Bistability breaks-off deterministic responses to intracortical stimulation during non-REM sleep

Pigorini, Andrea
2015-05-15


http://hdl.handle.net/10138/203730
https://doi.org/10.1016/j.neuroimage.2015.02.056

Downloaded from Helda, University of Helsinki institutional repository.
This is an electronic reprint of the original article.
This reprint may differ from the original in pagination and typographic detail.
Please cite the original version.
Bistability breaks-off deterministic responses to intracortical stimulation during non-REM sleep

Andrea Pigorini, Simone Sarasso, Paola Proserpi, Caroline Szymanski, Gabriele Arnulfo, Silvia Casarotto, Matteo Fecchio, Mario Rosanova, Maurizio Mariotti, Giorgio Lo Russo, J. Matias Palva, Lino Nobili, Marcello Massimini

ARTICLE INFO

Article history:
Accepted 24 February 2015
Available online 4 March 2015

Keywords:
Causality
Consciousness
Down-state
Intracranial
Phase-locking

ABSTRACT

During non-rapid eye movement (NREM) sleep (stage N3), when consciousness fades, cortico-cortical interactions are impaired while neurons are still active and reactive. Why is this? We compared cortico-cortical evoked-potentials recorded during wakefulness and NREM by means of time–frequency analysis and phase-locking measures in 8 epileptic patients undergoing intra-cerebral stimulations/recordings for clinical evaluation. We observed that, while during wakefulness electrical stimulation triggers a chain of deterministic phase-locked activations in its cortical targets, during NREM the same input induces a slow wave associated with an OFF-period (suppression of power > 20 Hz), possibly reflecting a neuronal down-state. Crucially, after the OFF-period, cortical activity resumes to wakefulness-like levels, but the deterministic effects of the initial input are lost, as indicated by a sharp drop of phase-locked activity. These findings suggest that the intrinsic tendency of cortical neurons to fall into a down-state after a transient activation (i.e. bistability) prevents the emergence of stable patterns of causal interactions among cortical areas during NREM. Besides sleep, the same basic neurophysiological dynamics may play a role in pathological conditions in which thalamo-cortical information integration and consciousness are impaired in spite of preserved neuronal activity.

© 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction

Once awakened from NREM stage N3 early in the night people often deny that they were experiencing anything at all (Stickgold et al., 2001). Experimentally, this reduction of consciousness upon falling into NREM is associated with a significant impairment of the ability of distributed groups of cortical neurons to sustain reciprocal interactions. This alteration has been detected by analyzing functional connectivity within cortical networks in resting condition (Boly et al., 2012; Spoormaker et al., 2011; Tagliazucchi et al., 2013) and becomes obvious when one applies perturbations directly to the cerebral cortex. Hence, measurements performed with transcranial magnetic stimulation (TMS) and electroencephalography (EEG) have shown that a single magnetic pulse triggers a complex chain of causal interactions that propagate through a distributed network of cortical areas during wakefulness, but a simple response that remains either local (Massimini et al., 2005) or spreads like an oil-spot (Massimini et al., 2007) during NREM. This altered response to TMS has been subsequently observed in other conditions in which consciousness is lost, such as general anesthesia and the vegetative state (Casali et al., 2013; Rosanova et al., 2012). Also in these cases, a differentiated pattern of deterministic interactions is replaced by the occurrence of a stereotypical slow wave. Yet, the neurophysiological mechanisms underlying this reduction of the brain's capacity to sustain complex patterns of cortical activation remain unclear.

The most obvious feature of spontaneous brain activity during NREM is the appearance of EEG slow waves associated with brief periods of hyperpolarization and neuronal silence (down-states) (Steriade et al., 1993). In principle, the spontaneous occurrence of cortical down-states may in itself be enough to prevent reliable information transmission among cortical areas. However, while this mechanism may be crucial in the case of general anesthesia, during which cortical activity...
is frequently interrupted by long hyperpolarized down-states (Lewis et al., 2012), it may not be sufficient in other conditions such as sleep. In fact, intracellular recordings in cats show that even during deep NREM, cortical neurons can spend most of their time in a depolarized, wakefulness-like up-state, which is only occasionally interrupted by hyperpolarized, silent down-states (Chauvette et al., 2011). Local field potentials (LFP) and multiunit recordings in humans have also demonstrated that most slow waves and the associated down-states are local (Hangya et al., 2011; Nir et al., 2011), and that some brain areas may be active for long stretches while others are asleep (Nobili et al., 2011). Finally, the active periods observed during NREM share several fundamental features with ongoing activity during wakefulness, e.g. mean firing rate (Destexhe et al., 1999; Hobson and McCarley, 1971; Steriade et al., 2001), spectral profile (Csercsa et al., 2010) and synchronization level (Destexhe et al., 1999; Steriade and Amzica, 1996) so much so that the depolarized states of the sleep slow oscillation have been referred to as “fragments of wakefulness” (Destexhe et al., 1999, 2007). Thus, it is not obvious that the spontaneous occurrence of silent periods may substantially impair the brain's capacity for internal communication.

