Biological Psychology 82 (2009) 70–81

Contents lists available at ScienceDirect

Biological Psychology

journal homepage: www.elsevier.com/locate/biopsycho

Non-musicians' perception of phrase boundaries in music: A cross-cultural ERP study

Yun Nan [a], Thomas R. Knösche [b,*], Angela D. Friederici [b]

[a] State Key Laboratory of Cognitive neuroscience and Learning, Beijing Normal University, Beijing, China
[b] MPI for Human Cognitive and Brain Sciences, Leipzig, Germany

ARTICLE INFO

Article history:
Received 4 January 2008
Accepted 10 June 2009
Available online 18 June 2009

Keywords:
Music perception
Phrase structure
Cultural differences
Event-related potentials
Non-musicians

ABSTRACT

The present study investigates neural responses to musical phrase boundaries in subjects without formal musical training, with special emphasis on the issue of cultural familiarity (i.e., the relation between the enculturation of the subjects and the cultural style of the presented music). German and Chinese non-musicians listened to Western and Chinese melodies which contained manipulated phrase boundaries while event-related potentials (ERP) were measured. The behavioral data clearly showed that melodic phrase boundary perception is influenced by cultural familiarity. The ERP revealed a series of positive and negative peaks with latencies between 40 ms and 550 ms relative to the phrase boundary offset, all of which were influenced by the phrase melodic structure type. In contrast, cultural familiarity only influenced phrase boundary processing at longer latencies, reflected by a P3-like component peaking at 280 ms.

At about 450–600 ms post phrase boundary offset, we observed a slightly right-lateralized music closure positive shift (CPS), which has been reported as a marker for phrase boundary processing in musicians in earlier studies. This study demonstrates for the first time that the music CPS can be elicited in non-musicians, suggesting that the underlying phrase boundary processing does not require formal musical training.

1. Introduction

Over recent decades, the question of when and to what extent particular experiences affect the on-line processing of speech has been quite extensively discussed. Effects of different language experiences have been reported to be present quite early during development (Friederici et al., 2007; Näätänen et al., 1997) and to persist throughout life (Werker et al., 1992). In the domain of music, the influence of a specific music experience provided by different cultures has so far rarely been investigated (but see Lynch et al., 1990; Nan et al., 2006) although the role of musical training has been the subject of research (e.g., Besson and Faita, 1995; Schön et al., 2004; Trainor et al., 1999). For a recent review of the roles of musical enculturation and formal musical training during ontogeny, see Hannon and Trainor (2007).

1.1. Musical phrasing

In order to successfully parse music and speech, phrases which structure the incoming information are important (Chiappe and Schmuckler, 1997; Dowling, 1973; Stoffer, 1985; Tan et al., 1981; Wilson et al., 1999) not only for adults but also for young infants (Hirsh-pasek et al., 1987; Jusczyk and Krumhansl, 1993; Krumhansl and Jusczyk, 1990). The fast and accurate recognition of phrase boundaries is crucial for the reliable processing of the input underlying phrase structure. Phrase boundaries are marked by both structural (global) and acoustic (local) cues (Trainor and Adams, 2000). For language, the structural cues include grammatical rules and phonotactic constraints, while for music, they consist of functional relations between the different notes of the scale. Acoustic cues to phrase boundaries in both speech (Cooper and Sorensen, 1977; Scott, 1982) and music (see, e.g., Boltz, 1993; Jusczyk and Krumhansl, 1993) include the insertion of a pause, lengthening of the final element(s) of the phrase, decrease in intensity at the end of the phrase, and changes in pitch. A number of studies on music perception indicate that, quite similar to language (Nakatani and Schaffer, 1978), in general, duration-based cues seem to be more important than those based on intensity (Drake and Palmer, 1993; Frankland and Cohen, 2004; Trainor and Adams, 2000), The single most important cue is clearly the pause

* Corresponding author at: Max Planck Institute for Human Cognitive and Brain Sciences, P.O. box 500355, 04303 Leipzig, Germany. Tel.: +49 0 341 35521735; fax: +49 0 341 35521740.
E-mail address: knoesche@cbs.mpg.de (T.R. Knösche).

0301-0511/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.biopsycho.2009.06.002

(Frankland and Cohen, 2004; Neuhaus et al., 2006; Riemann, 1900), but the relevance of other cues has been demonstrated in behavioral (Boltz, 1993; Cuddy et al., 1981; Jusczyk and Krumhansl, 1993; Tan et al., 1981) as well as electrophysiological (Neuhaus et al., 2006) studies.

