Topography of functional connectivity in human multichannel EEG during second language processing

Ernesto Pereda[1,*], Susanne Reiterer[2,3], Joydeep Bhattacharya[4,5]

[1]Department of Basic Physics, University of La Laguna, Tenerife, Spain.
eperdepa@ull.es
[2]Department of Neuroradiology, Section of Experimental MR of the CNS,
University of Tübingen, Germany
[3]Hertie Institute for Clinical Brain Research, General Neurology,
University of Tübingen, Germany
susanne.reiterer@med.uni-tuebingen.de
[4]Department of Psychology, Goldsmiths College, University of London,
London, United Kingdom
[5]Commission for Scientific Visualization, Austrian Academy of Sciences, Vienna, Austria
j.bhattacharya@gold.ac.uk
* Corresponding author.

Abstract. We analyze the topography of nonlinear interdependence in the EEG of two group German-native speakers, divided according to their English proficiency level (high or low), when listening to one text in German and one in English. Global functional connectivity was assessed in the full band EEGs using the nonlinear correlation integration entropy, an index of multivariate interdependence derived from the normalized cross-mutual information between every two electrodes within each region of interest (ROI): three interhemispheric (frontal, centro-temporal and parieto-occipital) and two intrahemispheric ones (left and right hemisphere). The results show clear topographic differences between the interhemispheric ROIs, but no differences between the intrahemispheric ROIs Furthermore, there were also differences in language processing that depend on the proficiency level. We discuss these results and their implications along with recent findings about phase synchronization in the gamma band during second language processing.

Keywords: EEG, second language processing, functional connectivity, joint entropy

1 Introduction

Non-linear multivariate time series analysis methods have been extensively and successfully used during the last decade to study brain dynamics from EEG and MEG records in different situations (see, e.g., [1] for a review). Indeed, the term *functional* *connectivity*[1] has been coined to refer to the existence of statistical dependencies between the signal recorded from distinct units (ranging from single neurons to whole brain areas) within a nervous system [2]. Initial works in this line of research concern themselves with the analysis of the statistical interdependence between two units using bivariate nonlinear indexes of, e.g., generalized or phase synchronization. However, with more and more experiments in which an increasing number of sites were simultaneously recorded, it became apparent the need for new, truly multivariate approaches that allow the characterization of the collective dynamics of more than two units [3-5]. We have recently used one of these approaches to characterize the global phase synchronization in the gamma band of the EEG during second language processing and its dependence on the proficiency level of the subjects [6]. In this work, we complement and extend this result by studying, using an index of nonlinear correlation based on the concept of mutual information, the topography of the functional connectivity during second language processing of full-band EEG.

2 Material an Methods

2.1 Groups of subjects

The two groups of subjects contrasted have been described elsewhere [6], thus we only describe them briefly here. 38 university students with comparable educational levels were divided in two groups of 19 subjects each according to their second language proficiency level (L2 = English). The ‘high proficiency group’ subjects (HP) were advanced university language students studying English language and linguistics for a master’s degree (last year, 5-6 years completed). Their level of English proficiency was ‘‘very good’’ (so-called “native speaker-like” performance) or “good” according to their performances at university: they all had high levels of linguistic training and knowledge at the time of experiment. Additionally, they had spent abroad in an English speaking country an average of 10 months. By contrast, participants in the ‘low-proficiency group’ (LP) were university students of various disciplines other than English language and linguistics. They displayed medium to low level of second language skills (corresponding to the three rating-system groups “medium”, “lower-level” and “lowest-level”), which were sufficient to let them pass their school leaving exams (“Matura”, an equivalent to “A levels”), but since then were not developed any further. They were able to lead basic level conversations in English, but their speech was non-fluent. The average amount of time LP participants spent abroad in an English speaking country was 5 weeks. With regard to the country where they had spent some time, the groups were homogeneous.

[1] The definition of functional connectivity given here, is the most commonly accepted nowadays, although a search in Google of the expression “functional connectivity” (into inverted commas) produces, as of May 1st, 2011, no less than 190.000 results, some of them with different definitions of this concept.

