Project

Impact des connaissances sur le locuteur dans la compréhension du langage non-littéral

Maude Champagne-Lavau(1), Madelyn Klein(1)

(1) LPL

Abstract

L’objectif de cette étude est d’explorer quand et comment des informations linguistiques et extralinguistiques concernant des connaissances sur le locuteur (i.e. le métier qu’il occupe) sont intégrées lors de la compréhension d’énoncés ironiques. Pour répondre à cet objectif, l’EEG est utilisée. En accord avec la littérature un composant positif tardif/P600 de plus grande amplitude est attendu pour les énoncés interprétés comme ironiques en comparaison à ceux interprétés comme littéraux en fonction du contexte précédant l’énoncé cible. Une modulation de cette positivité liée à l’ironie, en fonction du métier du locuteur, suggèrerait que des informations extralinguistiques portant sur le locuteur affectent les étapes de traitement tardives dans la compréhension de l’ironie (Regel et al., 2010). En plus de l’analyse classique des potentiels évoqués (PEs), il s’agit aussi d’étudier l’activité oscillatoire neuronale par une analyse en Temps-Fréquence (ATF) afin de distinguer des traitements potentiels non visibles dans l’analyse des PEs et de mieux caractériser les processus d’intégration mis en jeu dans la compréhension de l’ironie (Bastiaansen et al., 2012) (cf. Justification de cette analyse ci-dessous). L’analyse en Temps-Fréquence permet d’étudier la synchronisation et la désynchronisation de populations neuronales liées au couplage et au découplage d’ensembles de neurones fonctionnellement liés (Bastiaansen et al., 2012). Il est supposé qu’en réponse à l’arrivée d’un événement, l’activité des ensembles de neurones fonctionnellement liés se synchronise dans une bande de fréquence donnée. Ceci conduit à des changements fréquence-spécifique (par exemple, augmentation ou diminution de la puissance) de l'activité neuronale oscillatoire mesurables à la surface du scalp.

Publications

-

Fiche-résumé contribution CREx

Ironie

How the processing of non-literal language is impacted by one's knowledge of the speaker.

Description: Study of the pragmatics of irony processing

Investigators: Maude Champagne-Lavau, Madelyn Klein

Duration : 6 mois – ongoing

(

- Pre-processing of continuous electrophysiological data in preparation for time-frequency analysis.
- Time-Frequency decomposition.

Objectives: *The aim of this study is to understand how linguistic and extra-linguistic information (such as context and knowledge of the interlocutor) are integrated into the comprehension of non-literal language, in this case, ironic speech. In particular the researchers are interested in uncovering the time latency in relation to the critical stimulus at which this integration occurs and the neural substrates of this integration. In order to gain a deeper insight into the processes underlying this integration, the researchers wish to carry out time-frequency analysis.*

Preprocessing – The continuous EEG data, which had already been resampled to 512Hz and re-referenced to the average of the left and right mastoids were band-pass filtered between 0.3 and 120Hz using a windowed sinc filter. An upper filter limit of 120Hz was chosen as the researchers are interested in looking at activity in the high gamma band also (>50Hz) as well as the lower frequency bands: theta (4-7Hz), alpha (8-13Hz) and low gamma (30-45Hz).

Once filtered the data was epoched based on the 6 experimental conditions:

1. Sarcastic Profession + Ironic Context – **SI** (trigger code = 7)
2. NonSarcastic Profession + Ironic Context –**NSI** (trigger code = 9)
3. Sarcastic Professtion + Literal Context – **SL** (trigger code = 8).
4. Non-Sarcastic Profession + Literal Context –**NSL** (trigger code = 10).

5. First-name + Ironic Context – **NI** (trigger code = 11).
6. First-name + Literal Context – **NL** (trigger code = 12).

Total epoch duration was 3500ms; baseline of 1000ms and a post-stimulus interval of +2500ms. Such large epochs were defined so as to facilitate the analysis of the lower frequency bands, in particular the theta band (4-7Hz). However, baseline normalization was carried out using a pre-stimulus baseline of -250 to -50ms.

