NeuroImage 277 (2023) 120196

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Microsynt: Exploring the syntax of EEG microstates

Fiorenzo Artoni [a,b,*], Julien Maillard [d], Juliane Britz [c,b], Denis Brunet [a,b], Christopher Lysakowski [d], Martin R. Tramèr [d], Christoph M. Michel [a,b]

[a] Functional Brain Mapping Laboratory, Department of Basic Neurosciences, University of Geneva, Campus Biotech, Switzerland
[b] CIBM Center for Biomedical Imaging, Switzerland
[c] Department of Psychology, University of Fribourg, Fribourg, Switzerland
[d] Division of Anesthesiology, Department of Anesthesiology, Clinical Pharmacology, Intensive Care and Emergency Medicine, Geneva University Hospitals, and University of Geneva, Geneva, Switzerland

ARTICLE INFO

Keywords:
EEG
Microstates
Syntax
Complexity
Entropy
Anesthesia

ABSTRACT

Microstates represent electroencephalographic (EEG) activity as a sequence of switching, transient, metastable states. Growing evidence suggests the useful information on brain states is to be found in the higher-order temporal structure of these sequences. Instead of focusing on transition probabilities, here we propose "Microsynt", a method designed to highlight higher-order interactions that form a preliminary step towards understanding the syntax of microstate sequences of any length and complexity. Microsynt extracts an optimal vocabulary of "words" based on the length and complexity of the full sequence of microstates. Words are then sorted into classes of entropy and their representativeness within each class is statistically compared with surrogate and theoretical vocabularies. We applied the method on EEG data previously collected from healthy subjects undergoing propofol anesthesia, and compared their "fully awake" (BASE) and "fully unconscious" (DEEP) conditions. Results show that microstate sequences, even at rest, are not random but tend to behave in a more predictable way, favoring simpler sub-sequences, or "words". Contrary to high-entropy words, lowest-entropy binary microstate loops are prominent and favored on average 10 times more than what is theoretically expected. Progressing from BASE to DEEP, the representation of low-entropy words increases while that of high-entropy words decreases. During the awake state, sequences of microstates tend to be attracted towards "A – B – C" microstate hubs, and most prominently A – B binary loops. Conversely, with full unconsciousness, sequences of microstates are attracted towards "C – D – E" hubs, and most prominently C – E binary loops, confirming the putative relation of microstates A and B to externally-oriented cognitive processes and microstate C and E to internally-generated mental activity. Microsynt can form a syntactic signature of microstate sequences that can be used to reliably differentiate two or more conditions.

1. Introduction

An ever increasing number of works provide evidence that complex information processing involves large-scale distributed brain networks (Bressler and Menon, 2010; Mesulam, 2008). Even at rest, these networks seem to be inherently active in an organized manner replaying or preparing for information processing. Alterations in the spatial and temporal organization of large scale resting state networks (RSNs) have been related to changes of brain functions in neurological and psychiatric diseases (Cabral et al., 2014; Fox and Greicius, 2010).

Electroencephalography (EEG) is noninvasive, relatively inexpensive to set up, has a good temporal resolution and, in combination with microstates analysis, has gained popularity as a useful tool for investigating spatial and temporal properties of RSNs at rest or in association with different cognitive states and their alterations (Michel and Koenig, 2018). Microstates can be described as short time periods of stable scalp potential fields, generated by a network of synchronously active sources (Seeber and Michel, 2021) representing the electrophysiological correlate of the spontaneous chain of switching, transient, metastable states, even in the absence of input (Deco and Jirsa, 2012; Jirsa et al., 1998; Michel and Koenig, 2018; Tognoli and Kelso, 2014).

After determining the optimal number of microstates that best describe the data, spontaneous EEG is translated into a sequence of microstate topographies, alternating in discrete periods of approximately 60–100 ms in duration - see (Michel and Koenig, 2018) for a full review. Microstate sequences are conventionally described using a series of parameters, such as: "Duration" (i.e., the average time that a given microstate remains stable), "Coverage" (i.e., the fraction of total record-

* Corresponding author.
E-mail address: fiorenzo.artoni@unige.ch (F. Artoni).

https://doi.org/10.1016/j.neuroimage.2023.120196.
Received 22 January 2023; Received in revised form 16 May 2023; Accepted 25 May 2023
Available online 5 June 2023.

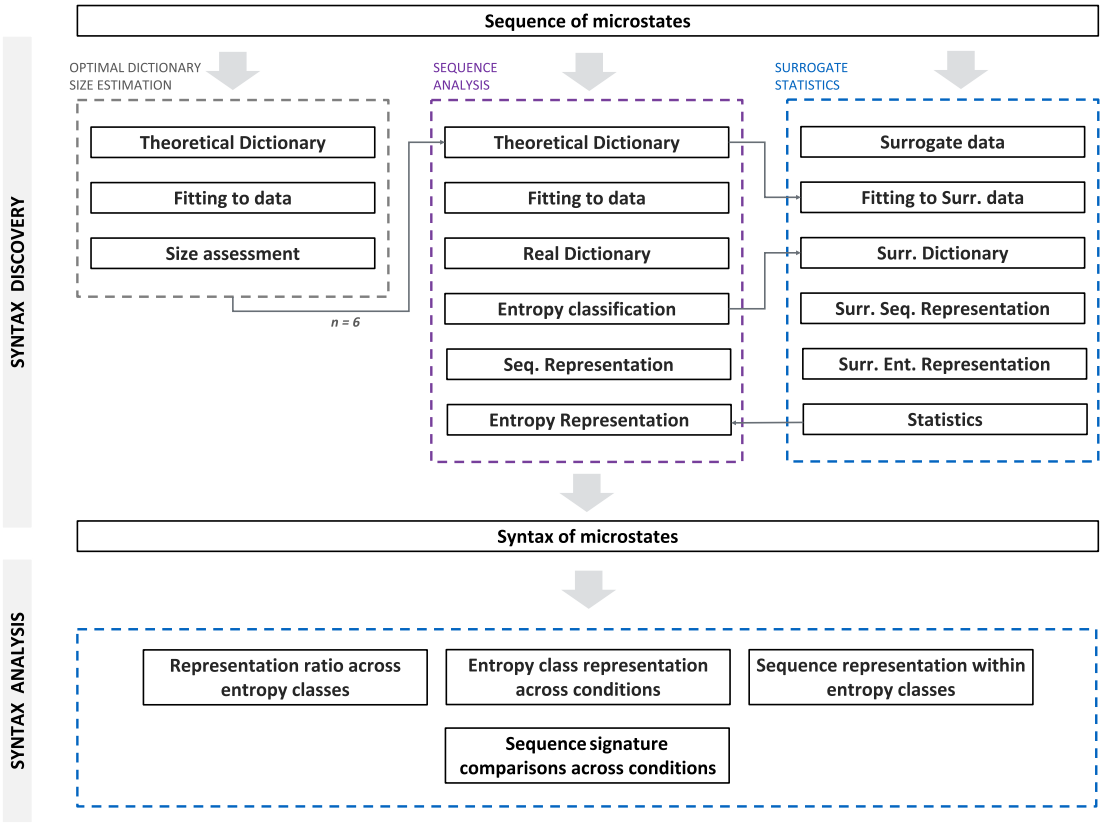

Fig. 1. Analysis pipeline
Schematics representing the steps performed for the analysis of sequence of microstates to extract their syntax (see "Methods"). The "Syntax Discovery" pipeline focuses first on estimating the optimal dictionary and sequence size for the analysis ("Optimal Dictionary Size Estimation"). The "Sequence Analysis" steps (violet dotted line) include the extraction of all elements of size 6, their pooling according to entropy and the quantification of their representation within the data. The "Surrogate Statistics" steps (blue dotted line) focus on creating the necessary statistics, based on bootstrap, to determine which sequences are significantly more or less represented in the real data than in theoretical/synthetic data. The "Syntax Analysis" pipeline includes the analysis of (i) the representation ratio across entropy classes and against surrogate data, (ii) the entropy class representation across conditions (Fig. 3), (iii) the sequence representation within entropy classes (Fig. 4), (iv) the syntax signature comparisons across conditions (Fig. 5). These analyses also allow to establish which microstate sequence "hubs" best differentiate different states of consciousness. .

ing time for which a specific microstate is dominant), "Occurrence" (i.e., the number of occurrences per unit of time for each microstate, independent of its duration), "Global Explained Variance" (i.e., the variance explained by each microstate) "Correlation" (i.e., the average correlation of each microstate topography with the raw data topographies relative to the time said microstate is dominant). These measures however fail to effectively highlight the rules that govern their temporal structure.

Transition probabilities (TP), i.e., the conditioned probability of a microstate given the appearance of another, obtained using hidden Markov models (Lehmann et al., 2005), constitute a useful step towards a deeper analysis of the temporal structure of microstate sequences and are gaining popularity as a tool to characterize brain activity under different conditions (Gschwind et al., 2015). For example, (Lehmann et al., 2005; Wackermann et al., 1993) showed that TP are not random except perhaps in patients with Alzheimer-type dementia (Nishida et al., 2013), and are altered, for example, by schizophrenia (Lehmann et al., 2005; Tomescu et al., 2015). There is general consensus, however, that further modeling of microstate sequences should extend beyond the first and second-order descriptions of short-term interactions highlighted by TP.

For this reason, there has been a growing interest in analyzing what could be called "microstates syntax". In literature, the term "syntax" usually refers to the study of frequencies of transition pairs (Lehmann et al., 2005; Nishida et al., 2013; Schlegel et al., 2012), conditional transition probabilities (Antonova et al., 2022; Bréchet et al., 2019; Lehmann et al., 2005; Tomescu et al., 2018, 2015), or properties of the transition matrix (von Wegner et al., 2017). However, if microstates metaphorically represent "atom of thoughts", (or "phonemes" in a hypothetical "language of the brain"), then studying microstates syntax amounts to studying whether a set of rules (e.g., grammatical relations, hierarchical sentence structure etc.) can be found to govern how these phonemes form larger units such as words, that can form phrases and sentences, with a meaning greater than the sum of the meaning of each word, considered in isolation. Indeed, the study of transition probabilities should be generalized to higher order loops (e.g., A->*B*->*A*->C->B->…). On the other hand, relaxing the model order constraints can rapidly render the problem computationally intractable either due to computing power limitations or the finite nature of microstate sequences.