All subjects were right-handed (measured by the Edinburgh handedness inventory; Oldfield, 1971) female students with German as their native language. We rigidly controlled for the variable gender in order to avoid possible influences of gender onto the processing of language and its neural representations. Mean (SD) age was 24 years (2.3 years and 2.7 years respectively for two groups) for both groups. They were also matched for socio-cultural background and education: all participants had similar social (middle class), educational (university students), and cultural (living in Vienna) background.

We strictly controlled for the variable "age of onset" of L2 learning. The average (± s.d.) age of onset was 9 yrs (1 yr) and was matched between the two groups. Further controlled variables were: age, handedness, gender, mother tongue, socio-educational and cultural background and region of residence.

The study was in compliant with the Code of Ethics of the World Medical association (Declaration of Helsinki) and the experimental protocol was approved by the local ethics committee. All subjects gave their written informed consent for the study.

2.2 Stimulus

We used coherent spoken speech (radio news) as stimuli in a listening comprehension and discourse processing paradigm. In cooperation with the English department at the University of Vienna, the speech samples were matched for syntactic complexity, semantic contents & genre, discourse structure and gender of the speaker (all male speakers). Within the framework of a block design, six blocks of coherent speech (2.0 – 3.2 min each) with randomly inserted baseline blocks (acoustic noise, 2.0 min each) were presented in randomized order: three blocks in condition L2 English and three blocks in condition L1 German were auditorily presented in randomized order over earphones.

2.3 Data recording and pre-processing

We recorded multivariate EEG signals during L1 and L2 processing in a quiet, dimly-lit sound-proof experimental room. Participants were monitored through a video control system during the recording session in order to control for possible movements. Nineteen gold-disc electrodes were carefully attached to the scalp with adhesive electrode gel, positioned according to the international 10/20 System (Fig. 1); one additional frontal electrode was used as a ground, and two separate electrodes, at the right and left ear-lobe, were used as reference electrodes. The recordings were re-referenced against the algebraic mean of the two ear-lobe electrodes [7]. Eye movements were additionally controlled for by a piezo-electric device attached to the eyelid. Using a conventional Nihon-Kohden 21 channel recorder, the EEG was amplified, filtered (time constant 0.3 s.), displayed and recorded at a sampling rate of 128 Hz. Electrode's impedance was kept below 5 kΩ. A notch filter at 50 Hz was used for the elimination of power line contamination. Finally, we rejected those epochs containing samples of voltages higher than 70 μV (absolute value), plus additional epochs where 2% or more samples deviated more than 3 standard deviations from the mean value.

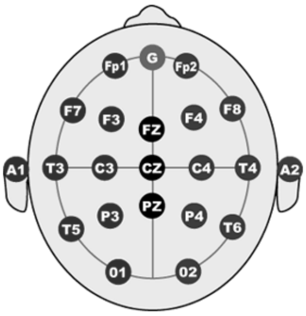

Fig. 1. Placement of the 19 electrodes recorded in the study (G represents the ground electrode; A1 and A2 are the linked earlobes used as reference)

2.4 Data analysis

2.4.1 Assessment of multivariate functional connectivity: the nonlinear correlation information entropy, I_R

In order to assess the functional connectivity in our multivariate data set, we made use of the so-called nonlinear correlation information entropy [5], whose calculation is outlined henceforth.

Given two discrete variables $X=[x_i]_{i=1,..,Ns}$ and $Y=[y_i]_{i=1,..,Ns}$, from which *Ns* samples have been obtained, we first sort, in ascending order, these samples and bin them into *b* ranks, with the first *Ns/b* samples of each variable placed in the first rank, the second *Ns/b* samples placed in the second rank, and so on. Then, the sample pairs $[(x_i,y_i)]_{i=1,.., Ns}$ are placed into a *b x b* rank grids by comparing each sample pair to the rank sequences of *X* and *Y*. The revised entropy of the variable *X* is defined as:

$$H^r(X) = -\sum_{i=1}^{b} \frac{n_i}{N_s} \log_b \frac{n_i}{N_s} \tag{1}$$

and the revised joint entropy of the two variables *X* and *Y:*

$$H^r(X,Y) = -\sum_{i=1}^{b} \sum_{i=1}^{b} \frac{n_{ij}}{N_s} \log_b \frac{n_{ij}}{N_s} \tag{2}$$

where n_{ij} is the number of samples in the *ij*th rank grid. The nonlinear correlation coefficient NCC(*X;Y*) is:

$$NCC(X;Y) = H^r(X) + H^r(Y) - H^r(X,Y) \tag{3}$$

here $H^r(Y)$ is defined in complete analogy with (1). Due to the binning scheme, n_i is invariant for both *X* and *Y* and equal to *Ns/b*. Thus, *NCC(X;Y)* reduces to:

$$NCC(X;Y) = 2 + \sum_{i=1}^{b^2} \frac{n_{ij}}{N_s} \log_b \frac{n_{ij}}{N_s} \quad (4)$$

If the sample sequences are exactly the same, the last right-hand side of the above equation equals -1 and thus, $NCC(X;Y)=1$, whereas if the two variables are completely uncorrelated, the sample pairs distribute equally into the *b x b* ranks, the sum equals to -2 and $NCC(X;Y)=0$.

In the case of $k>2$ variables (e.g., more than two EEG channels), we obtain a symmetric squared *k x k* matrix of nonlinear correlation coefficients:

$$R = \{NCC_{ij}\}_{i,j=1,..,k} \quad (5)$$

were NCC_{ij} is the nonlinear correlation coefficient between signals *i* and *j*, and $NCC_{ij}=NCC_{ji}$. Besides, NCC_{ij} =1 if $i=j$, and $0 \leq NCC_{ij} \leq 1$ when $i \neq j$. Thus, *R* is a Hermitian matrix, which is also positive semidefinite. The sum of its eigenvalues equals the trace, i.e.:

$$\sum_{n=1}^{k} \lambda_n = k \quad (6)$$

Recent studies on multivariate EEG analysis have taken advantage of the spectral properties of this kind of matrixes (see, e.g., [4, 8] for the equal time correlation matrix, and [3] for the bivariate phase synchronization matrix). The underlying idea is easy to understand if we analyze the two extreme cases of *k* completely correlated and *k* completely independent signals. In the first case, $NCC_{ij} = 1$, $\forall\, i,j = 1,..,k$, and $\lambda_1=k$, $\lambda_n=0$ $(n=2,..,k)$. Conversely, in the second one, $NCC_{ij} = 0$ whenever $i \neq j$, and $\lambda_n=1$ $(n=1,\ldots,k)$. In a (more realistic) intermediate case, a subgroup of the higher eigenvalues, which characterize dynamical clusters of functionally connected EEG channels, is greater than 1, whereas the rest are lower than 1. Additionally, the study of the corresponding eigenvectors makes it possible to define a participation index that assigns each electrode to a given cluster [3].

In the case of (5), this spectral property can be used to define a nonlinear index of multivariate correlation among $k>2$ signals. The *nonlinear joint entropy* of the *k* variables is derived from *R* as follows:

$$H_R = -\sum_{i=1}^{k} \frac{\lambda_i}{k} \log_k \frac{\lambda_i}{k} \quad (7)$$

From the properties of the eigenvalues, it follows that (7) is 0 if the *k* signals are completely correlated and 1 if they are completely independent

Thus, I_R, defined as:

$$I_R = 1 - H_R = 1 + \sum_{i=1}^{k} \frac{\lambda_i}{k} \log_k \frac{\lambda_i}{k} \qquad (8)$$

is an index of multivariate nonlinear correlation among k signals (termed the *nonlinear correlation information entropy)*, which equals 1 if they are completely correlated, and 0 if they are completely independent. In an intermediate case, one has $0<I_R<1$, with the index closer to 1 the more correlated are the signals. Note that, since (8) is obtained from the eigenvalues of R (whose elements are nonlinear correlation indexes), I_R is sensitive to both linear and nonlinear correlations among the k signals. This represents an advantage of I_R over similar indexes such as the one described in [4], which are only sensitive to linear correlations.