Besides changes in spontaneous activity, cortical neurons undergo a more profound modification upon falling asleep as they become bistable upon changes in their intrinsic properties (Compete et al., 2003; Sanchez-Vives and McCormick, 2000; Timofeev et al., 2001). Due to underlying bistability, cortical networks have the tendency to fall into a silent down-state in response to transient increases in activity (Compete et al., 2003; Sanchez-Vives and McCormick, 2000). An intriguing possibility is that this tendency may specifically impair the emergence of stable patterns of causal interactions among cortical areas. Specifically, we hypothesize that (i) during NREM a group of neurons that receives a cortical input rapidly plunges into a down-state and that (ii) this period of silence breaks-off the causal effects of the initial input.

Intracranial single-pulse electrical stimulation (SPES) and simultaneous stereotactic EEG (SEEG) recordings offer a unique opportunity to test this hypothesis. First, intracranial recordings allow a reliable, although indirect, detection of cortical down-states as a significant suppression of high frequency power above 20 Hz in the LFP (Cash et al., 2009; Valderrama et al., 2012). Second, intracranial perturbations with SPES permit to assess cortico-cortical and/or cortico-subcortico-cortical interactions from a causal perspective by calculating the phase locking factor (PLF) (Sinkkonen et al., 1995) (Fig. S1). We thus analyzed by means of time–frequency decomposition and phase-locking analysis the SEEG responses to SPES in 8 patients undergoing neurosurgical evaluation for intractable epilepsy.

Materials and methods

Patients and data acquisition

Data included in the present study derived from a dataset collected during the pre-surgical exploration of eight neurosurgical patients with a history of drug-resistant, focal epilepsy (Table S1). All subjects were candidates for surgical removal of the epileptic focus. During the pre-surgical evaluation all patients underwent individual investigation with simultaneous single pulse electrical stimulation (SPES) and recordings performed by stereotactically implanted depth multi-lead electrodes (Stereo-EEG, SEEG) for the precise localization of the epileptogenic zone and connected areas (Cossu et al., 2005). The investigated hemisphere, the duration of implantation and the location and number of stimulation sites were determined based on the non-invasive clinical assessment.

SEEG activity was recorded from platinum–iridium semiflexible multi-contact intracerebral electrodes, with a diameter of 0.8 mm, a contact length of 1.5 mm, an inter-contact distance of 2 mm and a maximum of 18 contacts per electrode (Dixi Medical, Besancon France—Figs. 1A–B). The individual placement of the electrodes was ascertained by post-implantation tomographic imaging (CT) scans (Fig. 1B). In addition scalp EEG activity was recorded from two platinum needle electrodes placed during surgery on the scalp at standard “10–20” positions Fz and Cz. Electro-ocular activity was recorded from the outer canthi of both eyes, and submental electromyographic activity was also recorded. Both EEG and SEEG signals were recorded using a 192-channel recording system (NIHOKHODEN NEUROFAX-110) with a sampling rate of 1000 Hz. Data were recorded and exported in EEG NIH-Kohden format. Recordings were referenced to a contact located entirely in the white matter.

SPES is a clinical procedure increasingly employed for the identification of abnormal cortical excitability in patients with epilepsy (David et al., 2010; Valentín et al., 2002). In the present dataset, SPES was performed five days after electrode implantation both during wakefulness and during NREM (Silber et al., 2007) given the possible nocturnal nature of the seizures. Specifically, a 5 mA current was applied through one pair of adjacent contacts, while SEEG activity was simultaneously recorded from all other bipolar contacts. A single stimulation session consisted of 30 consecutive single pulses delivered at varying inter-stimulus intervals (1 to 5 s). The stimulation, recording and data treatment procedures were approved by the local Ethical Committee (protocol number: ID 939, Niguarda Hospital, Milan, Italy). All patients provided written informed consent.

