THE INFLUENCE OF CONTEXTUAL INFORMATION: DISCOVERING SIMILARITIES IN MUSIC AND LANGUAGE PERCEPTION BY MEANS OF ERPS

Björn-Helmer Schmidt
Thomas Gunter
Sonja A. Kotz
Max Planck Institute of Cognitive Neuroscience, Leipzig, Germany

ABSTRACT

In two experiments, we examined the influence of expectancy in language and music by means of event-related brain potentials (ERPs). Therefore, we used familiar stimuli in both domains - idioms in language and well known tunes in music - and violated the endings of the acoustic stimuli (language: idiom vs literal- vs semantic violated ending; music: normal tune vs counterpoint- vs harmonic shift ending). In both studies the ERP response varied as a function of expectancy (more vs less familiar) and condition. In both domains, the data show an early negative component between 100 and 300 ms which is modulated by expectancy and condition. We assume that within this time period music as well as language stimuli are processed at a comparable, basic acoustic level. Thereafter the ERPs show different pattern: primarily a late positivity in language and a frontally accentuated negative shift in music.

1. BACKGROUND

The comparison of music and language has been used as a metaphor to refer to the undescribable aspects of musical experiences. Through modern neuropsychological methods this view again became a useful perspective [1,2]. The perception of both music and language starts with a linear acoustic perception process forming a highly complex hierarchical system in the listener‘s mind. This process is associated with the assembly of contextual information. The easier it is to build up a contextual frame the easier it is to predict the following unit in the acoustic stream. These predictions in turn allow to formulate syntactic and semantic hierarchies. What happens though when these predictions fail?

Here we investigated highly predictable musical and language phrases (familiar tunes and idiomatic phrases) in contrast to unpredictable endings of these phrases while recording ERPs.

It was expected that the violation of highly predictable, cloze probable language phrases would result in two prominent negative ERP components, namely an early negativity, reflecting the reaction to unexpected acoustic phonological features (the phonological mismatch negativity - PMN) [3] and an N400 elicited by unexpected or semantically violated phrase endings [4].

ERP results on music violation appear to be different. The violation of a familiar cloze probable tune by one tone at the end of a phrase results in an early negative component around 200 ms [5] and a large positive deflection (P300 or late positive component - LPC) [6,7].

In our studies, we decided to use larger musical units to violate familiar tunes. These units, so called motives or phrases, are more than a single tone and therefore build up an acoustic gestalt such as words or phrases do in language.

2. AIMS

The focus of these two experiments was to investigate both music and language nearly at the same methodological level with a similar paradigm. This should allow to compare ERPs in both domains in order to find similarities and differences.

3. METHOD

3.1. Material

Idiomatic phrases and well known tunes were rated for familiarity resulting in 40 more and 40 less familiar idioms, and 40 more or less familiar tunes.

For the language study, we created two variations for each idiom: a literal plausible ending and a semantically implausible ending. Thus, we got 240 critical items completed by 80 semantically implausible filler phrases. The phrases were embedded into matrix sentences with the same syntactical structure (see fig. 1). The experimental items were spoken by a female speaker and digitally recorded.

example: ins Gras beißen (to bite the dust)

idiomatic condition:
Mike denkt, dass Clara ins Gras beißen würde.
Mike thinks, that Clara in the grass bite would. (literal translation)

literal condition:
Mike denkt, dass Clara ins Gras gehen würde.
Mike thinks, that Clara in the grass go would. (literal translation)

extreme semantic violation:
Mike denkt, dass Clara ins Gras lachen würde.
Mike thinks, that Clara in the grass laugh would. (literal translation)

Figure 1. example sentence with its conditions

ISBN 3-931852-67-9
ISSN 1617-6847

For the music study, we composed two variations for each familiar tune: the counterpoint of the ending of the phrase and its harmonic shift (see fig. 2). All stimuli were embedded into a neutral unfamiliar melodic environment. This resulted in 240 critical items which were completed with 80 unfamiliar filler items. The musical stimuli were played by a cello player and were digitally recorded.