2.4.2 Regions of interests

We studied the topography of functional connectivity in both group of subjects and both conditions by defining three different, non-overlapping interhemispheric regions of interest (ROIs) (frontal region (FR), which includes electrodes Fp1, Fp2, F7,F3, Fz, F4 and F8; centro-temporal region (CT), which includes electrodes T3, C3, Cz, C4, T4, T5 and T6; parieto-occipital region (PO), which includes electrodes P3, Pz, P4, O1 and O2), and two intrahemispheric ROIs (left hemisphere (LH), including electrodes Fp1, F7, F3, C3, T3, C3, P3 and O1; right hemisphere (RH), including electrodes Fp2, F4, F8, C4, T4, P4, T6, and O2).

2.4.3 Practical aspects of I_R calculation

One practical issue that is necessary to deal with when estimating I_R from experimental data is that it is a *parametric* index, i.e., it depends on two parameters: the number of data samples, *Ns*, and the number of ranks, *b*. Typically, entropy estimations based on data binning may be strongly biased if either the total number of data of the average number of data in each bin are not long enough, which gives rise to high entropy values (see, e.g., [9], for a review of entropy estimation methods from data samples). Thus, it is necessary to determine a priori which are suitable values of both parameters to avoid (or at least, reduce as much as possible) such overestimation. Fig. 2 exemplifies, using the FR region of one subject, the values of I_R as a function of *b* and *Ns*. As can be seen for the figure, lower values of Ns and high values of b tend to produce higher values of I_R. According to this result, we took N_s=4000 (which correspond to 39.6 s) and *b=20.*

Thus, for every subject and ROI, we slid a moving window of size *Ns* along the whole record, and calculate I_R as the average of this index for the N_W windows[2]:

[2] The Matlab© script to calculate I_R is available upon request from the corresponding author.

$$I_R = \frac{1}{N_W}\sum_{i=1}^{N_W} I_{R_i} \tag{9}$$

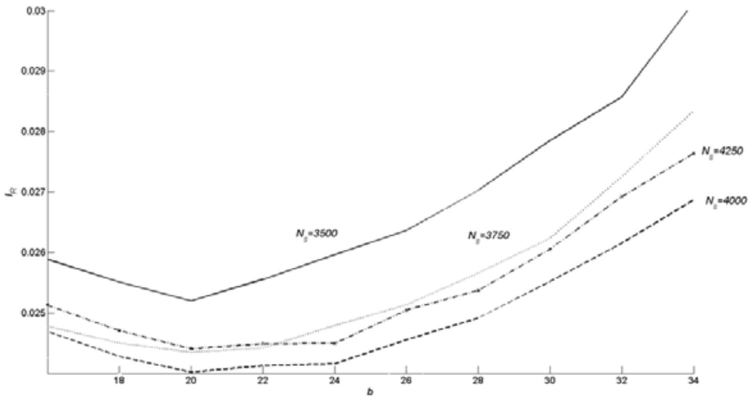

Fig. 2. I_R as a function of *Ns* and *b* for the FR region of one subject

2.5 Statistical comparisons

A multivariate analysis of the variance (MANOVA) test was used to check for the existence of between and within group differences. Thus, interhemispheric ROIs were analyzed with proficiency (H vs. L) as between group factor and Language (L1 vs. L2) and region (FR, CT and PO) as dependent factors. Likewise, intrahemispheric ROIs were analyzed by substituting in the above scheme the three intrahemispheric ROIs by the two interhemispheric ones. We used the conservative Bonferroni post-hoc test, when appropriate, to get further insight into the origin of these differences, which were considered significant for $p<0.05$.

As an additional precaution against false positives, we used a Levene's test to check the homogeneity of the variances of the different groups before applying the MANOVA test. All the statistical calculations were carried out using the data analysis software system STATISTICA[3] (StatSoft, Inc. (2008)) version 8.0.

3 Results

The Levene’s test was not significant for either the interhemispheric or the intrahemispheric ROIs analysis, which indicates that the variance is homogeneous in all cases.

