
P4	19/M	10	PTCS	+	+	+	–	+	+
P5	34/M	16	VS/UWS	–	+	–	+	+	+
P6	28/F	10	MCS+	+	+	+	–	+	+
P7	45/M	14	MCS+	+	+	–	+	+	+
P8	33/M	8	PTCS	+	+	+	+	+	+
P9	32/M	13	MCS+	–	–	+	+	+	+
P10	24/M	13	MCS–	–	+	–	–	+	+
P11	22/F	14	PTCS	+	+	–	–	+	+
P12	18/M	6	PTCS	+	+	+	–	+	+
P13	29/M	8	MCS+	+	+	+	+	–	+
P14	33/M	4	PTCS	+	+	+	+	+	+
P15	25/M	183	MCS+	+	+	+	+	+	+
P16	22/F	162	VS/UWS	–	+	+	–	+	+
P17	28/M	1173	MCS–	–	–	–	–	+	+
C1	27/M			+	+	+	+	+	+
C2	21/M			+	+	+	+	+	+
C3	27/M			+	+	+	+	+	+
C4	21/M			+	+	+	+	+	+
C5	19/F			+	+	–	+	+	+
C6	19/M			+	+	+	+	+	+
C7	34/M			+	+	+	+	+	+
C8	28/F			+	+	+	+	+	+
C9	45/M			+	+	+	+	+	+
C10	33/M			+	+	+	+	+	+
C11	32/M			+	+	–	+	+	+
C12	24/M			+	+	+	+	+	+
C13	22/F			+	+	+	+	+	+
C14	27/F			+	+	+	+	–	+
C15	18/M			+	+	+	+	+	+
C16	51/M			+	+	+	+	+	+

Patient ID	Age/Sex	Days Since Injury	Behavioral Diagnosis	Behavioral reference	Composite reference	Classifier	Forgacs	Reactivity	Synek
Agreement with Behavioral Reference Standard (Kappa)	–	–	–	–	–	0.40*	0.34*	0.08	0.21
Agreement with Composite Reference Standard (Kappa)	–	–	–	–	–	1.64	1.18	1.39	0.37*

Table 2.
Sensitivity and specificity based on the behavioral reference standard for patients, healthy subjects and for the combined cohort of patients and healthy subjects using each method.

DoC: Disorders of Consciousness; EEG: electroencephalography; Se: sensitivity; Sp: specificity

Method Measured against Behavioral Reference Standard	Patients with DoC (n=17)		Healthy Subjects (n=16)		Patients with DoC and Healthy Subjects (n=33)	
	Se	Sp	Se	Sp	Se	Sp
Automated EEG Classifier	80.0	57.1	87.5	—	84.6	57.1
Forgacs EEG Assessment	50.0	57.1	100	—	80.8	57.1
Synek EEG Assessment	100	14.3	100	—	100	14.3
EEG Reactivity	90.0	14.3	93.7	—	92.3	14.3

Table 3.
Sensitivity and specificity based on the composite reference standard for patients, healthy subjects and for the combined cohort of patients and healthy subjects using each method.

DoC, Disorders of Consciousness; EEG, electroencephalography; Se: sensitivity; Sp: specificity

Method Measured against Composite Reference Standard	Patients with DoC (n=17)		Healthy Subjects (n=16)		Patients with DoC and Healthy Subjects (n=33)	
	Se	Sp	Se	Sp	Se	Sp
Automated EEG Classifier	69.2	50.0	87.5	—	79.3	50.0
Forgacs EEG Assessment	46.1	50.0	100	—	75.9	50.0
Synek EEG Assessment	100	25.0	100	—	100	25.0
EEG Reactivity	92.3	25.0	93.7	—	93.1	25.0