To this aim, here we present a novel method, "Microsynt", to highlight higher-order interactions that form a preliminary step towards understanding the syntax of microstates sequences. Microsynt allows to infer the optimal size of the vocabulary of "words" based on the length and complexity of the full sequence of microstates. Based on this estimation, Microsynt extracts the full vocabulary of the interlinked words forming the original sequence. To make the problem tractable in real-data scenarios, words are sorted into classes of entropy. Subsequently, the representativeness of each class of entropy is statistically compared with surrogate and theoretical vocabularies. Microsynt is therefore applicable to microstate sequences of any length and complexity, it de-

termines the optimal word size and statistically significant sequences, using a Pareto-like approach (Hochman and Rodgers, 1969), both in relation to surrogate data or when comparing different conditions. Here we applied Microsynt on EEG data previously collected from healthy adults undergoing propofol anesthesia for surgery, and compared their fully awake and fully unconscious states before surgery commenced.

2. Materials and methods

A Experimental setup and data collection

A detailed description of experimental setup, data collection, data preprocessing and extraction of microstates can be found in (Artoni et al., 2022) and is briefly reported here. Fig. 1 also shows a diagram of the processing steps. The analyses were performed on EEG data collected from twenty-three not premedicated patients (mean age 30 years, range 20–47 years) scheduled for minor elective surgery (ear-nose-throat/plastic surgery). All patients gave informed consent and the study protocol was approved by the Ethics Committee of Geneva University Hospitals (CER 12–280). 64-channel EEG data with active Ag/AgCl electrodes (actiCap; BrainProducts) were recorded in an extended 10–10 system under the control of neuroscientists (Oostenveld and Praamstra, 2001). Prior to the arrival in the operation room, subjects were instructed to stay with closed eyes and to relax as much as possible. After a resting period of 10 min, a baseline EEG (BASE) was recorded (5 min duration). Subsequently, the anesthetic propofol was administered intravenously using a Target Controlled Infusion device (Base Primea, Fresenius-Vial, Brezins, France) which also served to estimate the propofol concentrations in the plasma and at the effect-site (brain). Effect-site concentrations were increased in steps of 0.5 μg ml^{-1} until loss of consciousness. After each increase, the "steady-state" was maintained for 5 min, and after this period, a five-minute EEG recording was done with a band pass filter between DC and 1000 Hz and was digitized at 5 kHz, with an online reference at FCz. The degree of alertness between fully alert condition (BASE) to general anesthesia (DEEP) was measured using a modified Observer's Assessment of Alertness/Sedation (OAA/S) scale (ranging from 5 [fully awake] to 1 [deep sedation]) (Chernik et al., 1990). No other drugs were administered during the recording period.

A Preprocessing and extraction of microstates

Data were preprocessed with custom MATLAB scripts based on routines from the EEGLAB toolbox (Delorme and Makeig, 2004) and within Cartool (Brunet et al., 2011) following an increased-stability procedure tested in previous works (Artoni et al., 2017) and described in (Artoni et al., 2022). First, raw data were preliminarily bandpass filtered (1–45 Hz) and processed using a Reliable Independent Component Analysis (RELICA) approach with an AMICA core (Artoni et al., 2014) to label artifact independent components (ICs), channels and portion of artifact data. Then raw data were more conservatively filtered (0.2 – 24 Hz) and previously identified ICs, channels and portions of artifact data removed. Missing channels were then interpolated, clean data were downsampled to 250 Hz and spatially filtered within Cartool to improve the Signal to Noise Ratio (SNR) of the data (Michel and Brunet, 2019) before undergoing the microstates extraction procedure.

EEG microstate segmentation was performed using the standard procedure also described in (Michel and Koenig, 2018) and shown in Fig. 2. For each condition and participant, the Global Field Power (GFP) peak maps (channel values at the timestamp corresponding to the GFP peak) were extracted to ensure a high signal-to-noise ratio (Koenig et al., 2002) and were clustered via modified k-means to extract distinct templates (Murray et al., 2008; Pascual-Marqui et al., 1995). Within this step, the spatial correlation between each GFP map and each randomly generated template was calculated while ignoring the polarity of maps (Michel and Koenig, 2018). Each template was iteratively updated by averaging the GFP maps that presented the highest correlation with the template. At the same time, the Global Explained Variance (GEV) of template maps

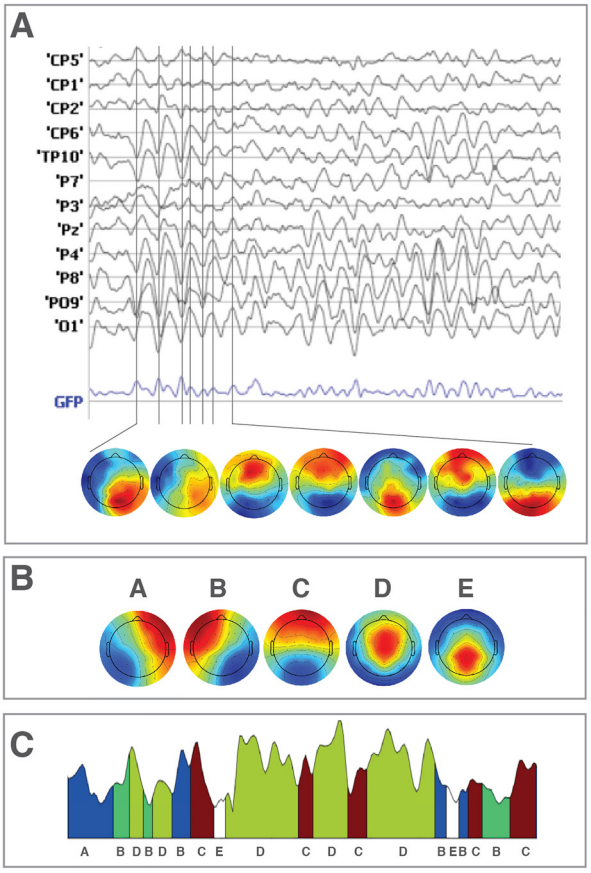

Fig. 2. Schematic representation of the microstates extraction procedure (Panel A) sample 2s scalp channel data (gray horizontal lines) with relative Global Field Power (GFP – blue horizontal line). Example representation of 7 GFP peak maps (each showing the head-surface potential maps at one peak of the GFP curve). (Panel B) Template dominant topographies (EEG microstates) obtained after submitting the GFP peak maps to the clustering algorithm (see Methods). (Panel C) Results of the assignment of each original maps at the peaks of the GFP curve to one of the microstate classes, A, B, C, D or E based on the degree of the spatial similarity with the microstate maps (fitting procedure).

was calculated, and the process was iterated until the stability of GEV was reached. For both conditions of alertness (BASE and DEEP), the optimal number of microstate classes was determined using an optimum MetaCriterion implemented and published with the Cartool toolbox and discussed in depth in (Bréchet et al., 2019). Given the high correlation between paired maps across conditions and the similar assessment of the optimal number of microstates yielded by the meta-criterion, the data of all conditions were pooled and the extraction process repeated. Finally, spatial correlation between the templates identified at the group level (ALL) and each data point of the EEG of each subject was computed (Fig. 3). EEG frames were labeled using a "winner-takes-all" strategy (Michel and Koenig, 2018) according to the group template it best corresponded to (a temporal constraint - Segments Temporal Smoothing - of 6 samples/24 ms was applied and no labeling was performed at correlations lower than 0.5), which generated the microstate sequence for further analysis.

A Sequence preprocessing

Microstate sequences were extracted for both conditions (BASE, DEEP). As briefly outlined also in (Artoni et al., 2022), section K- "Materials and methods", each sequence can be modeled as concatenation of microstate template maps (A,B,…, E), each appearing a number *m* of times. For example, a sequence such as [AAACCCD] would be constituted by the microstate maps "A", "C", "D" of durations "3″, "3″,

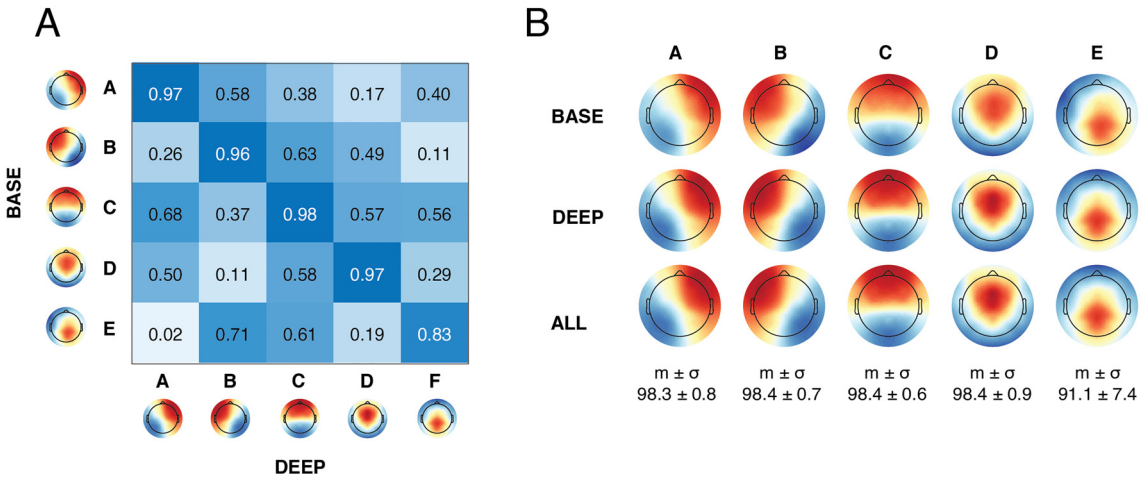

Fig. 3. Microstates
(Panel A): Similarity matrix (absolute correlation coefficient), representing the absolute correlation of microstate maps between BASE and DEEP conditions (Panel B): Topographic representation of the 5 microstates obtained, respectively, for conditions "BASE", "DEEP" and both ("ALL"). The average correlation of ordered microstates maps and the standard deviation are reported at the bottom of each column. Figure adapted (with permission) from (Artoni et al., 2022) .

and "1″ respectively. After removing unlabeled points of the sequence (i.e., [ABB-C-DDD] -> [ABBCDD]), each microstate lasting less than 5 successive time points (<20 ms) was removed. This preprocessing step was devised to remove from the sequence unstable microstates, i.e., active for less than or equal to 20 ms. The resulting sequence was then transformed into a "no-permanence" sequence by setting to 1 the size of each microstate, thus removing the repetition of microstates. For example, a sequence such as [AAAAACCCCCCBBBBB] would be transformed into [ACB]. This transformation ignores the different duration of the microstates, i.e. it only considers the temporal sequence of the states, independent of their individual length.