Selection of recording contacts and stimulation session

In each subject the dataset derived from the pre-surgical exploration included several sessions (up to 20), in which different contacts were stimulated, and multiple contacts (up to 189) were recorded. For the present analysis, selection of recording sites and stimulation sessions was based on the following criteria: we excluded from the analysis those contacts that (i) were located in the epileptogenic zone (as confirmed by post-surgical assessment), (ii) were located over regions of documented alterations of the cortical tissue (e.g. Taylor Dysplasia) as measured by the radiographic assessment, or (iii) exhibited spontaneous or evoked (Valentin et al., 2002) epileptiform SEEG activity during wakefulness or NREM (visual inspection performed by L.N. and P.P.). Also contacts located in white matter, as assessed by means of MRI, were excluded (by two of the authors: L.N. and C.S.) from further analysis. Anatomical locations of bipolar derivations were confirmed by means of customized MATLAB scripts, using Destrieux ATLAS (Fig S1A).

With the given contact selection, we analyzed the stimulation sessions that: (i) were delivered through a contact far from the epileptogenic zone (as confirmed by post-surgical assessment); (ii) were delivered through a contact that did not show spontaneous interictal activity (by visual inspection); (iii) did not evoke epileptic responses (Valentin et al., 2002) either in wakefulness or NREM; (iv) did not elicit muscle twitches, sensations or evident motor/cognitive effects (e.g. language comprehension/production or complex motor sequences) during presurgical evaluation, including single pulse and repetitive (50 Hz) stimulation (Cossu et al., 2005); (v) occurred during a period of N3 sleep; and (vi) did not disrupt sleep depth as assessed by comparing the power spectra of scalp EEG spontaneous activity before and after the stimulation train (Fig S1B). If in a given subject more than one stimulation session fulfilled these criteria, we selected the one that triggered significant cortico-cortical evoked potentials (CCEPs, Matsumoto et al., 2004) in the largest number of contacts as in Keller et al. (2011).

Data preprocessing

Data recorded during both wakefulness and NREM were imported from EEG NIH Kohden format into Matlab and converted using a customized Matlab-based script. Data were subjected to linear detrend and bandpass filtering (0.5 – 300 Hz), using a third order Butterworth filter. Bipolar montages were calculated by subtracting the signals from adjacent contacts of the same depth-electrode to minimize common electrical noise and to maximize spatial resolution.
(Cash et al., 2009; Gaillard et al., 2009) (see Fig. 1A). Single trials were obtained by using a digital trigger simultaneous to each SPES delivery. Finally, stimulation artifact was reduced by applying a Tukey-windowed median filtering, as in Chang et al. (2012), between $-5$ and 5 ms (Fig. S1C).

Methodological rationale and data analysis

First we assessed the number of contacts showing significant response to SPES by applying the same methodology as in Keller et al. (2011). Specifically, at the level of individual contacts, we tested different thresholds (from 2 to 14 standard deviations of the mean baseline activity) calculated from the rectified, mean amplitude of the CCEPs. Then, in order to assess quantitatively the differences in the dynamics triggered by SPES in wakefulness and NREM, we quantified (1) the amplitude of the low frequency components ($<4$ Hz), (2) the suppression of high frequency power ($>20$ Hz) and (3) the PLF (Fig. S1). Indeed, these three measures indicate respectively the possible presence of SPES-evoked slow waves, the correspondent occurrence of a cortical down-state (Cash et al., 2009; Csercsa et al., 2010; Valderrama et al., 2012) induced by SPES and the causal effects of SPES at the level of individual contacts. In the following, as in previous works (Nir et al., 2011; Vyazovskiy et al., 2009), we use the terms “OFF” periods instead of “down” (or “hyperpolarized”) states (Steriade et al., 2001) because silent periods were defined on the basis of extracellular activity (suppression of high frequency in the LFP) rather than based on a direct measure of membrane potential.

Amplitude of low frequency component

To assess and quantify the presence of SPES-evoked slow waves we calculated, both for wakefulness and NREM, the amplitude of low frequency components ($<4$ Hz) (Figs. S1 and 2) for each contact as the average of absolute value of filtered (4 Hz low-pass, third order Chebyshev filtering) single trials. We rectified the slow components because slow waves could be either positive or negative, depending on contacts locations (Fig. S2). Then we set to zero all the non-significant time points using a trial-based bootstrap statistical analysis ($\alpha < 0.01$, number of permutations $= 1000$) with respect to baseline (from $-300$ ms to $-50$ ms).