As the researchers were interested in analyzing activity in the gamma band, line-noise removal using a notch filter around 50Hz was not possible. The data, however, was characterized by very strong line-noise interference occurring on a wide range of electrodes so two techniques were applied to correct the very significant line-noise:

1. Adaptive Regressive Filtering Approach
2. Blind Source Separation: PCA-ICA

The **adaptive regressive filtering approach** (Mitra and Bokil, 2007) estimates the sinusoidal noise in the data in an adaptive manner and then removes these noise components. Briefly, a sliding window is applied to the data within which the instantaneous frequency spectrum is calculated using multi-tapers. By considering the sinusoidal line noise as a deterministic component that is mixed in white noise we can create a regression of the spectrum of the sinusoid component, as calculated with multi-tapers, onto the multi-taper spectrum of the original data at a given frequency. The regression coefficient, being complex, gives the phase and amplitude of the deterministic sinusoidal component. This allows us to reconstruct a time-domain version of the sinusoid and to subtract it from the data. Determining the exact line frequency can be difficult as it may not fall exactly on the specified 50Hz frequency component. To mitigate this, a Thompson F-statistic is applied to ascertain if the regression coefficient amplitude is significant (above zero) and small bandwidth above and below the defined sinusoid frequency is searched using the Thompson F-statistic to locate significant sinusoidal components. The method used here has been also adapted for use with both continuous and segmented data.

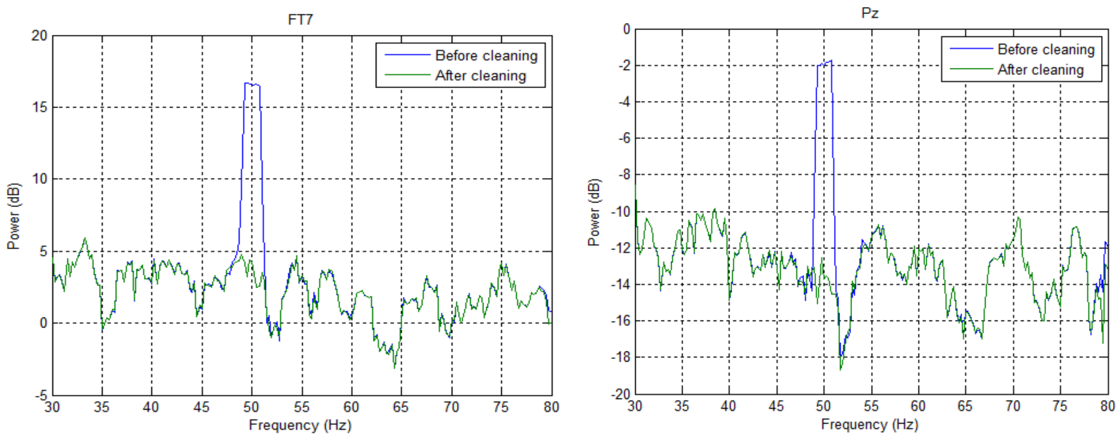

Figure 1: *Power spectrum of single-trial data for four different electrodes before (blue) and after (green) line noise removal using the adaptive regressive filtering approach.*

However, for certain subjects the line-noise proved too strong to be corrected using this approach (possibility of defective electrodes) and **PCA-ICA** was applied. In a first step PCA (Principle

Components Analysis) was applied to reduce data dimensionality and, as a consequence, the space and time complexities; this yielded 15 principle components. In a second step, ICA (Independent Component Analysis) was applied to separate the PCA data into 15 mutually independent components. Components corresponding to the undesired line noise were isolated by means of the components' frequency spectrum and time course. Figure 2b shows subject 13 after correction of the line noise component using ICA.