A The randomness of the sequence

The resulting no-permanence sequence for each subject and condition underwent a Lempel Ziv complexity (MS-LZC) estimation to determine its randomness (and therefore the potential information content), using a sliding-window approach similar to (Artoni et al., 2022), and averaged across subjects. The MS-LZC measure, introduced in (Artoni et al., 2022) allows to determine the complexity of non-binarized sequences by implementing the Lempel-Ziv-Markov chain algorithm (LZMA2) for lossless data compression (Pavlov, 2013a) with maximum compression level, 64 MB dictionary, 64 FastBytes, BT4 MatchFinder, BCJ2 Filter (Pavlov, 2013b). The MS-LZC is defined as the compressed size (in bytes) of a microstate sequence. The MS-LZC was compared against that obtained with synthetic sequences ("Random" and "Surrogate"). For each subject and condition, the synthetic sequences were obtained from the real sequence by randomly drawing each sample from the pool of 5 microstates (A,B,C,D,E) so that each microstate had equal probability of appearance (Random sequence) or the same probability of appearance as in the original sequence (Surrogate sequence). The following paragraph describes the steps of Microsynt, i.e., "syntax discovery" and "syntax analysis" (Fig. 1).

A Syntax Discovery – Optimal Dictionary Size Estimation

The first key step to syntax discovery was to estimate the optimal size of the words (and consequently dictionary) with which to analyze the sequences. A dictionary is defined as a set of elements (words), each one a sequence of size *n,* and formed by drawing each sample from the pool of e.g., 5 microstates (A,B,C,D,E) without having two or more contiguous microstates repeated (e.g., {AA, BB} etc.). For example, a theoretical dictionary of word size $n = 1$ is {A; B; C; D; E} and would have a dictionary size of 5 (5 words), a theoretical dictionary of word size $n = 2$ without repetition, would be {AB; AC; AD; AE; BA; BC; BD; BE; CA; CB; CD; CE; DA; DB; DC; DE; EA; EB; EC; ED} with a dictionary size of 20. In general, a theoretical dictionary, with word size *n*, without repetition, considering 5 microstates (A, B, C, D, E) will have size $5 * 4^{n-1}$.

Each word of the dictionary can be assigned an integer number representing the number of times it appears in a real sequence. For example, considering the dictionary of word size 2 outlined above, the input sequence [ABCAB]" would have the following number of appearances of each word: {AB=2; AC=0; AD=0; AE=0; BA=0; BC=1; BD=0; BE=0; CA=1; CB=0; CD=0; CE=0; DA=0; DB=0; DC=0; DE=0; EA=0; EB=0; EC=0; ED=0}. After removing the words with null representation, it is possible to say that the "real dictionary with word size 2″ of the input sequence [ABCAB] is {AB; BC; CA}, it has size 3 (3 words) and the sequence representation is {AB=2; BC=1; CA=1}. For obvious reasons, the "theoretical" dictionary size is always greater or equal to that of the "real" dictionary, extracted from an input sequence.

All the theoretical dictionaries of word sizes n ranging from 1 to 10 were fitted to the input sequences for all subjects and conditions (to estimate the "real dictionary size") and corresponding random sequences of the same size (to estimate the "random dictionary size").

The optimum word size *n* to use in subsequent analyses was then selected as the maximum value for which there was no difference in terms of random dictionary size and real dictionary size. As the size of a theoretical dictionary grows exponentially with the size *n* of the words composing it, an exponentially longer original sequence is needed to obtain a good representation of the words of the dictionary. In fact, an exceedingly low size would fail to capture the complexity of the input sequence, while an exceedingly high size would result in an insufficient representation of the dictionary words. Here we chose a word size $n = 6$, however a lower or higher number can also be used (see Fig. 9).

A Syntax discovery - Sequence Analysis

Once the dictionary sequence size n was defined, the theoretical dictionary was fitted to the input sequences for each subject and condition to obtain the real dictionaries and corresponding sequence representations. In order to estimate the "randomness" of the sequences, the words of the dictionary were divided into "classes" according to their entropy, defined as $H = -\sum_{i=1}^{5} p(x_i) \log p(x_i)$ where $p(x_i)$ is the probability of appearance of microstate "*i*". Intuitively, low-entropy sequences are constituted mostly by repeating blocks (e.g., [ABABABAB…]) while high-entropy sequences are more "random" (e.g., [ABCBEDCE…]) with

no discernable pattern. By pooling the input sequence representations according to classes of entropy it is therefore possible to obtain the "Entropy Representation" of each input sequence, that is the number of words extracted from a sequence that belongs to each entropy class. It thus allows to determine the non-randomness of an observed microstate sequence. For example, let's consider the input sequence [ABABAC] and its dictionary representation using words of size 3 {ABA=2; BAB=1; BAC = 1}. For words of size 3 without repetition there are only two classes of entropy for words in the form "XYX" (low entropy) and "XYZ" (high entropy) respectively. Since the words [ABA] and [BAB] belong to a "low-entropy" class (i.e., "Class 1″) and BAC to a "high-entropy" class (i.e., "Class 2″), the Entropy Representation would be {EntropyClass1 = 3, EntropyClass2 = 1}.

Both Sequence Representation and Entropy Representation can be expressed in absolute numbers (number of occurrences) or percentages (proportion of occurrences). Using the examples above, the "Size 3 Representation" of the sequence [ABABAC], in terms of percentages, would be {ABA = 50%; BAB = 25%; BAC = 25%}. Similarly, its "Size 3 Entropy Representation" would be {EntropyClass1 = 75%, EntropyClass2 = 25%}. The Representation for each entropy class can be also expressed as "Entropy Representation Ratio", that is the ratio between the Entropy Representation of a class and its Theoretical Entropy representation, that is the Entropy Representation calculated on the theoretical dictionary of words of the same size (size 3 in this case). In other words, the Entropy Representation Ratio for one entropy class is calculated as the ratio between the number of sequences from the real dictionary belonging to that class and the total number of sequences forming the corresponding theoretical dictionary.

For instance, using the example above of the sequence [ABABAC], the Theoretical Dictionary of element size 3, with microstates "A", "B" and "C" would be {ABA, ABC, ACA, ACB, BAB, BAC, BCA, BCB, CAB, CAC, CBA, CBC}. Since [ABA], [ACA], [BAB], [BCB], [CAC], [CBC] belong to the low-entropy class ("Class 1″) and [ABC], [ACB], [BAC], [BCA], [CAB], [CBA] to the "high-entropy" class ("Class 2″), the theoretical Entropy Representation of the two entropy classes would be {EntropyClass1 = 50%, EntropyClass2 = 50%}. The representation ratio for entropy Class 1 would be RR1 = 75/50 = 1.5; the representation ratio for Class 2 would be RR2 = 25/50 = 0.5. In other words, in this case, it is possible to say that the sequence [ABABAC] is better represented by the low entropy class (Class 1), which has a representation ratio above 1 than the high entropy class (Class 2), with a representation ratio below 1.

Note that the number of possible entropy classes depends on the word size, the number of microstates and whether permanence is allowed. Considering the example above, i.e., words of size 3 with repetition, there would be three classes of entropy for words in the form "XXX" (lowest entropy), "XYX"/"XXY"/"YXX" (mid-entropy), "XYZ" (highest entropy). However, if only two types of microstates are available to choose from (e.g., A, B), then the form "XYZ" could not be obtained. Generally, the higher the number of microstates available and the higher the word length, the higher the number of entropy classes. It is also worth mentioning that the numerical "distance" between the minimum and maximum entropy value is not evenly sampled by the mid-entropy classes in-between. For example, considering no-permanence size 6 words with 5 microstates, the 6 possible entropy classes (Cl.1 – Cl.6) would respectively have the following entropy values: "1; 1.46,

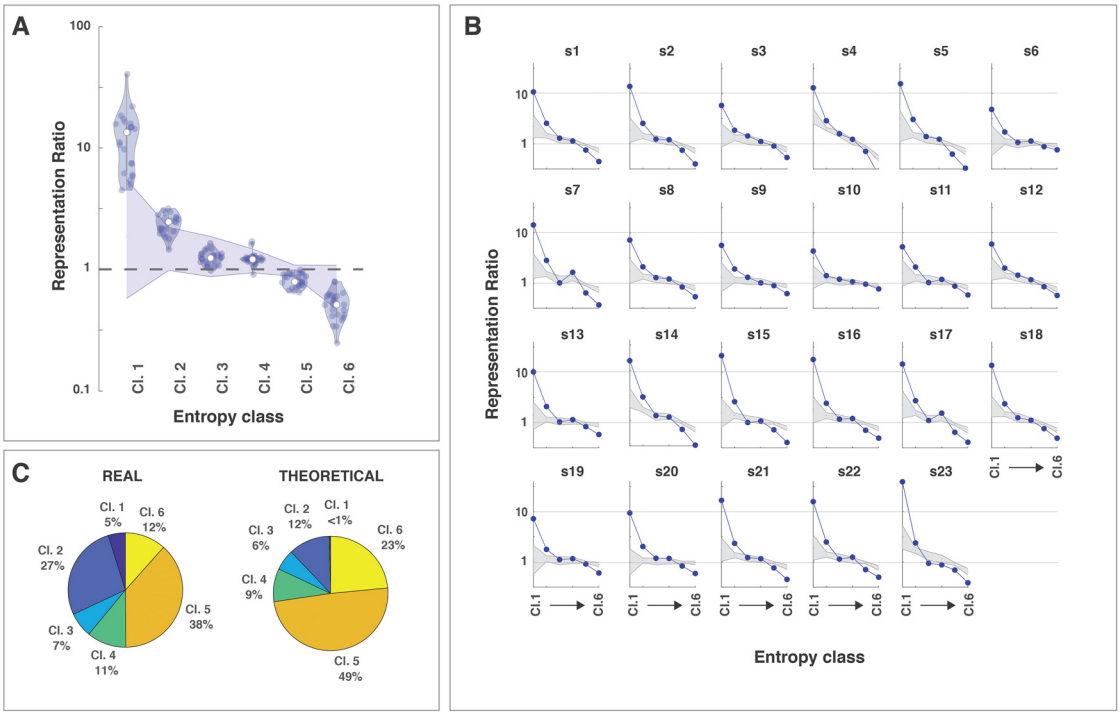

Fig. 4. Entropy Representation Ratio

(Panel A) Violin distribution of the Representation Ratio (ratio between the real Entropy Representation of a class and its theoretical Entropy Representation – see methods) for all subjects and condition "BASE" across all entropy classes (Cl.1 = 1; Cl. 2 = 1.46; Cl.3 = 1.58; Cl. 4 = 1.79; Cl. 5 = 1.92; CL.6 = 2.25) bits. The y axis is logarithmic for enhanced representation. The shaded region underlines the [2.5 – 97.5] percentile bilateral Entropy Representation Ratio surrogate distribution ($p = 0.05$ significance threshold), the horizontal dotted line indicates the threshold that discriminates between a higher (above the line) or lower (below the line) Representation Ratio with respect to neutral (i.e., no difference between the real Representation and the theoretical one). (Panel B) Representation Ratio for each subject individually and condition "BASE" for all entropy classes. Similarly to panel A, the shaded gray region represents the bilateral surrogate distribution. Representation Ratio values above or below the shaded gray region are significant (significance $\alpha = 0.05$). (Panel C) Representation of each entropy class within the real data (left) and theoretical representation of each entropy class i.e., considering the theoretical dictionary (right). .