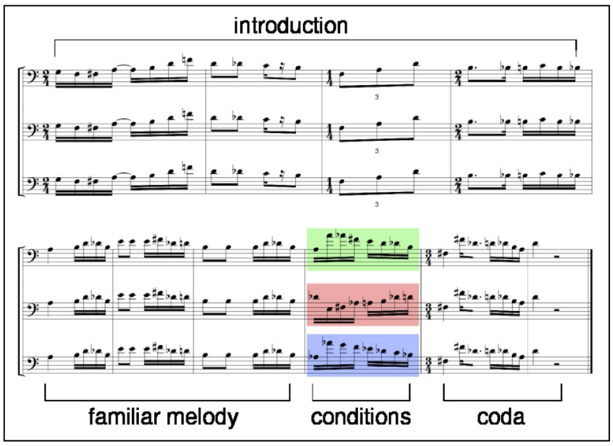

Figure 2. displays a representative stimulus. The familiar melody is "Can Can" by Jaques Offenbach. (upper staff: normal condition; middle staff: counterpoint condition; lower staff: semitone shift condition)

For both studies, we recruited 24 volunteers each (language study: 13 females; music study: 15 females). They were all right handed with normal auditory and normal or corrected to normal visual capacity.

All participants of the language study were native German speakers. The participants of the music study were students of the University of Music „Felix Mendelssohn Bartholdy", Leipzig and played different instruments professionally.

3.3. Procedure

In each study the participants sat in a sound attenuated room in front of a screen. They were instructed to focus on a cross in the middle of the screen and to prevent eye blinks during the auditory presentation of the items. Stimuli were presented in a controlled randomised order. The conditions of each item (plus filler-items) were presented in four presentation blocks. Repetition within each block was prevented. The serial order of the blocks was permutated resulting in a personal block order for each participant within each study. During the experimental presentation we recorded a triggered electroencephalogram (EEG) and additionally the electrooculogram (EOG) to exclude EEG epochs with eye artefacts from the following analysis. The EEGs were off-line averaged in mean ERPs for all subjects and each condition.

4. RESULTS & DISCUSSION

Early negativity in language and music

In both studies, we found a large posterior negative deflection between 100 and 300 ms. This early negativity varied in each case as a function of violation type and familiarity. In *language*, it seems probable that this component results from two factors a) the phonological mismatch detection and b) a semantic integration problem. However, for *music* it is more likely to classify this component in relation to a basic acoustic mismatch reaction reflected in an MMN [8]. We propose that this early negativity is mainly driven by a fast mismatch detection process on a more basic acoustic as well as phonological level modulated by the grade of expectancy. However, it is not a classical mismatch negativity due to its posterior scalp distribution.

Later positivity in language

Only in the language study, this negative component was followed by a posterior positivity which was seen in the extreme semantic violation. This positivity could be viewed as a P600 which is often described as a response to syntactic violations in language. However, due to the task (plausibility judgement), we are not able to discriminate whether this positivity reflects a syntactic violation (P600) or a target detection (P300). The fact that only the extreme violation shows this component might indicate that the response reflects target detection, as a number of idiomatic phrases show syntactical implausibilities, which did not result in a P600. Therefore, we argue that this positivity reflects a target P300. A positivity in the same time window in music did not reach significance.

Negative shift in music

In the music study we found a long lasting negative shift at all electrodes which varied as a function of expectancy. The maximal amplitude of the negative shift had a frontal distribution and was significant between 800 and 1400 ms. Frontal negative shifts are often interpreted as reflecting memory processes [9]. Thus, we would like to argue that this frontal negative shift is reflecting an ongoing memory retrieval and integration process as shown for the visual domain [10]. A visible posterior shift for literal variations of idiomatic phrases in relation to the idiomatic phrases was not significant.