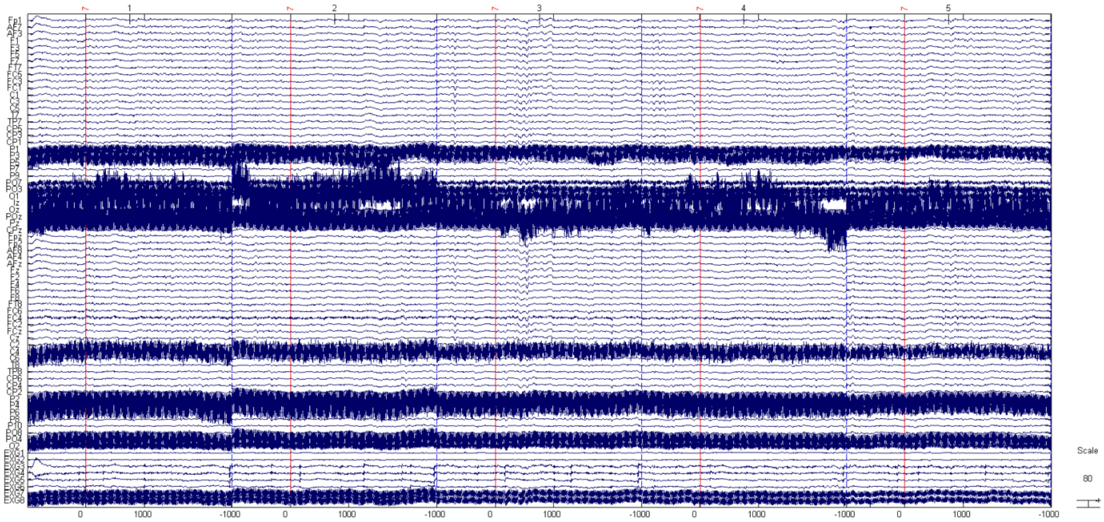

Figure 2a: *Very strong line noise (50Hz) interference of subject 13 dataset.*

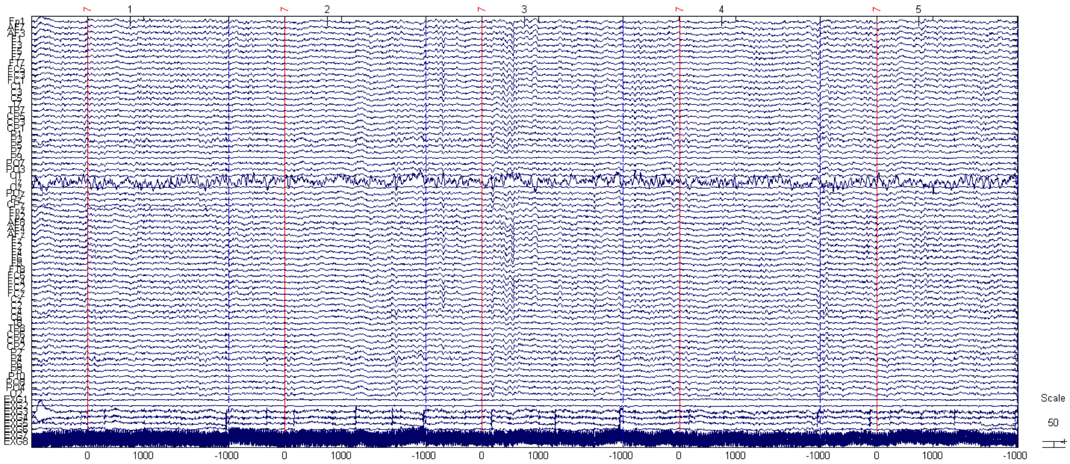

Figure 2b: *Subject 13 after ICA correction of line noise interference.*

In certain cases, even after application of PCA-ICA certain electrodes remained noisy (such as Iz in figure 2b) and were removed from the data.

Noisy epochs were also removed in a semi-automatic manner: visualization and based on the robust z-scores, analysis of the spectrum and total above-limit (>75mV) time per-epoch.

Analysis – Time frequency decomposition was carried out at the level individual epochs for each experimental condition and on all 20 subjects, which allows the investigation of both phase-locked and time-locked but non-phase locked data, otherwise known as the **induced time-frequency** data.

The time-frequency decomposition (TFD) was carried out over two wide frequency bands:

- 4Hz to 40Hz using time-domain convolution of Morlet wavelets
- 30Hz to 80Hz using multi-taper fft.

Baseline Normalisation of Time-Frequency Data

The interval -500ms to -100ms was used as baseline for the normalization of time-frequency data. The post-stimulus data was defined in terms of the relative change in relation to the baseline interval. The relative is calculated as follows:

$$X_{blc} = \frac{X_{all} - \overline{baseline}}{\overline{baseline}},$$

Where $\boldsymbol{X_{blc}}$ is the baseline corrected data, $\boldsymbol{X_{all}}$ is all epoch data and $\overline{\boldsymbol{baseline}}$ is the mean of the baseline.

Spatial Cluster-based Permutation Test

The mean time-frequency data of all subjects were entered into a two-tailed cluster-based permutation test taking into account all 64 electrodes, which implies the application of spatial clustering. The aim here is to isolate groups of electrodes that present significant oscillatory behavior for a particular experimental condition relative to another.

These analyses were carried out for the following contrasts:

- **Non-sarcastic Profession + Ironic** vs. **Non-sarcastic Profession + Literal**: Integration of extra-linguistic information with non-sarcastic profession.
- **Sarcastic Profession + Ironic** vs. **Sarcastic Profession + Literal:** Integration of extra-linguistic information with sarcastic profession.
- **First-name + Ironic** vs. **First-name + Literal:** Comprehension of non-literal language without integration of extra-linguistic information.
- **First-name + Ironic** vs. **Non-sarcastic + Ironic:** Integration of extra-linguistic information.
- **First-name (Ironic + Literal)** vs. **Sarcastic Profession (Ironic + Literal):** Not sure?
- **First-name (Ironic + Literal)** vs. **Non-sarcastic Profession (Ironic + Literal):** Integration of extra-linguistic information?