1.58, 1.79, 1.92, 2.25″ bits. Depending on the number of entropy classes available, further analyses can be performed at the entropy class level (e.g., considering each class separately), or by grouping multiple entropy classes first (e.g., based on absolute numerical distance or behavior across conditions). In case of a larger number of entropy classes (e.g., with words of size 10), grouping of entropy classes with a close numerical value of entropy might be unavoidable to obtain a sufficient representation.

A Syntax discovery - Surrogate Statistics

To determine the statistical significance of Sequence and Entropy representations, here we used, for each subject and condition, a surrogate-bootstrapping data approach. First the input unprocessed sequence is surrogated by randomly shuffling the order of the microstates. For instance, starting from the sequence [AABCCC], examples of surrogate would be [AACCCB] or [BAACCC]. Then the surrogate sequence undertakes the same preprocessing steps described in section **II.C.** to obtain its no-permanence version as well as the steps described in sections **II.E** and **II.F** to obtain the "size n" Surrogate Sequence Representation, Surrogate Entropy Representation, expressed in percentages (proportion of occurrences), as well as the Surrogate Entropy Representation Ratio for each entropy class. A surrogate sequence may result in a different probability of each microstate with respect to the original no-permanence sequence. For example a sequence such as a = [AAAAABBBBBAAAAACCCCC], yielding the no-permanence sequence [ABAC], could be surrogated into a' = [AAAAAAAAAABBBBBCCCCC], yielding [ABC]. For this reason the whole surrogate-bootstrapping data approach is then repeated 1000 times to estimate the surrogate distributions of each measure. For example, statistical significance on the Entropy Representation Ratio was set at values outside the 95th percentile of the Surrogate Entropy Representation Ratio distribution i.e., [2.5% – 97.5%]).

A Syntax analysis

The Entropy Representation Ratio was compared across entropy classes both considering each subject separately (Fig. 4, Panel B) and all subjects together, for condition "BASE" (Fig. 4, Panel A).

Then, the representation of the low-entropy classes (Cl.1 and Cl.2) and high-entropy classes (Cl.5 and Cl.6) was compared across conditions (BASE, DEEP) and the corresponding surrogate distribution (Fig. 5). Normalization was performed in such a way that the sum of the representation for one condition and entropy classes considered amounted to 100%. After testing the null hypothesis of data Normal distribution over each group using a Kolmogorov Smirnov test (significance set at $\alpha = 0.05$), these measures were compared across conditions using a Kruskal Wallis test followed by a Tukey's Honest Significant Difference (HSD) criterion for post-hoc comparisons and multiple testing correction.

The most significantly represented entropy class, compared to the theoretical distribution, was Cl.1, representing by the lowest entropy words (Fig. 6). These words are formated as [XYXYXY], and considering the 5 microstates, only 10 combinations are possible representing microstates "hubs" [AB], [AC], [AD], [AE], [BC], [BD], [BE], [CD], [CE], [DE]. Twin words ([ABABAB] and [BABABA]) were pooled together. In fact, a sequence such as [ABABABA] would originate two "size 6″ words, [ABABAB] and [BABABA], which should probably be considered together as belonging to a single "AB" hub given their similarity. This precaution allows to compensate for the possible artifact generated when unpacking the original sequence using a fixed size. The different representations of these "oscillating" sequences within CL.1 (normalized to 100%) were tested for significance across conditions (BASE and DEEP) as previously described (Kolmogorov Smirnov test, Kruskal Wallis and Tuckey's HSD).

Extending the analysis to the other classes of entropy, it is possible to observe that size 6 words can be represented using a 5-point graph, with lines connecting the microstates with width proportional to the number of repetitions of the same "path". For example, sequences in the form [ABABAC] would be represented by a wide line connecting A and B (hub) and a thinner line connecting A and C (satellite). To extract what could be called "the syntax signature" that best differentiates the two conditions, BASE and DEEP, all possible words of the dictionary were ranked in terms of the absolute value of the difference of representation within the data across the two conditions (averaged across all subjects), after removing those not reaching significant representation (against surrogate distribution). The topmost 30 words (15 with representation for BASE greater than for DEEP and, vice versa, 15 where the representation for BASE was lower than for DEEP) were retained. This pool of sequences (regardless of which entropy class they belonged to) totaled an overall representation of 10% and could be defined as the "10% syntax signature" that best differentiates the two conditions. The "absolute values of the difference" represented in Fig. 7 were further normalized to 100%. The error bars represent the "standard error of the difference".

A Analysis with permanence and different word sizes.

We also demonstrated the use of Microsynt also without removing permanence, (Fig. 8) i.e., without the "sequence preprocessing" steps described in **section II.C,** and/or by selecting a different word size ranging from 3 to 7 (Fig. 9).

3. Results

A K-means cluster analysis

The k-means cluster analysis and the meta-criterion revealed 5 microstates as best explaining the data (see (Artoni et al., 2022) for details of this analysis). The microstate maps for both conditions and the mean maps across the 2 conditions are shown in Fig. 3. The microstates were ordered according to the canonical microstates described earlier, e.g. (Artoni et al., 2022; Bagdasarov et al., 2022; Bréchet et al., 2019, 2020; Tomescu et al., 2022). Note that the 5th microstate is labeled here as microstate E instead of microstate F as in Artoni et al., 2022.

A Microstate sequences are not random

The Lempel-Ziv complexity analysis of microstates sequences reveals that, even at rest, sequences are actually not random. In fact, the complexity of the input sequences is significantly lower than the complexity of derived random or surrogate sequences. Given the length of the input sequences, the analyses showed that "6″ is the optimal word size, that is the highest size that allows to capture the complexity of the input sequence, while maintaining sufficient representation of each dictionary element. In fact, starting from size 7, the "random dictionary size" diverges from the "real dictionary size".

A Spontaneous brain activity favors low-entropy (simpler) microstates syntax

As shown in Fig. 4, panel C, Entropy Cl. 1 and Cl.2 respectively represent the 5% and 27% of the real data, as opposed to the 1% and 12% of the (expected) theoretical representation. Accordingly, the Representation Ratio (Fig. 4, panel A) of entropy Cl. 1 is on average above 10, meaning that Cl. 1 is roughly 10 times more represented in the data than it would be, in a surrogate sequence of the same length. The higher Representation Ratio of low-entropy classes (Cl. 1 and 2) is balanced by a lower Representation Ratio of high-entropy classes (Cl. 5 and 6). It is possible to infer that brain activity is best captured by low-entropy sequences of microstates (words), or, in other terms, that a microstate sequence, based on EEG data, tends to be simpler than artificial/random/surrogate synthetic derived sequences. Fig. 4, panel B demonstrates how this behavior holds true for all patients with only minor differences. Cl. 1 and Cl. 2 are always significantly more represented than in surrogate data. On the contrary, Cl. 5 (except for a few subjects) and Cl. 6 are significantly less represented than in surrogate data. Due to the common behavior of Cl.1/Cl.2 and Cl.5/Cl.6, throughout

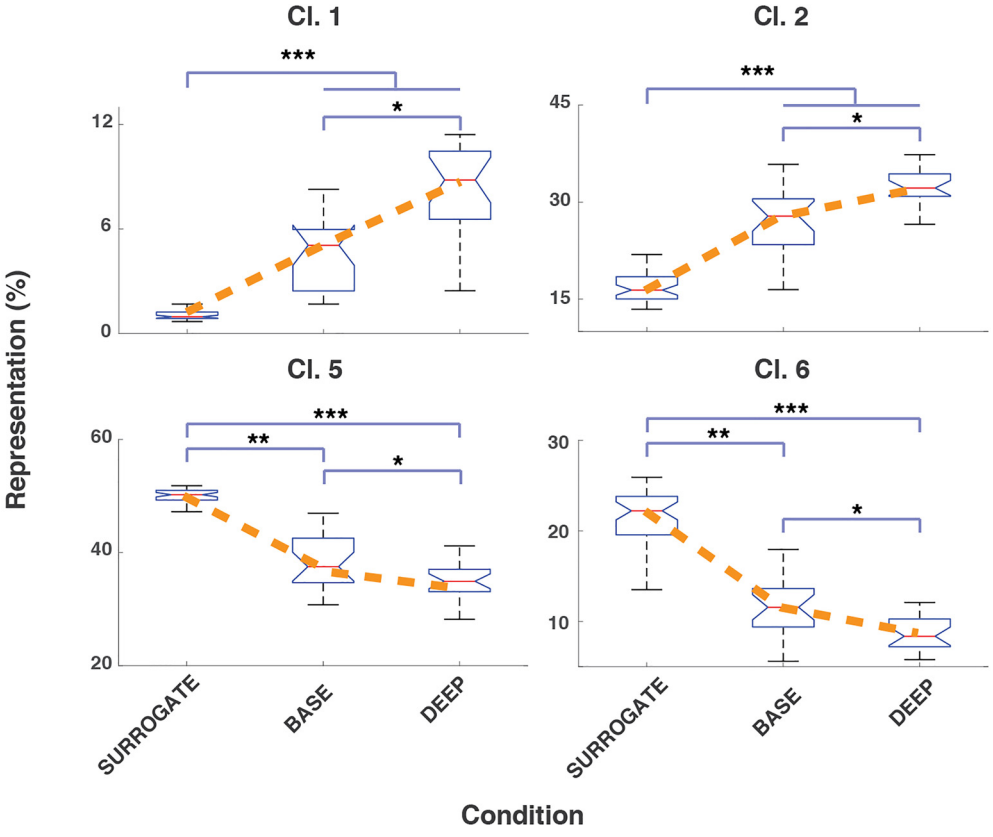

Fig. 5. Entropy class representation across conditions

Comparison of Entropy Representations for low entropy classes (Cl.1, Cl. 2) and high-entropy classes (Cl. 5, Cl. 6) across conditions (BASE, DEEP) and surrogate data (SURROGATE) and considering all subjects. Values are normalized to 100% (e.g., the sum of the representation for BASE and Cl.1, Cl. 2, … Cl. 6 is 100). * = $p<0.05$; ** = $p<0.01$; *** = $p<0.001$.