ISBN 3-931852-67-9
ISSN 1617-6847

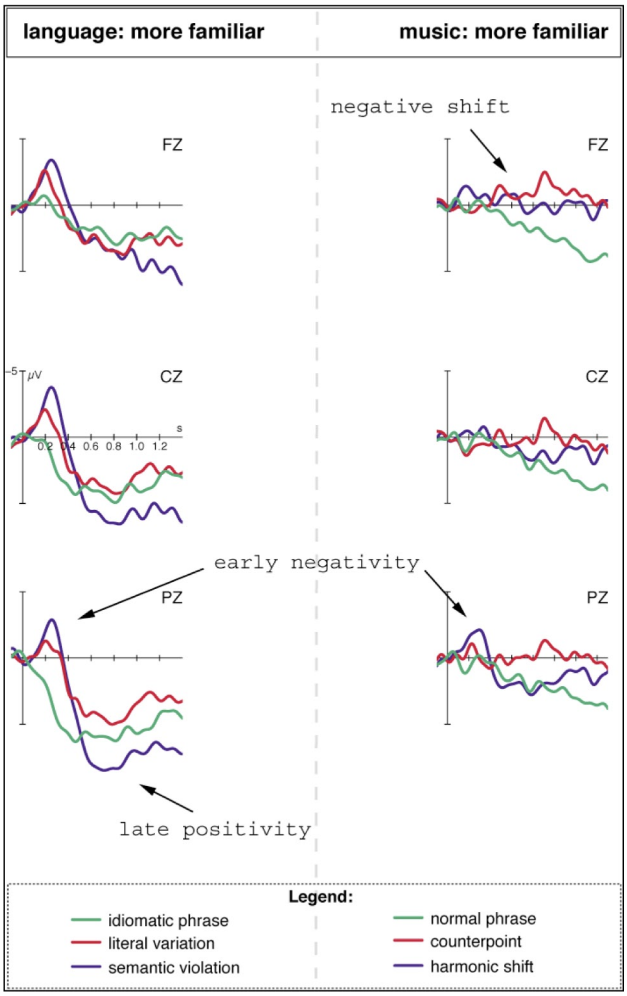

Figure 3. mean ERPs at midline electrodes for more familiar stimuli. Left: language study; right: music study.

5. CONCLUSION

The data show that music and language share common cognitive processes. However, these processes are obviously limited to a very small time window between 100 and 300 ms. It is plausible that this negativity is reflecting a cognitive mechanism in which basic acoustic features are checked against an established template build by a particular context. If the template does not fit the incoming acoustic features (e.g., frequency, onset-characteristics, timbre) the system responds with this relatively fast mismatch reaction.

Thereafter, it seems that the cognitive processes of music and language start to differ. The main difference seems to result from the very special timing dimensions of language and musical units. It is obvious that the length of words is limited by its plausible number of syllables in contrast to musical units where motives or phrases are not so strongly limited in their timing dimension. The cognitive strategy of listeners of music seems to be, to hold musical information as long as possible in memory so that a musical unit can be finished as a unitary gestalt. This might be reflected by the long lasting negative shift in the ERP.

The imaging methods in neuropsychology offer new possibilities to benefit from the old language-music metaphor. One possible insight should be to put away the old overstatement ‚music might be a universal language'. Music shares some structural aspects with the language system, which should result in sharing mutual cognitive resources.

6. REFERENCES

1. Besson, M. & Macar, F (1987). An event-related potential analysis of incongruity in music and other non-linguistic contexts. *Psychophysiology*. 24, 1, 14-25.
2. Zatorre, R.J., Belin, P. & Penhune V.B. (2002). Structure and function of auditory cortex: music and speech. *TRENDS in Cognitive Science*. 6, 1, 37-46.
3. Connolly, J.F. & Phillips, N.A. (1994). Event-related potential components reflect phonological and semantic processing of the terminal word of spoken sentences. Journal of Cognitive Neuroscience. 6, 3, 256-266.
4. Kutas, M. & Hillyard, S.A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 207, 203-204
5. Paller, K.A., McCarthy, G. & Wood, C.C (1992). Event-related potentials elicited by deviant endings to melodies. *Psychophysiology*. 29, 2, 202-206.
6. Granot, R. & Donchin, E. (2002). Do Re Mi Fa Sol La Ti - Constraints, Congruity, and Musical Training: An Event-Related Brain Potentials Study of Musical Expectancies. *Music Perception*. 19, 4, 487-528.
7. Besson, M. & Fa(ii)ra, F. (1995). An event-related potential study of music expectancy: comparison of musicians and nonmusicians. *Journal of Experimental Psychology: Human Perception and Performance*. 21, 6, 1278-1296.
8. Näätänen, R. (1992). *Attention and brain function*. Lawrence Erlbaum Associates Pub. Hillsdale, New Jersey.
9. Vos, S.H., Gunter, T.C., Kolk, H.H.J. & Mulder, G. (2001). Working Memory constraints on syntactic processing: An electrophysiological investigation. *Psychophysiology*. 38, 41-63.
10. Gunter, T.C., Schmidt, B.-H. & Besson, M. (in press). Let's face the music: A behavioural and electrophysiological exploration of score reading. *Psychophysiology*.

ISBN 3-931852-67-9
ISSN 1617-6847