High-frequency suppression

To assess and quantify the presence of SPES-evoked cortical OFF-periods we calculated the amount of significant (bootstrap method; $\alpha < 0.05$, number of permutations $= 1000$) high-frequency ($>20$ Hz) suppression, an established extracellular marker of the neuronal down-state that characterizes sleep bistability (Cash et al., 2009; Csercsa et al., 2010; Mukovski et al., 2007; Nir et al., 2011; Valderrama et al., 2012; Vyazovskiy et al., 2009). Hence, as in Cash et al. (2009), we calculated for each contact the mean value of the ERSP above 20 Hz (Figs. S1 and 2) over a peri-stimulus time-window ranging between $-300$ and $+700$ ms.

Phase locking factor

To quantify and compare the duration of the deterministic effects of SPES during both wakefulness and NREM for each contact, the instantaneous PLF was calculated as in Palva et al. (2005). PLF is an adimensional
A. Pigorini et al. / NeuroImage 112 (2015) 105–113

(range 0–1) index defined as the absolute value of the average of the Hilbert Transform of all single trials; in the case of evoked potentials (here CCEPs), it reflects the ability of an external stimulus (here SPES) to affect the phase of ongoing oscillations across trials. Statistical differences from baseline (from −300 ms to −50 ms) were assessed (for each contact) by assuming a Rayleigh distribution of the values of the baseline. Then, given a statistical threshold set at $\alpha = 0.05$, those PLF values below threshold were set to zero.

**Results**

In all patients (Table S1) SPES was delivered both during wakefulness and NREM through one pair of adjacent (2 mm apart) contacts pertaining to the same depth-electrode, while SEEG recordings were
obtained from all other bipolar contacts (Figs. 1A–B; see 2. Materials and methods and Supplementary material for a detailed description of the number and location of the stimulating and recording contacts). In each individual, depending on clinical needs, a number of cortical sites were stimulated with 30 pulses at frequencies between 0.2 and 1 Hz during wakefulness preceding lights off as well as during NREM (scored according to Silber et al., 2007). Off-line sleep scoring using one scalp EEG derivation, together with one bipolar electrooculographic (EOG) and one electromyographic (EMG) derivation confirmed that all considered sessions were recorded during NREM N3. The stability of stage N3 throughout each stimulation session was further assessed by comparing the power spectra of the scalp EEG recorded immediately (~40-s epochs) before and after the SPES train (Fig. S1B).

As illustrated in Fig. 1C, cortico-cortical CCEPs were highly reproducible from trial to trial and were characterized by a high signal-to-noise ratio in intracranial recordings. As a first step, we asked whether the ability of SPES to trigger significant activations across recording contacts differed in NREM compared to wakefulness. To do so, we detected at each contact the presence of significant CCEPs employing the same criteria described by Keller et al. (2011); for each individual contact, we tested different thresholds applied to the rectified amplitude of the CCEPs calculated on the mean amplitude signal. Across patients, the percentage of significantly active contacts in both wakefulness and NREM decreased monotonically with increasing thresholds (from ~100% at 2SD to ~50% at 14SD) and tended to be lower during NREM (Fig. 1D). However, no significant difference was observed between the two conditions at any applied threshold. Despite this similarity in the number of contacts affected by the initial stimulation, the CCEPs recorded during wakefulness and NREM were substantially different in terms of their wave shape.

SPES evokes a slow wave-like response in NREM but not in wakefulness

Generally, during wakefulness SPES evoked a composite response made of recurrent waves of activity that persisted until ~500 ms (Fig. 1E). Conversely, during NREM, CCEPs consisted of a simpler and slower wave, composed of three consecutive events, which we will henceforth call components 1, 2 and 3 (Figs. 1E and 2). The polarity of these components could be inverted depending on the location of the recording contacts (Fig. S2). In all cases, component 1 was a sharp peak (between 10 and 50 ms), component 2 was a prominent rebound of opposite polarity (peaking ~200 ms), and component 3 was an ensuing, smoother deflection in the same direction as component 1. Quantitatively, cortical responses to SPES during NREM were characterized by a prevalent low frequency (0.5–4 Hz) oscillation that was invariably reduced during wakefulness (Fig. 2A). Overall, both in terms of period and peak amplitude, the CCEPs recorded during NREM closely resembled spontaneously occurring sleep slow waves (Fig. S2).