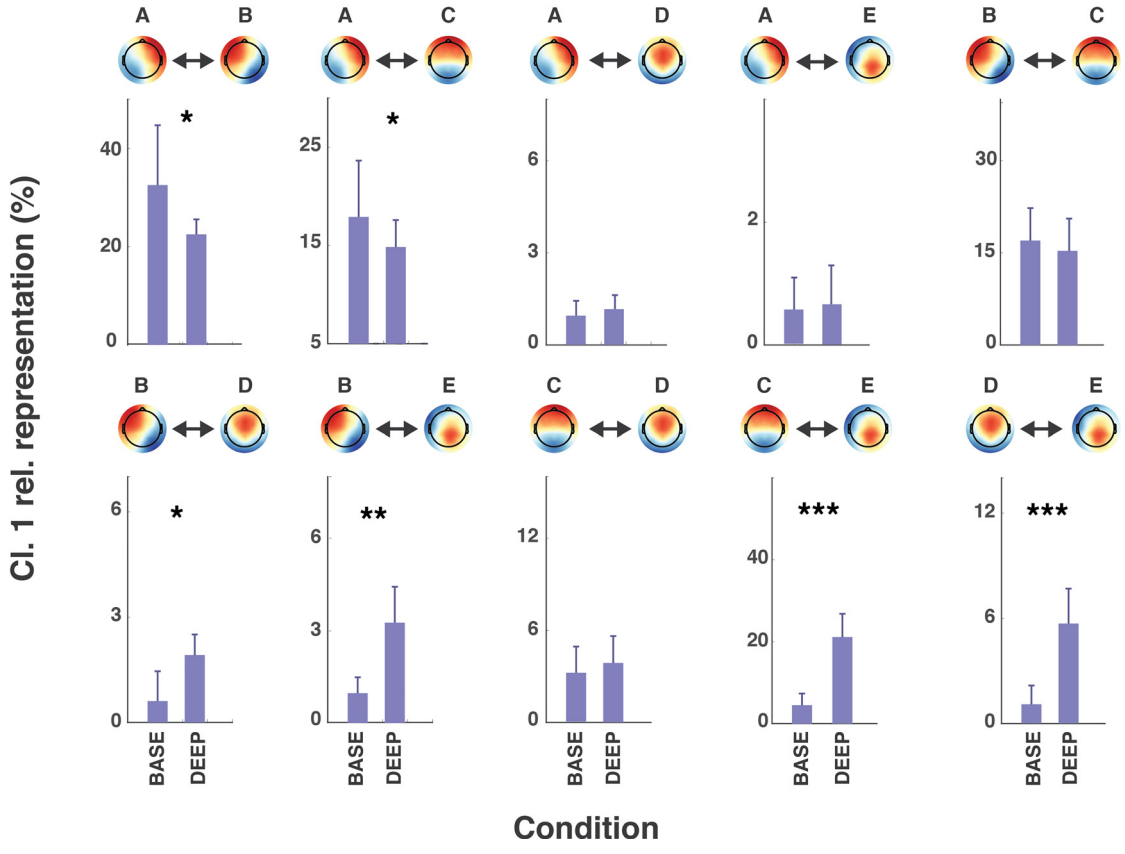

Fig. 6. Sequence representation within entropy Cl. 1

Relative representation (%), within entropy class 1 (Cl. 1), considering all subjects, for each of the 10 types of size 6 elements. For example, the topmost left panel represents the sequences [ABABAB] and [BABABA] pooled together. Values for each condition are normalized to 100% (e.g., the sum of the representation for BASE and *A*<->*B*, *A*<->C,…, *D*<->*E* is 100). * = $p<0.05$; ** = $p<0.01$; *** = $p<0.001$.

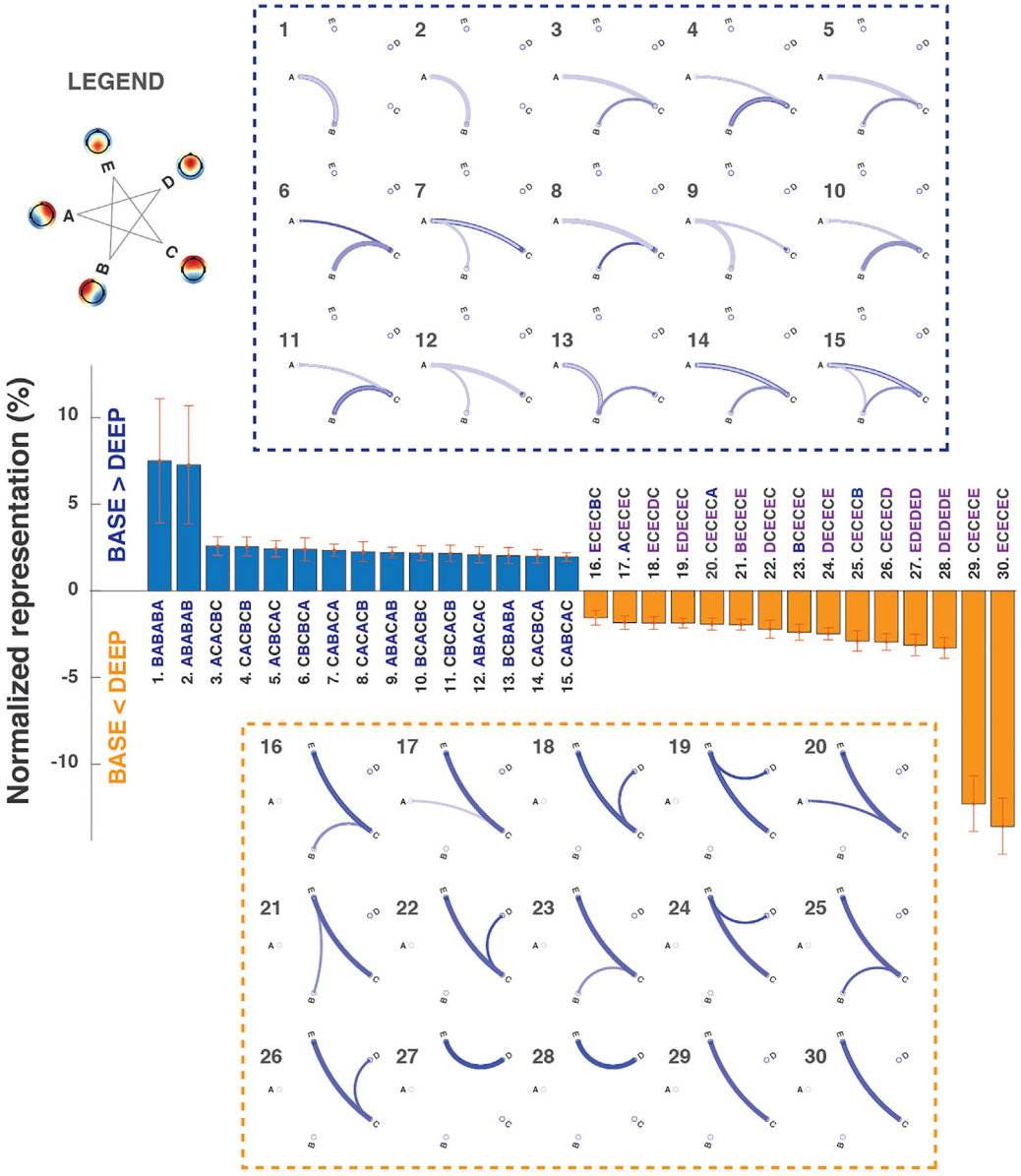

Fig. 7. Syntax signature comparison across conditions
Normalized representation (to 100%) across all subjects of the best 30 elements (words) of size 6 that best differentiate conditions BASE and DEEP, and accounting for 10% of the full dictionary representation. Each bar represents the average representation across subjects of the word indicated in the caption (e.g., ACACBC) for BASE minus that of DEEP. Positive and negative bars respectively represent words that best identify BASE and DEEP conditions and are colored in blue and orange. Error bars represent the standard error of the measure.
Words are also represented in a starfish diagram by a minimum of 1 and maximum of 3 lines connecting a minimum of 2 and a maximum of 3 points of the diagram. Sequences such as [ABABAB] are represented by a single wide line connecting "A" and "B". Sequences such as [ABABCA] are represented by a wide line connecting "A" and "B" and a thin line connecting "A" and "C".

the manuscript we refer to them as "low-entropy" and "high-entropy" classes to simplify the description of the results.

A Low and High-entropy sequences are sensitive to loss of consciousness

As shown in Fig. 5, the representation of entropy classes Cl.1, Cl.2, Cl. 5 and Cl. 6 is significantly modulated by the conditions (BASE and DEEP). Also considering surrogates, for low entropy classes (Cl. 1 and Cl. 2) it is possible to observe a significant upward trend differentiating BASE/DEEP and SURROGATE ($p<0.001$), as well as BASE and DEEP ($p<0.05$). The opposite is true for high entropy classes (Cl. 5 and Cl. 6) with a marked downward trend where SURROGATE > DEEP ($p<0.001$); SURROGATE > BASE ($p<0.01$); BASE > DEEP ($p<0.05$).