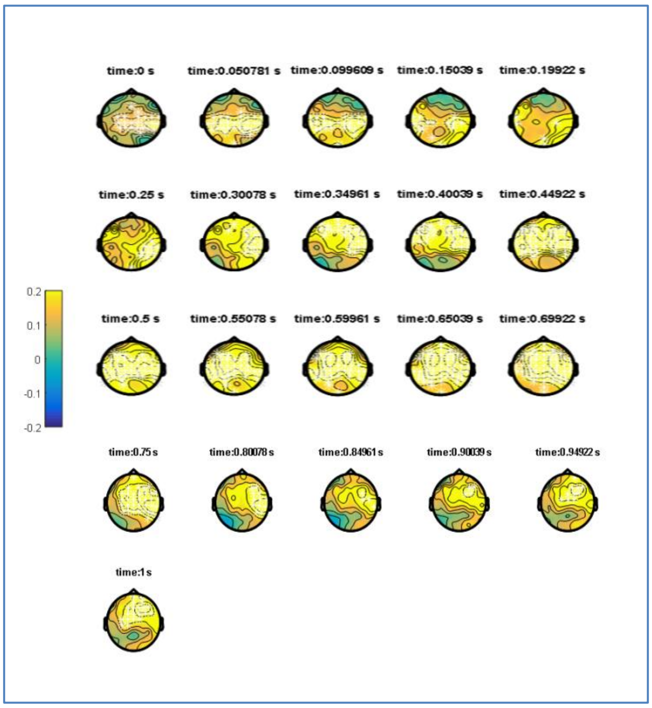

Figure 3: *Electrode clusters presenting significantly greater synchronization for **NSI** condition than **NSL** condition at 5Hz (theta) according to a cluster-based permutation test ($p \leq 0.025$, two-tailed) with 2000 permutations. The clusters are marked in a white X and are plotted against topographies presenting the raw effect: NSI –NSL.*

Analysis of patterns of synchronization and desynchronisation

To analyze the relative patterns of synchronization and desynchronisation by comparing the baseline and post-stimulus intervals cluster-corrected z-scores ($p \leq 0.05$) were calculated. The baseline applied was -250ms to -50ms. Figure 4 presents the cluster-corrected z-scores for conditions SI (left) and SL (right). The blue areas that we can observe in SI in particular indicate statistically significant ($p \leq 0.05$) desynchronisation in the alpha band around 500ms, while the yellow areas indicate significant synchronization and can be observed for the theta band in the 300 to 500ms time for SL in particular.

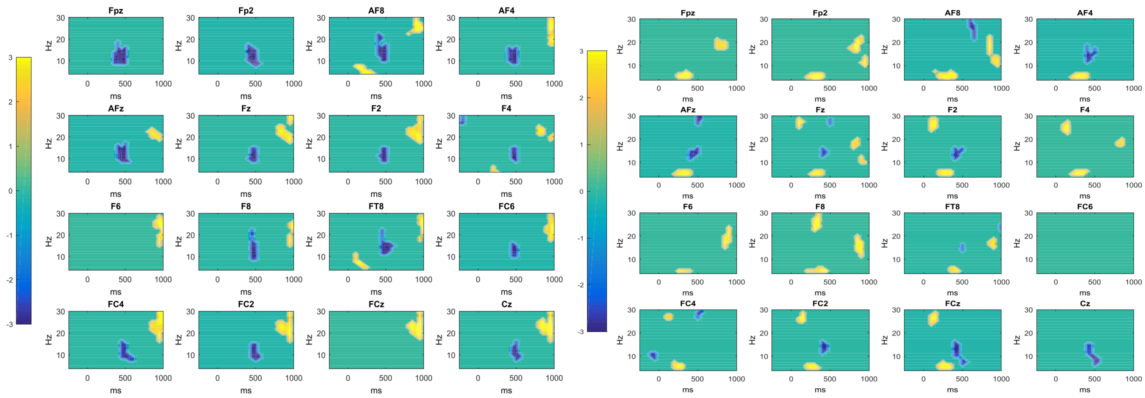

Figure 4: ***Left** : Significant patterns of synchronisation (yellow) and desynchronisation (blue) for condition **SI** for right fronto-central electrodes. **Right**: Significant patterns of synchronisation (yellow) and desynchronisation (blue) for condition **SL** for right fronto-central electrodes. A mask was applied so that only statistically significant activity according cluster-corrected z-score calculation is shown.*