SPES induces suppression of high-frequency power in NREM but not in wakefulness

To further explore the relationships between the CCEPs obtained during NREM and sleep slow waves, we capitalized on previous animal and human intracranial recordings. These studies show that the silent hyperpolarized state that characterizes the cortical OFF-period of spontaneously occurring sleep slow oscillations is marked by a suppression of high-frequency power (~20 Hz) in the LFP (Amzica and Steriade, 2002; Cash et al., 2009; Csercsa et al., 2010; Mukovskii et al., 2007; Nir et al., 2011; Steriade et al., 1993) (see also Fig. S2). Hence, we performed time–frequency decomposition of CCEPs (Cash et al., 2009) to compare the power modulation (as assessed by event-related spectral perturbation—ERSP) during wakefulness and NREM. Results obtained in one patient at three representative cortical targets during wakefulness and NREM are shown Fig. 2A.

While the composite CCEPs recorded during wakefulness corresponded to increases in spectral power as compared to baseline, the slow wave-like response elicited in NREM was associated with an alternation of positive and negative significant power modulations (see Methods, bootstrap statistics, α < 0.05). Specifically, component 1 of the NREM response coincided with a transient broadband increase of spectral power. Component 2 was associated with a significant suppression of high-frequency oscillations (~20 Hz) (Figs. 2, S1, and S2), irrespective of stimulation frequency (Fig. S3) and of the level of background activity (Fig. S4). Finally, during component 3 spectral power recovered and rebounded to a level comparable to wakefulness in a broad range of frequencies (between 8 Hz and 100 Hz), including alpha, spindle and beta rhythms (Fig. 2A and S5B).

Phase-locking is short-lasting during NREM but sustained during wakefulness

Next, we tested our main hypothesis, i.e. that bistability break-off of the causal effect of cortical inputs during NREM. To quantify and compare the duration of the deterministic effects of SPES during both wakefulness and NREM, we employed PLF applied to the LFP activity in the ~8–100 Hz frequency band, the same range characterizing (i) the power modulation (as assessed by event-related spectral perturbation—ERSP) during wakefulness and NREM, (ii) the correlation between the maximum amplitude of the evoked slow wave-like response elicited in NREM and high-frequency power in the slow wave frequency band (0.5–10 Hz) (Figs. 2, S1, and S2), irrespective of stimulation frequency (Fig. S3) and of the level of background activity (Fig. S4). Finally, during component 3 spectral power recovered and rebounded to a level comparable to wakefulness in a broad range of frequencies (between 8 Hz and 100 Hz), including alpha, spindle and beta rhythms (Fig. 2A and S5B).

Amplitude, power and phase-locking modulations of cortical responses are reproducible across contacts and patients

Overall, the results obtained for low frequency amplitude, high-frequency power suppression, duration of PLF and their differences between wakefulness and NREM were reproducible across cortical contacts (Fig. 3B), significant across contacts at the single subject level (Fig. 3B, histograms—Wilcoxon ranksum test, p < 0.05 and Fig. S5) and consistent at the population level (Fig. 3C—Wilcoxon ranksum test, p < 0.05). Hence, cortical responses to SPES during NREM were characterized by a prominent low frequency (0.5–4 Hz) oscillation, by a significant suppression of high-frequency (20–100 Hz) activity and by an early (~200 ms) obliteration of phase-locked, deterministic effects. By contrast, during wakefulness, when the low frequency component was reduced and the suppression of high frequency was absent, the PLF remained significant until ~500 ms after SPES.

Amplitude, power and phase-locking modulations of cortical responses are related during NREM

We finally asked whether the three distinctive features of the NREM response (i.e. the presence of a slow wave-like response, high frequency suppression and shorter PLF duration) were related. To this aim, for each subject, we selected the contacts showing the largest (top 50%) power in the slow wave frequency band (0.5–4 Hz). We then computed (i) the correlation between the maximum amplitude of the evoked slow wave (max SWA) and the maximum level of suppression of high-frequency power with respect to baseline (max SHFp) as well as (ii)
the correlation between the timing of the maximum high frequency suppression (max \(SHF_t\)) and the latency at which PLF dropped below significance level (max \(PLF_t\)). This analysis detected significant correlations in each single subject showing (i) that larger evoked slow waves corresponded to more pronounced suppressions of high frequencies and (ii) that earlier suppressions corresponded to an earlier dampening of PLF (Fig. 4).