A Microstate sequences favor oscillations between two microstates

The analysis of the words forming Cl. 1, which has the highest representation ratio, allows to determine which "hubs" of microstates best discriminate different conditions of consciousness. Fig. 6 shows a significantly higher representation within CL.1 of hubs C<->*E* and C<->*D* ($p<0.001$), *B*<->*D* ($p<0.05$) and *B*<->*E* ($p<0.01$) for DEEP with respect to BASE. Conversely, a significantly higher representation of hubs *A*<->*B* and *A*<->C ($p<0.05$) can be found in BASE with respect to DEEP,

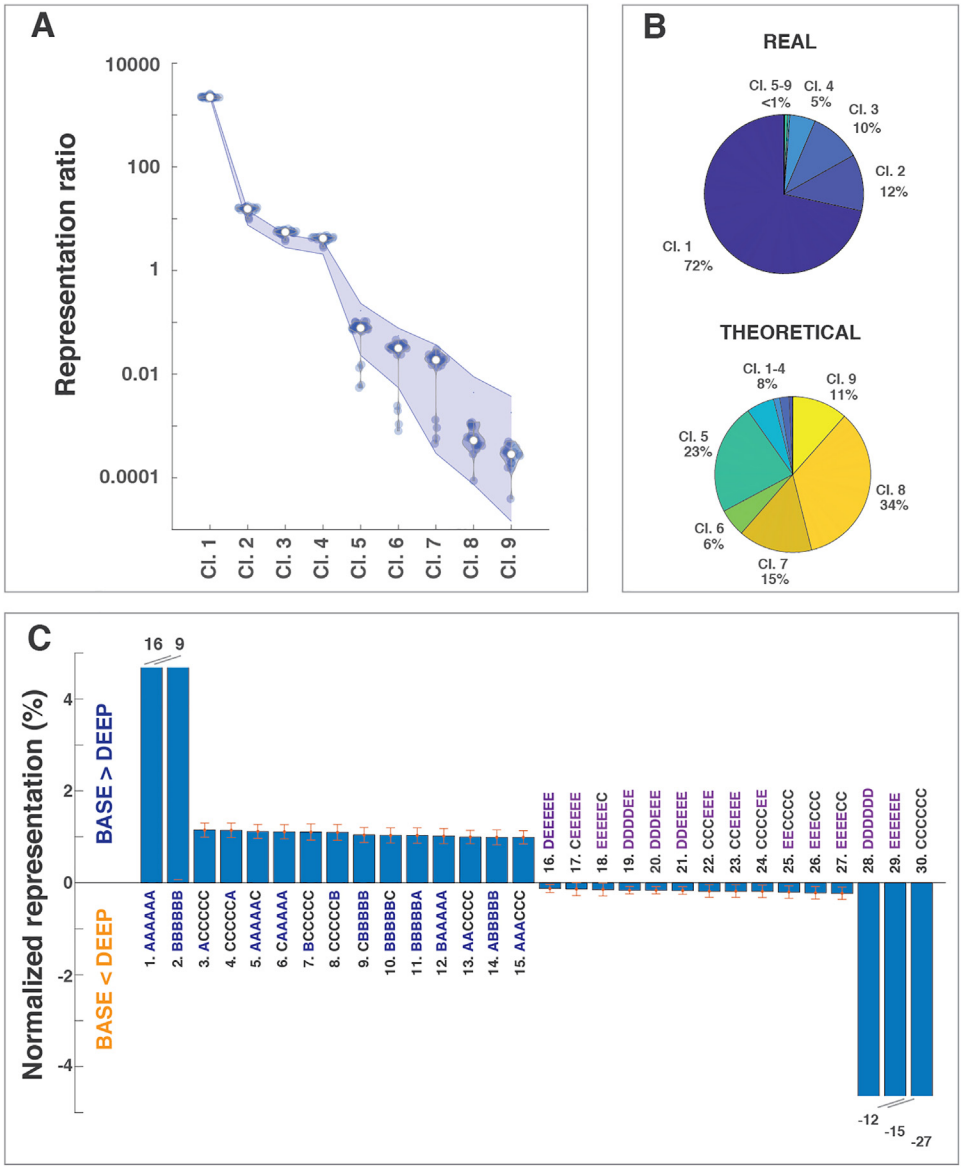

Fig. 8. Syntax signature comparison and representation with permanence

(Panel A) Violin distribution of the Representation Ratio for all subjects and condition "BASE" across all entropy classes, considering permanence. Entropy values (in bits) for each class are: Cl.1 = 0; Cl. 2 = 0.65; Cl.3 = 0.92; Cl. 4 = 1; Cl. 5 = 1.25; Cl.6 = 1.46; CL.7 = 1.58; CL.8 = 1.79; CL.9 = 1.92). See also caption of Fig. 3.

(Panel B) Representation of each entropy class within the real data (top) and theoretical representation of each entropy class i.e., considering the theoretical dictionary (bottom). See also caption of Fig. 3.

(Panel C) Normalized representation (to 100%) across all subjects of the 30 size 6 words that best differentiate conditions BASE and DEEP, and accounting for 10% of the full dictionary representation. See also caption of Fig. 6.

suggesting a shift towards microstates C and E during loss of consciousness.

A Low-entropy sequence hubs of microstates

A further analysis was carried out on the statistically significant words that best differentiated BASE and DEEP conditions, forming the overall "10% syntax signature", regardless of their belonging to a specific entropy class. Fig. 7 shows how words composed of microstates A and B are most represented while fully awake with respect to fully unconscious. The opposite is true for words composed of microstates C and E that are mostly prominent while fully unconscious with respect to fully awake. Microstate C functions as a network hub for both conditions. Fig. 8 panel C also confirms the importance of microstate C in both conditions, with the greatest representation for the word [CCCCCC] in DEEP (~27%). Interestingly it is possible to qualitatively observe how the normalized representation of words "ABABAB" and "BABABA" present greater variability across subjects than the words "CECECE" and "ECECEC".

A Results without removal of permanence and with different words sizes

Microsynt can be applied to any type of sequence regardless of removal of permanence. Fig. 8 panel A shows the entropy representation ratio of the 9 entropy classes obtained using size 6 words for conditions BASE. The lowest entropy class (corresponding to a sequence of the type [XXXXXX]) has the greatest representation ratio (~3000), while the highest entropy class (corresponding to a sequence of the type [XYZKWY]) the lowest representation ratio (~1/5000). As Panel B shows, Cl.1 represents ~72% of the data and Cl.1–4 almost 99%. On the contrary, the theoretical representation of Cl.1–4 is around ~8%. Considering the statistically significant words that best differentiated BASE and DEEP conditions (Fig. 8, panel C), the ones composed by microstates A and B remain the most represented microstates characterizing the fully-awake condition, while D and E best characterize the fully-unconscious condition. Microstate C seems to function as a network hub for both conditions. The higher Representation Ratio of low-entropy classes, balanced by a lower Representation of high-entropy classes can also be observed with different word sizes (Fig. 9). For example, considering $n = 3$, Cl.1 words of type [XYX] have a significant representation ratio of 1.5, while words of type [XYZ] have a representation ratio of 0.8. The higher the word size the higher the representation

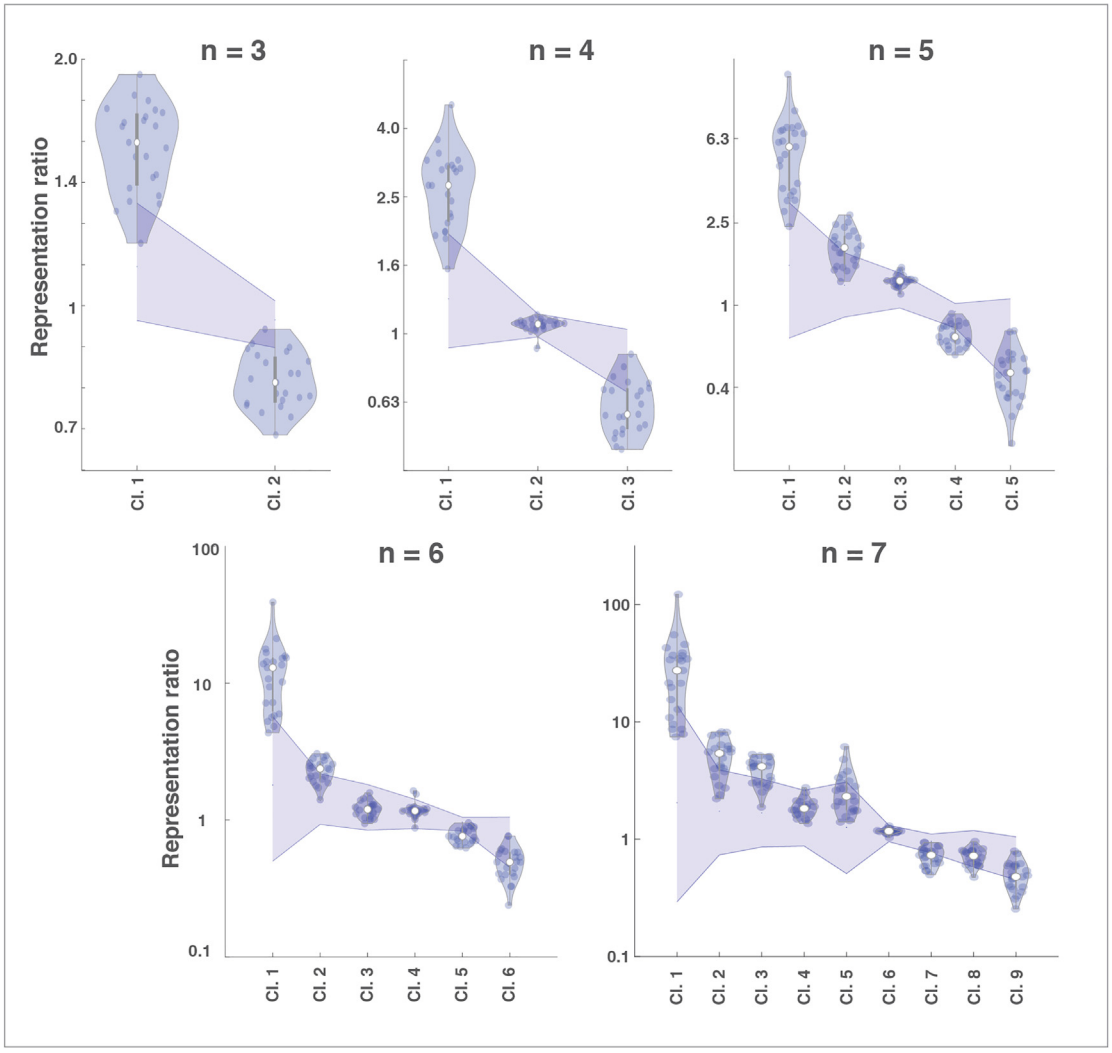

Fig. 9. Representation Ratio with different word sizes.
Violin distribution of the Representation Ratio (ratio between the real Entropy Representation of a class and its theoretical Entropy Representation – see also methods and caption of Fig. 3) for all subjects and condition "BASE" across all entropy classes and word sizes ranging from 3 to 7. Entropy values (in bits) are as follows.
n = 3: Cl.1 = 0.92, Cl.2 = 1.58;
n = 4: Cl.1 = 1, Cl.2 = 1.5, Cl.3 = 2;
n = 5: Cl.1 = 0.97; Cl.2 = 1.37, Cl.3 = 1.52, Cl.4 = 1.92, Cl.5 = 2.32
n = 6: Cl.1 = 1; Cl.2 = 1.46; Cl.3 = 1.58; Cl.4 = 1.79; Cl.5 = 1.92; Cl.6 = 2.25;
n = 7: Cl.1 = 0.99; Cl.2 = 1.38; Cl.3 = 1.45; Cl.4 = 1.56; Cl.5 = 1.66; Cl6 = 1.84; Cl.7 = 1.95; Cl.8 = 22.13; Cl.9 = 2.24

ratio of the lowest entropy classes (~10/100 for Cl. 1 and n = 7), in relation to other classes. This confirms that a microstate sequence, both with and without permanence, tends to have lower entropy than any artificial/random/surrogate derived sequence.

