The polyphonic brain: Extracting the neural representation of rhythmic structure for separate voices of polyphonic music using ERP.

Irene Sturm1,2,3, Matthias Treder2, Sven Dähne2, Daniel Miklody2, Hendrik Purwins5, Benjamin Blankertz2,4
Gabriel Curio1,3,4
1Berlin School of Mind and Brain
2Technische Universität Berlin, Neurotechnology Group
3Charité Medicine Berlin, Department of Neurology, Neurophysiology Group
4Bernstein Focus Neurotechnology, Berlin
5Sound and Music Computing Group, Aalborg University Copenhagen

ERPs responses to the onset structure of music

Rapid changes in the stimulus envelope (indicating tone onsets) elicit an N1-P2 ERP response, as has been shown for clicks and sine waves [1], musical tones [2] and for speech [3]. Canonical Correlation Analysis with temporal embedding (kCCA) as a multivariate correlation-based method, allows to extract brain responses to these changes in continuous auditory stimuli.

Here, we
- (1) probe, whether kCCA can be applied to track changes in the stimulus envelope in the EEG of subjects who were presented with semi-artificial monophonic music clips of three instruments.
- (2) On polyphonic trials, composed of the same parts as in (1), we explore, whether the kCCA-filters derived in (1) can recover a representation of each instrument’s part from the EEG where subjects listened to the polyphonic stream.
- (3) We explore, whether, eventually, such a representation is influenced by focused attention.

Experiment (N=11), 64-channel EEG

Re-analysis of data from a ‘Musical’ brain computer interface application (MusicBCI) [4]: control of a user interface by shifting attention to one of three musical instruments (drums, keyboard, bass) that form a complex semi-naturalistic, polyphonic music stimulus.

- Music clips of 40 s duration: sequences of a repetitive standard and a deviant pattern, resembling a minimalistic version of Depeche Mode’s “Can’t get enough” (1980s Electro Pop).
- 63 polyphonic trials, 14 monophonic trials for each instrument
- Task: count silently number of deviant patterns in target instrument.

Methods: Canonical Correlation Analysis with temporal embedding (kCCA)

Idea:
- Design data-driven spatio-temporal filters [5] that maximize the correlation between the power slope of the music signal and EEG.
- Capture complex brain responses by integrate full set of electrodes and a range of time lags (50-250ms).
- Optimize and evaluate filter on separate portions of the data (cross-validation).

Steps of analysis:
- Pre-processing: Sampling frequency 100 Hz, HP filter at 1Hz
- Temporal embedding: Add to EEG data set \( A_t \), additional dimensions that are copies of \( X_t \), time-shifted by \( \ldots, -25 \) data points \( X_{t-25}, \ldots, X_{t} \). This allows to capture brain responses within a latency of 50 to 250ms.
- k-fold Leave-One-Out cross-validation: divide data set of k trials into training and test set, so that each trial is once the test set and the remaining clips are the training set.
- CCA (k folds): train spatio-temporal filter on training set, apply to test set. Correlate resulting uni-dimensional EEG projection with power slope of music signal.
- Non-parametric permutation-based testing: assess significance of derived correlation in permutation tests with surrogate data [6] that has the same spectral components, but random phases.
- Source reconstruction from kCCA patterns: MUSIC algorithm ([7])

Results: Reconstruction of monophonic trials

- Significant correlation of kCCA-filtered EEG (averaged for the 14 monophasic presentations of each instrument) with the respective power slope in 6/11 subjects for drums, in 9/11 subjects for bass, and in 11/11 subjects for keyboard.
- Grand Average: Significant correlation for drums \((n=0.61, p<0.002)\) and bass \((n=0.52, p<0.004)\).

subject drums bass keyboard
| S1 | 0.63 | 0.34 | 0.30 |
| S2 | 0.80 | 0.25 | 0.26 |
| S3 | 0.27 | 0.51 | 0.24 |
| S4 | 0.50 | 0.34 | 0.17 |
| S5 | 0.76 | 0.25 | 0.32 |
| S6 | 0.35 | 0.15 | 0.27 |
| S7 | 0.45 | 0.48 | 0.24 |
| S8 | 0.42 | 0.56 | 0.22 |
| S9 | 0.38 | 0.4 | 0.31 |
| S10 | 0.28 | 0.44 | 0.26 |
| S11 | 0.28 | 0.35 | 0.09 |

Results: Source reconstruction

Decomposing the kCCA patterns results in a fronto-central component, resembling the topography of the N1/P2 complex. The scalp pattern is consistent for the three instruments, the temporal evolution differs.

Drums

Bass

Keyboard

Results: Application of instrument-specific kCCA filters on polyphonic trials

- Applying the kCCA filters (trained on monophasic trials) on the polyphonic stimulus extracts the power slope of each single instrument with significant correlation in: 7/11 subjects for drums, in 3/11 subjects for bass, in 7/11 subjects for keyboard.
- Grand Average: Significant correlation only for keyboard \((p<0.001)\).
- In 7/11 subjects the power slopes of two instruments can be reconstructed in parallel from the (same) EEG where the polyphonic stimulus was presented.

subject drums bass keyboard
| S1 | 0.36 | 0.28 | 0.31 |
| S2 | 0.11 | 0.05 | 0.33 |
| S3 | 0.26 | 0.13 | 0.17 |
| S4 | 0.12 | 0.07 | 0.29 |
| S5 | 0.32 | 0.05 | 0.49 |
| S6 | 0.03 | 0.34 | 0.21 |
| S7 | 0.02 | 0.23 | 0.22 |
| S8 | 0.06 | 0.08 | 0.09 |
| S9 | 0.28 | 0.02 | 0.33 |
| S10 | 0.25 | 0.13 | 0.14 |
| S11 | 0.13 | 0.15 | 0.10 |

Results: Effects of attention in polyphonic stimuli

Correlation between EEG projection and power slope of an instrument was significantly enhanced if this instrument was the target of attention for:
- Drums: S3, \( p<0.007 \)
- Bass: S4, \( p<0.018 \), S6, \( p<0.029 \)
- Keyboard: S7, \( p<0.002 \), S9, \( p<0.002 \) and S10, \( p<0.03 \)

Within the group of subjects a significant effect of attention was present for bass \((p<0.046)\) and for keyboard \((p<0.012)\). The behavioral performance differs for the three instruments, with highest performance for keyboard.

Discussion & Conclusion

- Canonical Correlation Analysis with temporal embedding (kCCA) allows to extract neural responses to the onset structure of a continuous music stimulus.
- Extracted neural sources resemble a N1/P2 complex with a consistent scalp topography across instruments, but varying temporal characteristics. These may reflect the onset properties and the structural characteristics of each instrument’s part.
- In 10/11 subjects instrument-specific kCCA filters recovered the power slope of at least one instrument from the EEG of the polyphonic stimulus. This suggests that subtle characteristics of onsets and rhythmic structure that are ‘learned’ by the kCCA filter could be utilised for the perceptual segregation of a polyphonic piece of music.
- The power slope of keyboard was reconstructed best and most consistently within the group of subjects.
- Tentatively, this may point to an enhanced representation of maticonic instrument in the present P1 response.
- Our preliminary results suggest, that attention may enhance the present brain response to the onset structure of music as suggested in [8] for the typical N1-P2 component. Thus, kCCA may be a promising tool to investigate a listener’s spontaneous fluctuation of attention in continuous music.

References

Acknowledgments
This work is funded by the Berlin School of Mind and Brain, Humboldt Universität zu Berlin.