**Discussion**

In the present work we compared CCEPs recorded during wakefulness and NREM by means of time–frequency analysis and PLF in 8 epileptic patients implanted with SEEG electrodes for clinical evaluation. We observed that during wakefulness SPES triggers a chain of sustained effects, as indicated by a phase-locked response that lasted for about half

![Figure 3](image1.png)

**Panel A.** For the same representative subject, color coded plots of amplitude of the slow wave (≤4 Hz) and high frequency power (≤20 Hz) and PLF (≤8 Hz) are calculated as a function of time for all contacts both for wake (W—left column) and NREM (NREM—right column). For each plot the responses to SPES are ranked based on the amount of low frequency power during NREM. Panel B. On the left, the same three measures as in Panel A averaged across contacts for W (in red) and NREM (in blue). On the right, the top and the middle histograms indicate the average values for low frequency amplitude and high frequency power respectively (black bars indicate standard error of the mean) calculated over the time interval between \(-50\) ms and \(+50\) ms around the maximum of the low frequency wave detected at each contact (Panel A—top left box). The bottom histogram shows the average duration of phase-locking (black bars indicate standard error of the mean) calculated from the time of the latest statistically significant non-zero PLF for each contact, in W and NREM. ** indicates significant differences (Wilcoxon ranksum test \(p<0.01\)). Panel C. The same three measures shown in Panels A and B are tested across all subjects (W in red, NREM in blue). Black lines indicate the grand average across subjects. Black vertical bars indicate the standard errors. Differences between W and NREM at the population level were tested using a Wilcoxon ranksum test \((p<0.01)\) and were significant for all three measures.

![Figure 4](image2.png)

**Panel A.** For each subject the correlation between the maximum amplitude (converted in dB) of the evoked slow wave (max \(SW_a\)) and the maximum level of high-frequency power suppression with respect to baseline (max \(SHF_p\)) is shown. Below, the coefficient of determination \(R^2\) and the significance level \(p\) of the correlation are indicated for each subject. Panel B. For each subject the correlation between the timing of the maximum high frequency suppression (max \(SHF_t\)) and the latency at which PLF fell below the threshold for significance (max \(PLF_t\)) is shown. Below, the coefficient of determination \(R^2\) and the significance level \(p\) of the correlation are indicated for each subject. For each subject, both correlations were calculated considering the contacts that showed the highest (top 50%) low frequency amplitude during NREM.
a second. During NREM the same initial activation induces a slow wave and a cortical OFF-period in its cortical targets after which the phase-locked response breaks-off, in spite of restored levels of cortical activity.

**Bistability, cortical down-states and break-off of phase-locked responses**

Intracortical stimulation/recordings can be used to study cortico-cortical interactions from a causal perspective. SPES delivered to the grey matter is thought to elicit direct depolarization of the initial segment of the axons, which travels through direct or indirect cortico-cortical pathways generating CCEPs in adjacent and remote areas (Keller et al., 2011; Matsumoto et al., 2004). Accordingly, in the present work, we observed evoked responses to SPES in targets both near and far from the stimulating contact, in ipsi-lateral and contra-lateral areas (see Fig. 2A). Significant responses to SPES could be elicited in a large number of contacts during both wakefulness and NREM (Fig. 1D). This finding is in agreement with previous animal studies (Richardson and Fetz, 2012) and with the hypothesis that the breakdown of complex interactions observed during NREM by means of TMS-EEG is not due to the interruption of structural cortico-cortical and/or cortico-subcortico-cortical connections but rather to changes in the dynamics of neuronal responsiveness (Massimini et al., 2012). Indeed, we observed that the initial activation triggered by SPES in its cortical targets was followed by a composite set of waves during wakefulness, but by a stereotypical slow wave that was associated with an extracellular marker of a neuronal down-state during NREM. Interestingly, slow waves and cortical OFF-periods could be triggered even on the background of a low-amplitude, wakefulness-like activated LFP (Fig. 54).

This observation is consistent with in vivo and in vitro studies showing that, due to bistability, networks that can display wakefulness-like activity tend to fall into a silent down-state upon transient increases in activation (Compte et al., 2003; Sanchez-Vives and McCormick, 2000; Timofeev et al., 2001). At the neuronal level, the bistability of NREM sleep is thought to be primarily due to the dynamics of activity-dependent potassium \( (K^+) \) currents, which become prominent when the neuromodulating milieu changes upon falling asleep (Compte et al., 2003; Sanchez-Vives and McCormick, 2000; Timofeev et al., 2001). In this condition, the stronger is the initial activation, the more \( K^+ \)-currents will tend to drive neurons into a hyperpolarized, silent state (Compte et al., 2003). Accordingly, in the present study, the degree of high frequency suppression was significantly correlated with the amplitude of the SPES-evoked potentials (Fig. 4A), indicating that larger responses were associated with stronger OFF-periods (i.e. deeper down-states).