4. Discussion

Here we described a novel method, "Microsynt", to advance our understanding of how microstates are combined into a sequence to form a relevant "syntax signature" that can both successfully characterize resting state data and differentiate conditions/brain states. One of the main features of Microsynt is the so-called "explanation facility", defined as "the ability [for a method/system] to provide explanations that clarify its functioning and recommendations" (Wick and Slagle, 1989). Depending on the length and complexity of the original sequence of microstates, it is possible to determine the optimal size of the descriptive sequence as the one that generates a vocabulary, sufficiently represented both by real and surrogate data (see methods). The problem of exponential vocabulary growth in relation to element size n of the dictionary could be solved by sorting the words into entropy classes, thus making it possible to focus further analyses to the classes that exhibited a significant Entropy Representation Ratio, either because outside the range of their surrogate distribution, or because they were significantly modulated across conditions. Given these considerations, here we applied Microsynt with word size of 6. However it is important to note that the method can be also used with lower (>2), or higher words sizes. As shown in Fig. 9, the lower the word size, the lower the number of available entropy classes that the data can be sorted into, the higher the representation of each entropy class and the lower the complexity of the analysis. Interestingly, all analyses, regardless of word size, agree on a higher Representation Ratio for low-entropy classes, balanced by a lower Representation of high-entropy classes (Fig. 9). A dictionary with word sizes n = 1 and n = 2 would correspond to calculating the state occurrence probability, and the state transition probability respectively.

As low-represented words are grouped, comparisons across conditions in terms of representation ratio are likely more relevant than when

only considering single words. However, with the increase of the word size and number of microstates available and the consequent increase of possible entropy classes, some entropy classes might still be insufficiently represented. In that case the user can choose to group some entropy classes together to increase representation. Grouping can be performed according to different factors, e.g., absolute entropy values, absolute representation ratio or representation ratio across conditions. Words can also be studied in terms of their representativeness within each entropy class. This facilitates the comparison of sequences belonging to the same "category" (e.g., see Fig. 6).

For example, the simplest words of a certain size n are those iterating the same couple of microstates, i.e., {XYXY…}, indicating a "binary" loop. Comparing the probability of occurrence of words belonging to the same class allows to determine which binary loop is prominent (e.g., "ABAB…" or "ACAC…").

On the other hand, it is possible that comparing the input sequence representation of words belonging to two different entropy classes could be biased by the a-priori theoretical probability of occurrence within each class. For example, considering a pool of 5 microstates, the number of possible binary loops is 20, but the number of possible ternary loops such as {XYZXYZ…} is 60. While a correction for the a-priori probability of each single word is possible, (i.e., without considering entropy classes), the effectiveness of this normalization may be hindered by a-priori probabilities close to zero, especially for high-entropy words. The normalization for the a-priori probability of occurrence (based on surrogates) for entropy classes ("Entropy Representation Ratio" - see methods and Fig. 4), instead, is much more reliable as many low-represented words are grouped together. Plus, representation comparisons of different words across conditions are easier if performed within a single entropy class independently. On this aspect, Fig. 5 highlights two opposed trends of representation comparing surrogates, fully awake and fully unconscious conditions within low entropy and high entropy classes.

Representation ratios of the top 10% single words (or another percentage of choice), in terms of significance and representativeness, regardless of their belonging to classes, can however be used to form a "syntax signature" (Fig. 7) which can be highly useful as a decoding feature for classifiers with the aim of differentiating two or more classes, while maintaining optimal explanation facility. The "star" representation of words (Fig. 7, top and bottom panels) gives an at-a-glance representation of the syntax signatures of the two conditions.

The results of the application of Microsynt on real EEG data show that microstate sequences, even at rest, are indeed not random but tend to behave in a more predictable way, favoring simpler sub-sequences, or "words". As Fig. 4 shows, lowest-entropy (class 1), binary microstate loops are prominent, (e.g., "ABABAB"…) and favored 10 times more, on average than what was theoretically expected. The inverse is true for high-entropy sequences. This suggests the presence of a reduced complexity, even at rest, with respect to what is expected from a random or surrogate sequence - which maintains the same a-priori probabilities of occurrence of each microstate. Murphy et al., using sample entropy on data collected from patients with early-course psychosis, also observed the prominent presence of A-B and C-D binary loops (Murphy et al., 2020). This binary periodicity preference may be also linked to the observation by Wegner et al. (von Wegner et al., 2017) that showed how resting state microstates sequences are non-Markovian processes which inherit periodicities from the underlying EEG dynamics.

The authors also observed that microstates sequences display more complex temporal dependencies than what is captured by the classic transition probabilities, but likely a finite memory content. Note that Microsynt analyzes the statistical significance (against surrogates) of recurring words in the sequence over time. This is a different approach with respect to the one presented by (Van de Ville et al., 2010), which focuses instead on long-range dependencies and fractal properties of sequences of microstates with permanence.

Interestingly, our results suggest that the representation of words classified by entropy classes can be used to differentiate patterns and trends across conditions. For example, Fig. 5 shows, also including surrogates, a positive trend of representation for low entropy classes and a negative trend of representation for high entropy classes. This is further proof of the simplification of the structure of sequences, or maybe the "language", induced by general anesthesia. However, DEEP and BASE are not just different in terms of complexity of their representation, but also in its contents. As Fig. 6 shows, microstate E, in relation with microstates C, D and B, forms significantly binary loops (e.g., CECECE), much more prominent during general anesthesia than rest. Conversely, binary loops involving microstates A, B are more prominent during awake than general anesthesia conditions. This shift towards loops involving microstates C and E during unconsciousness is also confirmed by the syntax signatures of the two conditions shown in Fig. 7, even by relaxing the constraint of having to belong to entropy class 1.

The dominance of microstate sequence A-B during wakefulness and sequence C-E during general anesthesia corresponds well to the putative functional significance of these microstates. Microstates A and B have been associated with externally-oriented sensory-cognitive processes. Simultaneous EEG-fMRI as well as high-density EEG source imaging revealed temporal brain areas including auditory cortices as generators of microstate A and occipital brain areas including the visual cortex as generators of microstate B (Britz and Michel, 2010; Custo et al., 2017). (Seitzman et al., 2017) demonstrated increase in presence of microstate B when transitioning from an eyes closed to an eyes open condition. (Bréchet et al., 2019) observed an increase of microstate B (source-localized in the visual cortex) during visual imagination of autobiographical memories. On the other hand, Microstates C and E were previously associated to the posterior and anterior part of the Default Mode Network, respectively (Bréchet et al., 2019; Custo et al., 2017) and have been interpreted as reflecting internally-generated mental activity (Tomescu et al., 2022). The increased iteration between microstates A and B during the awake state might thus reflect repetitive conscious scanning of sensory input, while the increased C-E sequences during general anesthesia reflect spontaneous internally-generated non-conscious brain activity. The significance of microstate C during full unconsciousness is also confirmed by the analysis with permanence which highlights the highest representation of the size-6 words composed only by the repeating microstate C [CCCCCC]. Note that the Mycrosynt approach already corrects for the a-priori probability of microstates during each condition by statistically comparing word representation ratios against surrogates of the original sequence, which preserves the original distribution of microstates within each condition. For example, the increased iterations between microstates A and B during the awake state reported are significantly higher than the same iterations between microstates A and B in surrogate sequences, and the same is true for every word forming the syntax signature reported in Fig. 7.

It is important to underline that Microsynt can be applied also to microstate sequences with permanence, leading to different avenues of analysis. While retaining the permanence would allow to consider the effect of the duration of microstates, removing the permanence allows to conveniently avoid repeating states in the dictionary as well as several other advantages outlined as follows

- First, with permanence, a much longer word size is required to capture sequences. For example, a sequence such as [AAAABBBBBB-BAAAAABBBBBAAAAABBBBB] would require a "size 30″ word to capture the ABABAB binary loop, with an exponentially increased computational power required for the analysis.
- Second, information about the sequence of microstates may be lost among the different classes of entropy generated by small differences in duration and thus diluted. For example, the two sequences a = [AAAABBAAABBBBAA]; b = [AAABBBAAABBBAAA], would both amount to c = [ABABA] when removing permanence (skipping

the filtering process of microstates lasting less than 5 successive time points). However, *a* and *b* would fall into different entropy classes if considered with permanence. Also, Fig. 8, panel B shows that Cl.1 sequences representing just one repeating microstate ([XXXXX]) account for ~75% of the data, basically leaving no room for the representation of higher-entropy words, unless a high word size is used.

- Third, whether to remove or not duration is a preprocessing choice that can result in different values of complexity/entropy. Higher entropy can be generated by both a reduced duration of microstates and a change in the complexity of the sequence. For example, the sequence *a* = [AAAABBBBAAAABBBBAAAABBBB], has a low entropy as it is composed by a 4-times repeating sequence of A and B microstates. An increased entropy could both be obtained by changing the durations of microstates (e.g., *b* = [AABBBAAAAAAABBBBABBBBBBB]), or by changing the actual sequence (*c* = [AAAABBBBCCCCAAAACCCCAAAA]). We believe that it is therefore best to consider both types of analysis into account. A first analysis may focus on first-order microstates features (i.e., density of occurrence, duration, global explained variance etc.) and second order ones (i.e., transition probabilities) permanence. A second analysis can then focus on the syntax, by applying microsynt on permanence-removed data. This second analysis allows to highlight changes in the data that *only* consider the structural properties of the sequence and are not influenced by duration.

Another aspect to consider when selecting a "permanence" or "no-permanence" approach is whether to allow duration of microstates as a differentiating factor in recurrence over time. For example, considering the permanence sequence [AAAA**BBCCBBCCAAAAAAA** DDDDEEEE**BBBBCCBCAAAAAA**BDBBBACCC], the parts of the sequence highlighted in bold would not be considered equal due to the different microstate durations, but they would be exactly equal in the corresponding no-permanence sequence [A**BCBCA**DE**BCBCA**BDBAC]. Removing permanence also allows Microsynt to capture a level of self-similarity (i.e., property of a sequence to have a similar pattern at a different time scale) in a sequence with permanence. For example, considering a variation on the example above, in the permanence sequence [AAAA**BBCCBBCCAAA**DDDDEEEE**BBBBCCCCBBBBCCCCAAAAAA** BDBBBACCC], the parts highlighted in bold, could be considered self-similar and would not be recognized as such by Microsynt, having different lengths. They would however be exactly equal in the no-permanence sequence [A**BCBCA**DE**BCBCA**BDBAC] and therefore could be captured by Microsynt.