[3] http://www.statsoft.com

Figure 3 presents the results corresponding to the interhemispheric ROIs, which can be summarized as follows: there are global within-group differences among ROIs ($p<0.001$), with a lower I_R for the CT region than for the other two regardless of the language and the proficiency, and a further increase of the index in the PO region for L2 (both groups) as well as L1 (HP group). Furthermore, I_R was lower for L2 as compared to L1 for the LP group in the FR region.

The results for the two intrahemispheric ROIs are shown in figure 4. In this case, there are neither between- nor within-group differences.

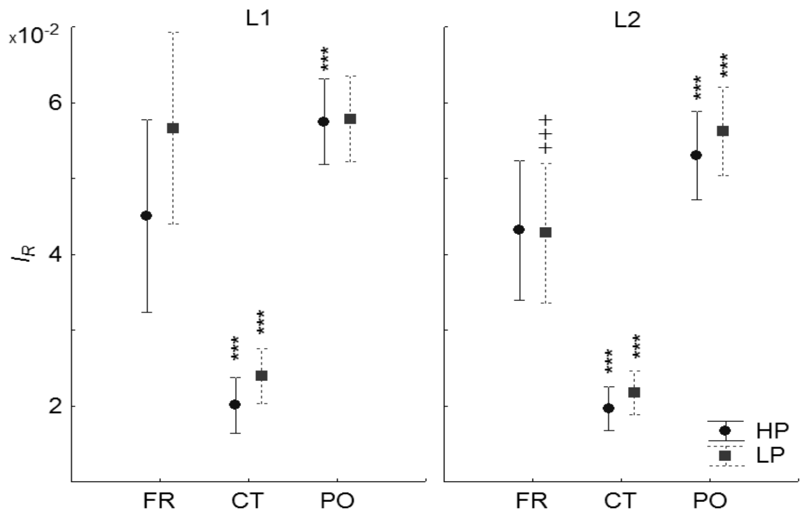

Fig. 3. Average I_R for the three interhemispheric ROIs (FR: frontal, CT: centrotemporal; PO: parieto-occipital) and both proficiency groups (HP: high proficiency, LP: Low proficiency) during L1 (left) and L2 processing (right). Vertical bars denote 0,95 confidence intervals. Asterisks stand for within-group regional differences (vs. CT). Crosses stand for L1 vs L2 differences. ***,+++:p<0.001 (Bonferroni post-hoc test).

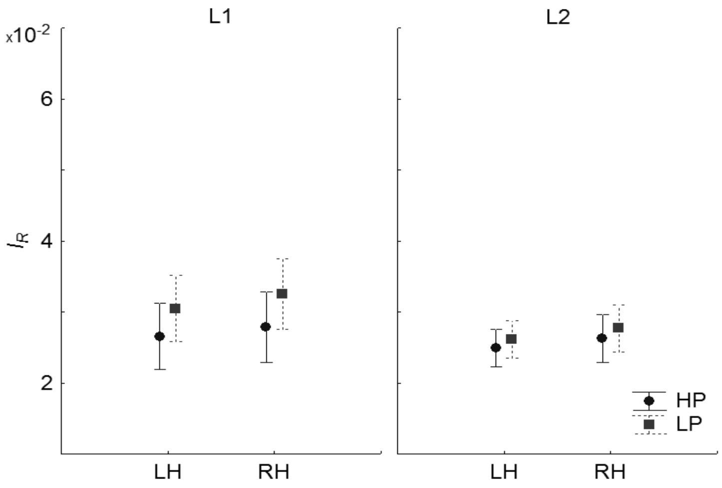

Fig. 4. Same as in Fig. 3 but for the intrahemispheric ROIs (LH: left hemisphere; RH: right hemisphere). We use the same upper and lower limits for the vertical axis as in Fig. 3 for comparability.

4 Discussion

We have shown in this work that functional EEG connectivity during language processing presents topographic interhemispheric (but not intrahemispheric) differences, with the FR and the PO regions showing greater collective cooperation than the CT region. Additionally, although we did not directly compare them, as it is apparent from fig. 3 and 4 the two former ROIs presented greater functional connectivity than the two intrahemispheric ROIs, indicating that cooperation within FR and PO regions is superior to that within the two hemispheres.