We also found that the timing of the occurrence of the OFF-period was correlated with the timing of the drop of PLF (Fig. 4B) pointing to a specific role of the down-state in the early disruption of causal interactions. Importantly, PLF values dropped below significance level after the OFF-period despite power recovering to levels comparable to wakefulness in the 8 to 100 Hz frequency band (Fig. 55B). Different neuronal mechanisms may account for this phenomenon. Indeed, in vivo, in vitro and in computo studies seem to suggest that the resumption of cortical activity after the silent down-state of the slow oscillation is a stochastic process (Chauvette et al., 2011; Compte et al., 2003; Sanchez-Vives and McCormick, 2000) possibly due to spontaneous neurotransmitter release (Timofeev et al., 2000), intrinsic properties leading to spontaneous firing of layer V neurons (Sanchez-Vives and McCormick, 2000), or selective synchronization of small neuronal ensembles (Cossart et al., 2003; Lueckaz et al., 2007).

An additional finding that is worth discussing is that the PLF measured during wakefulness remained significant up to ~500 ms even in the contacts adjacent to the stimulation site, which, unlike distant targets, displayed a large slow wave with a concurrent significant suppression of high frequency activity, possibly due to the local paraphysiological effects of intracranial electrical stimulation (Borchers et al., 2012). A plausible explanation for the persistence of deterministic effects induced by SPES following this local OFF-period is the feedback of phase-locked activity from the rest of the network during wakefulness. A recent study employing cortical microstimulation in the monkey visual cortex demonstrated that, while feed-forward interactions mainly occur in the gamma band, feed-backs are carried by alpha oscillations (van Kerkoerle et al., 2014). Interestingly, in the present study we found that phase-locking in the gamma band was short-lasting (~70 ms) and comparable between wakefulness and NREM, whereas phase-locking in the beta and alpha bands was sustained (~500 ms) only during wakefulness (Fig. S6). In future studies it will be important to elucidate the relative contribution of feed-forward versus feed-back network dynamics to the persistence of phase-locking during wakefulness as compared to NREM.

**Bistability and loss of consciousness**

Growing evidence suggest that the sleeping brain can still process sensory inputs at least to some extent (Bastuji et al., 2002; Kouider et al., 2014) and that it can actively support restorative functions (Tononi and Chiara, 2014) and memory consolidation (Destexhe et al., 2007; Inostroza and Born, 2013). Indeed, compared to general anesthesia, during which a saturation of slow waves (Purdon et al., 2013) and frequent down-states may lead to a complete fragmentation of neuronal networks (Lewis et al., 2012), the impairment of cortico-cortical and/or cortico-subcortico-cortical interactions indicated by the present measurements in NREM is more likely to be relative and graded. In practice, the dynamics revealed by SPES point to a general mechanism by which the brain’s potential to sustain large-scale, specific patterns of causal interactions may be impaired during NREM compared to wakefulness; due to bistability cortical circuits, upon receiving an input, tend to respond briefly, then hush and forget. We argue that, while this dynamics still allows a certain degree of deterministic interactions, it may specifically affect the level of consciousness. For example, stimulation (Libet, 1982) as well as recording (King et al., 2014) experiments in humans have shown that conscious perception requires the activity of cortical neurons to be stable in time for hundreds of milliseconds. The occurrence of a cortical OFF-period (i.e. down-state) observed here as early as 100 ms after SPES, may act on this time factor by curtailing the duration of stable neuronal responses. At the network level, bistability may interfere directly with the efficacy of recurrent processes among distributed cortical areas, another mechanism that is thought to be important for consciousness (Tononi and Edelman, 1998; Lamme et al., 1998). To the extent that reentry relies on the amplification of coherent activity across distributed set of neurons, the scrambling of phases operated by the down-states at each node may critically impair the emergence of this large-scale phenomenon (Lumer et al., 1997). Finally, the present data fit nicely with the postulate that consciousness depends on the brain's ability to integrate information which, in turn, relays on effective information (EI) among different group of neurons (Hoel et al., 2013; Oizumi et al., 2014). EI is a perturbation-based general measure of causal interactions that captures how effectively causes produce effects in the system and how selectively causes can be identified from effects. Thus, EI is maximal for systems that are deterministic (i.e. a given initial state produces a given effect) and not degenerate (i.e. different initial states produce different effects); EI decreases with degeneracy (i.e. different initial states produce the same effect) and/or indeterminism (i.e. when a given initial state produce different effects) (Hoel et al., 2013). Our results suggest that, when cortical neurons become bistable, any input invariably converges into a stereotypical down-state (degeneracy), after which the causal effects of the input are obliterated (loss of determinism).