A potential issue when exploring a sequence is where and how to define the beginning and end of a word. For example, the sequence [ABAB] might be considered (i) one word [AB], repeated two times, (ii) a single word [ABAB] repeated one time, or even (iii) a different word [BA] preceded and followed by unary words [A] and [B] respectively. This issue was partially solved by studying all possible words at each select size, with maximum overlap (e.g., the sequence [ABCAB] would generate the following size-3 sequences: {ABC; BCA; CAB}). As a further study It will be interesting to determine at which point (if any) the optimal word size would saturate its value, even considering longer recordings (e.g., hours or days). Finally, an extension of Microsynt would be that of considering both shorter and longer words at the same time (e.g., the representation of mixed vocabularies such as [AB; ABCABC; CDCD]).

Declaration of Competing Interest

None

Data availability

At the time of data collection, the signed consent form of the patients did not include the permission for data sharing. Sharing the data could be arranged with a specific data sharing agreement upon reasonable request.

Acknowledgments

The work was supported by the Swiss National Science Foundation (grant No. 320030_184677) and by the National Center of Competence in Research (NCCR) 'SYNAPSY - The Synaptic Bases of Mental Diseases' (SNF, grant n° 51AU40_125759). It was also supported by the Development Fund of the Medical Directories of the Geneva University Hospitals.

Credit Authors Statement

Fiorenzo Artoni: Conceptualization, Methodology, Software, Formal analysis, Data curation, Data analysis, Writing - original draft, Visualization, Investigation Project administration. Julien Maillard: Data collection. Juliane Britz: Data collection. Denis Brunet: Validation. Christopher Lysakowski: Project administration. Martin R. Tramèr: Resources, Project administration. Christoph M. Michel: Resources, Project administration, Funding acquisition, Writing –review & editing.

References

Antonova, E., Holding, M., Suen, H.C., Sumich, A., Maex, R., Nehaniv, C., 2022. EEG microstates: functional significance and short-term test-retest reliability. Neuroimage: Reports 2, 100089.
Artoni, F., Fanciullacci, C., Bertolucci, F., Panarese, A., Makeig, S., Micera, S., Chisari, C., 2017. Unidirectional brain to muscle connectivity reveals motor cortex control of leg muscles during stereotyped walking. NeuroimageNeuroimage 159, 403–416.
Artoni, F., Maillard, J., Britz, J., Seeber, M., Lysakowski, C., Bréchet, L., Tramèr, M.R., Michel, C.M., 2022. EEG microstate dynamics indicate a U-shaped path to propofol-induced loss of consciousness. Neuroimage, 119156.
Artoni, F., Menicucci, D., Delorme, A., Makeig, S., Micera, S., 2014. RELICA: a method for estimating the reliability of independent components. NeuroimageNeuroimage 103, 391–400.
Bagdasarov, A., Roberts, K., Bréchet, L., Brunet, D., Michel, C.M., Gaffrey, M.S., 2022. Spatiotemporal dynamics of EEG microstates in four-to eight-year-old children: age-and sex-related effects. Dev. Cogn. Neurosci., 101134.
Bréchet, L., Brunet, D., Birot, G., Gruetter, R., Michel, C.M., Jorge, J., 2019. Capturing the spatiotemporal dynamics of self-generated, task-initiated thoughts with EEG and fMRI. NeuroimageNeuroimage 194, 82–92.
Bréchet, L., Brunet, D., Perogamvros, L., Tononi, G., Michel, C.M., 2020. EEG microstates of dreams. Sci. Rep. 10, 1–9.
Bressler, S.L., Menon, V., 2010. Large-scale brain networks in cognition: emerging methods and principles. Trends Cogn. Sci. (Regul. Ed.) 14, 277–290.
Britz, J., Michel, C.M., 2010. Errors can be related to pre-stimulus differences in ERP topography and their concomitant sources. NeuroimageNeuroimage 49, 2774–2782.
Brunet, D., Murray, M.M., Michel, C.M., 2011. Spatiotemporal analysis of multichannel EEG: CARTOOL. Computational intelligence and neuroscience 2011.
Cabral, J., Kringelbach, M.L., Deco, G., 2014. Exploring the network dynamics underlying brain activity during rest. Prog. Neurobiol. 114, 102–131.
Chernik, D.A., Gillings, D., Laine, H., Hendler, J., Silver, J.M., Davidson, A.B., Schwam, E.M., Siegel, J.L., 1990. Validity and reliability of the Observer's Assessment of Alertness/Sedation Scale: study with intravenous midazolam. J. Clin. Psychopharmacol..
Custo, A., Van De Ville, D., Wells, W.M., Tomescu, M.I., Brunet, D., Michel, C.M., 2017. Electroencephalographic resting-state networks: source localization of microstates. Brain Connect. 7, 671–682.
Deco, G., Jirsa, V.K., 2012. Ongoing cortical activity at rest: criticality, multistability, and ghost attractors. J. Neurosci. 32, 3366–3375.
Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J. Neurosci. Methods 134, 9–21.
Fox, M.D., Greicius, M., 2010. Clinical applications of resting state functional connectivity. Front. Syst. Neurosci. 4, 19.
Gschwind, M., Michel, C.M., Van De Ville, D., 2015. Long-range dependencies make the difference-Comment on" A stochastic model for EEG microstate sequence analysis". NeuroimageNeuroimage 117, 449–455.
Hochman, H.M., Rodgers, J.D., 1969. Pareto optimal redistribution. Am. Econ. Rev 59, 542–557.
Jirsa, V.K., Fuchs, A., Kelso, J.A., 1998. Connecting cortical and behavioral dynamics: bimanual coordination. Neural Comput. 10, 2019–2045.
Koenig, T., Prichep, L., Lehmann, D., Sosa, P.V., Braeker, E., Kleinlogel, H., Isenhart, R., John, E.R., 2002. Millisecond by millisecond, year by year: normative EEG microstates and developmental stages. NeuroimageNeuroimage 16, 41–48.
Lehmann, D., Faber, P.L., Galderisi, S., Herrmann, W.M., Kinoshita, T., Koukkou, M., Mucci, A., Pascual-Marqui, R.D., Saito, N., Wackermann, J., 2005. EEG microstate duration and syntax in acute, medication-naive, first-episode schizophrenia: a multi–center study. Psychiatry Res. Neuroimaging 138, 141–156.
Mesulam, M., 2008. Representation, inference, and transcendent encoding in neurocognitive networks of the human brain. Ann. Neurol. 64, 367–378.
Michel, C.M., Brunet, D., 2019. EEG source imaging: a practical review of the analysis steps. Front. Neurol. 10, 325.

Michel, C.M., Koenig, T., 2018. EEG microstates as a tool for studying the temporal dynamics of whole-brain neuronal networks: a review. NeuroimageNeuroimage 180, 577–593.
Murphy, M., Stickgold, R., Öngür, D., 2020. Electroencephalogram microstate abnormalities in early-course psychosis. Biol. Psychiatry: Cognitive Neurosci. Neuroimag. 5, 35–44.
Murray, M.M., Brunet, D., Michel, C.M., 2008. Topographic ERP analyses: a step-by-step tutorial review. Brain Topogr. 20, 249–264.
Nishida, K., Morishima, Y., Yoshimura, M., Isotani, T., Irisawa, S., Jann, K., Dierks, T., Strik, W., Kinoshita, T., Koenig, T., 2013. EEG microstates associated with salience and frontoparietal networks in frontotemporal dementia, schizophrenia and Alzheimer's disease. Clin. Neurophysiol. 124, 1106–1114.
Oostenveld, R., Praamstra, P., 2001. The five percent electrode system for high-resolution EEG and ERP measurements. Clin. Neurophysiol. 112, 713–719.
Pascual-Marqui, R.D., Michel, C.M., Lehmann, D., 1995. Segmentation of brain electrical activity into microstates: model estimation and validation. IEEE Trans. Biomed. Eng. 42, 658–665.
Pavlov, I., 2013a. "7z Format". http://www.7-zip.org/7z.html.
Pavlov, I., 2013b. LZMA SDK. http://7-zip.org/sdk.html.
Schlegel, F., Lehmann, D., Faber, P.L., Milz, P., Gianotti, L.R., 2012. EEG microstates during resting represent personality differences. Brain Topogr. 25, 20–26.
Seeber, M., Michel, C.M., 2021. Synchronous brain dynamics establish brief states of communality in distant neuronal populations. eNeuroeNeuro 8.
Seitzman, B.A., Abell, M., Bartley, S.C., Erickson, M.A., Bolbecker, A.R., Hetrick, W.P., 2017. Cognitive manipulation of brain electric microstates. NeuroimageNeuroimage 146, 533–543.
Tognoli, E., Kelso, J.A., 2014. The metastable brain. NeuronNeuron 81, 35–48.
Tomescu, M., Rihs, T., Rochas, V., Hardmeier, M., Britz, J., Allali, G., Fuhr, P., Eliez, S., Michel, C., 2018. From swing to cane: sex differences of EEG resting-state temporal patterns during maturation and aging. Dev. Cogn. Neurosci. 31, 58–66.
Tomescu, M.I., Papasteri, C.C., Sofonea, A., Boldasu, R., Kebets, V., Pistol, C.A., Poalelungi, C., Benescu, V., Podina, I.R., Nedelcea, C.I., 2022. Spontaneous thought and microstate activity modulation by social imitation. NeuroimageNeuroimage 249, 118878.
Tomescu, M.I., Rihs, T.A., Roinishvili, M., Karahanoglu, F.I., Schneider, M., Menghetti, S., Van De Ville, D., Brand, A., Chkonia, E., Eliez, S., 2015. Schizophrenia patients and 22q11. 2 deletion syndrome adolescents at risk express the same deviant patterns of resting state EEG microstates: a candidate endophenotype of schizophrenia. Schizophrenia Res.: Cognition 2, 159–165.
Van de Ville, D., Britz, J., Michel, C.M., 2010. EEG microstate sequences in healthy humans at rest reveal scale-free dynamics. Proc. Natl. Acad. Sci. 107, 18179–18184.
von Wegner, F., Tagliazucchi, E., Laufs, H., 2017. Information-theoretical analysis of resting state EEG microstate sequences-non-Markovianity, non-stationarity and periodicities. NeuroimageNeuroimage 158, 99–111.
Wackermann, J., Lehmann, D., Michel, C., Strik, W., 1993. Adaptive segmentation of spontaneous EEG map series into spatially defined microstates. Int. J. Psychophysiol. 14, 269–283.
Wick, M.R., Slagle, J.R., 1989. An explanation facility for today's expert systems. IEEE Intell. Syst. 4, 26–36.