As for differences between L1 and L2 processing, they were found only in the FR region. Moreover, whereas the topographic differences are the same for the two groups, differences in language processing exist only for the LP group, where functional connectivity in the FR region decreases during L2 processing.

A straightforward conclusion of this latter result would be that proficiency L2 level correlates with the degree of frontal functional connectivity, because native-like L2 proficiency gives rise to a functional interhemispheric integration in this region that is essentially equal to that found during L1 processing. Yet in an earlier work we found that gamma band phase synchronization is greater in LP than in HP subjects during L2 processing [6], which we explained within the framework of the cortical efficiency hypothesis. According to it, persons who are good at a certain task, use a limited group of brain circuits or use their neuronal subroutines more efficiently, thus requiring fewer neuronal networks to accomplish a task, while poor performers (for whom problems are hard) use more circuits, which are inessential or inefficient for task performance and this is reflected in higher overall patterns of activity [10].

Taken together, past and present results on the relationship between functional EEG connectivity and L2 proficiency level suggest that high proficiency, native-like, processing of L2 is carried out with the same balance between functional segregation that is used during L1 processing. However, low-proficiency L2 level produces both an increase of frontal functional segregation and functional integration in the high frequency gamma band. This increase in functional connectivity in the gamma may be therefore a mechanism to compensate the reduced cooperation during L2 processing (as assessed by I_R) among the frontal areas of LP subjects in the full-band EEG. In contrast, HP subjects, who process L2 almost automatically L2, do it thanks to the proficient cooperation of their frontal areas. Thus, careful analysis of functional connectivity in full-band EEG is necessary to thoroughly characterize, on the one hand, the balance between integration and segregation of the brain areas that participate in L2 processing; and, on the other hand, the changes in functional connectivity that distinguishes LP from HP subjects.

5 References

1. Pereda, E., De La Cruz, D.M., De Vera, L., Gonzalez, J.J.: Comparing generalized and phase synchronization in cardiovascular and cardiorespiratory signals. IEEE Trans. Biomed. Eng. 52, 578-583 (2005)
2. Friston, K.J.: Functional and effective connectivity in neuroimaging: a synthesis. Hum Brain Mapp 2, 56-78 (1994)
3. Allefeld, C., Muler, M., Kurths, J.: Eigenvalue decomposition as a generalized synchronization cluster analysis. Int J Bifurcat Chaos 17, 3493-3497 (2007)
4. Müller, M., Baier, G., Galka, A., Stephani, U., Muhle, H.: Detection and characterization of changes of the correlation structure in multivariate time series. Phys Rev E 71, 046116 (2005)
5. Wang, Q., Shen, Y., Zhang, J.Q.: A nonlinear correlation measure for multivariable data set. Physica D: Nonlinear Phenomena 200, 287-295 (2005)
6. Reiterer, S., Pereda, E., Bhattacharya, J.: Measuring second language proficiency with EEG synchronization: how functional cortical networks and hemispheric involvement differ as a function of proficiency level in second language speakers. Second Language Research 25, 77-106 (2009)
7. Essl, M., Rappelsberger, P.: EEG coherence and reference signals: experimental results and mathematical explanations. Med Biol Eng Comput 36, 399-406 (1998)
8. Rummel, C., Baier, G., Muller, M.: Automated detection of time-dependent cross-correlation clusters in nonstationary time series. Europhys. Lett. 80, (2007)
9. Hlavackova-Schindler, K., Palus, M., Vejmelka, M., Bhattacharya, J.: Causality detection based on information-theoretic approaches in time series analysis. Phys. Rep. 441, 1-46 (2007)
10. Haier, R.J., Siegel, B.V., Jr., MacLachlan, A., Soderling, E., Lottenberg, S., Buchsbaum, M.S.: Regional glucose metabolic changes after learning a complex visuospatial/motor task: a positron emission tomographic study. Brain Res 570, 134-143 (1992)