In this perspective, cortical bistability seem to be in a key position to selectively impair the level of consciousness.

Practically, the present results provide a mechanistic account for the collapse of complex spatiotemporal interactions revealed by non-
invasive TMS/EEG experiments during NREM and other conditions (Casali et al., 2013). A particularly interesting possibility is that bistability may play a role in pathological states in which TMS invariably triggers a stereotypical EEG slow wave, such as the vegetative state/ unresponsive wakefulness syndrome (Casali et al., 2013; Rosanova et al., 2012). Indeed, brain lesions may indirectly induce bistability in intact cortical tissue in several ways such as by (i) impairing the function of brainstem activating systems, thus enhancing K+-conductances at the cortical level (Englot et al., 2010), by (ii) reducing the excitatory drive of thalamostriatal circuits on cortical neurons (Schiff, 2010), by (iii) altering the excitation/inhibition balance in favor of inhibition (Murase et al., 2004), and by (iv) severing subcortical white matter fibers (Timofeev et al., 2000). To the extent that bistability—a process that is in principle reversible—is involved in these conditions, it may represent a suitable target for novel therapeutic approaches in patients in whom consciousness is impaired despite preserved cortical activity.

**Limitations**

Our observations were derived from a population of epileptic patients whose clinical condition and ongoing treatment may affect the SEEG recordings. To minimize this confound, our results did not include any contact (i) located in the epileptic zone as verified by surgical resection, (ii) located over regions of documented structural brain damage, nor (iii) exhibiting interictal activity (see Section 2, Material and methods).

Another potential limitation is that our data did not include multiunit activity recordings, thus preventing a direct observation of neuronal silence. However, previous studies in animals (Mukovski et al., 2007; Vyazovskiy et al., 2009) and humans (Cash et al., 2009; Csicsra et al., 2010; Nir et al., 2011) have provided solid evidence that the typical down-state characterized by slow slow oscillations can be reliably detected based on spectral modulation in the high-frequency (20–100 Hz) range of the LFP signal.

The number of session and recording contacts included in this study was constrained by clinical needs and exclusion criteria. However, the recording contacts included in the present analysis showed a widespread distribution across the cerebral cortex and the stimulations were applied to different cortical areas including frontal, parietal, cingular and insular cortex. Thus, while we cannot rule out specific regional differences that may only become significant through a more intensive and systematic mapping of CCEPs during wakefulness and sleep, we may safely conclude that our results could be generalized to different cortical areas (see Fig. S1).

Finally, it is known that awakening from NREM sleep (especially stages N1 and N2) and from Rapid Eyes Movement sleep (REM) result in dream report (McNamara et al., 2010). Although bistability per se does not index directly the level of consciousness, it would be important to assess its degree in all these conditions. In the present work we focused on comparing the responses obtained during wakefulness and N3 early in the night because (1) our primary goal was to parallel previous experiments employing non-invasive stimulation and recordings with TMS-EEG, (2) because we observed that NREM sleep N2 was less stable and much more affected by spontaneous interictal activity and paroxysmal physiological pattern compared to N3 and (3) because, since we couldn’t wake up subjects to obtain a report during the clinical protocol, we capitalized on previous works showing that the most extreme reduction of conscious experience occurs during the first N3 episode of the night (McNamara et al., 2010; Stickgold et al., 2001). This specific focus represents a limitation of this study and future works should definitively try to assess differences among sleep stages with special regards on REM sleep.

**Conclusions**

The present results reveal a general neurophysiological mechanism—bistability and the associated down-states—by which reliable deterministic interactions within thalamo-cortical system may be impaired even in the presence of intact cortical connections and preserved levels of neuronal activity. In essence, cortical circuits, upon receiving an input, tend to adapt briefly, then hush and forget. Future studies should investigate whether the same mechanism may account for the collapse of thalamo-cortical complexity detected by TMS/EEG in pathological conditions such as the vegetative state/unresponsive wakefulness syndrome.

**Acknowledgments**

We thank Adenauer Casali, Giulio Tononi, Chiara Cirelli, Corey J. Keller, Giandomenico Iannetti and Adam Barrett for insightful discussion and comments on the manuscript. This study was funded by EU grant FP7-ICT-2011-9, n. 600806 “Corticonic”, James S. McDonnell Foundation Scholar Award 2013 #220020369, and PRIN 2010 to M.M #2010SH7H3F_004.

**Appendix A. Supplementary data**

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.neuroimage.2015.02.056.

**References**