Recently, electroencephalography (EEG) and magnetoencephalography (MEG) have been used to investigate the time course of the processing of musical phrase boundaries in musicians (Knösche et al., 2005; Nan et al., 2006; Neuhaus et al., 2006) and non-musicians (Neuhaus et al., 2006). In musicians, a centro-parietal positive event-related potential (ERP) component approximately 500 ms after the onset of the first tone following a phrase boundary was identified. Because it bears some resemblance in its mode of elicitation and scalp topography to the Closure Positive Shift (CPS) indexing the processing of intonational phrase boundaries in speech (Pannekamp et al., 2005; Steinhauer, 2003; Steinhauer et al., 1999; Steinhauer and Friederici, 2001), it was labeled *music CPS* (Knösche et al., 2005). Based on their similarity, it was suggested that language and music CPS might reflect related cognitive processes (i.e., the closure of a phrase). If so, this would indicate that phrase structure processing is a cognitive function shared by both music and language, similar to other aspects of structural processing of language and music for which similar neural substrates are reported. For example, ERP components reflecting syntactic processing have been found to be very similar for speech (Friederici et al., 1993) and music (Koelsch et al., 2000; Patel, 1998). Moreover, MEG source analysis (Maess et al., 2001) and functional magnetic resonance imaging (fMRI) (Koelsch, 2005; Koelsch et al., 2002a) showed that many constituents of the cortical 'language network', including Broca's area, are also involved in the processing of musical syntax.

The music CPS as a late ERP component has been found to be sensitive to a number of phrase boundary cues, in particular to pause length, length of the last tone preceding the pause, and harmonic function of this last tone (Neuhaus et al., 2006), and appears to reflect rather late, high-level aspects of phrase boundary processing.

However, effects of music phrase processing on the early auditory components were also observed (Knösche et al., 2005; Neuhaus et al., 2006). Since, in these studies, the critical condition (phrased condition) contained a pause whereas the control condition (unphrased condition) did not, these effects might reflect a confound with habituation and refractoriness of the neuronal populations in the auditory cortices

1.2. The influence of formal musical training

Although there is some work indicating that fundamental processes of music cognition may be similar in all normal humans regardless of training (Bigand et al., 1999; Bigand and Pineau, 1997; Koelsch et al., 2000; Regnault et al., 2001), other studies demonstrated that musicians and non-musicians differ in many aspects when processing music and also other auditory input. These aspects include detection performance for temporal and dynamic perturbations (Repp, 1995), perception of loudness, timbre and pitch (Hoover and Cullari, 1992), as well as brain activity during passive listening to musical stimuli (Morrison et al., 2003; Shahin et al., 2003) and in response to music expectancy violations (Besson et al., 1994; Fujioka et al., 2004; Hantz et al., 1992; Janata, 1995; Jongsma et al., 2005; Koelsch et al., 2005; Regnault et al., 2001; Trainor et al., 1999; van Zuijen et al., 2004). Musical training seems to play a role in music phrase structure perception as well, as indicated by various studies (Chiappe and Schmuckler, 1997; Frankland and Cohen, 2004; Tan et al., 1981). In this line, it has been demonstrated that the occurrence of the music CPS seems to depend on an individual's formal musical training. In an ERP study, Neuhaus et al. (2006) found that, in contrast to formally trained musicians, non-musicians did not exhibit a CPS in response to phrase boundaries, but an earlier and frontally distributed negativity was found. MEG recordings of non-musicians, on the other hand, did show a CPSm (magnetic equivalent to the CPS) qualitatively similar to the one for musicians, although much lower in amplitude. This apparent difference between EEG and MEG may be due to the fact that the two methods reflect different aspects of the neural substrate of psychological processes (see discussion in Shahin et al., 2003). MEG, for example, is most sensitive to sulcal activity and almost silent for sources on the crowns of gyri, while EEG yields highest amplitudes for sources which are located near the brain surface and point towards the skull (thus on the gyral crowns). The combined results appear to indicate that phrase boundary processing in musician and non-musician takes place in the time range of the CPS, but that the underlying processes differ. This interpretation is also supported by the finding that the time courses of the music CPS and CPSm in musicians differ considerably (single peak vs. double peak), suggesting that they are not entirely based on the same generators (Neuhaus et al., 2006). These findings suggest that the CPS (at least partially) reflects processes that are modulated by formal musical training and are thus more modifiable than previously thought.

A final conclusion, however, must await further research, as the reported result might have been due to the specific experimental setting. In the study by Neuhaus and colleagues (2006), participants were asked to scan the input for false (off-key) notes, a task that is particularly demanding for non-musicians and requires a rather local, element-directed focus of attention. This might have prevented then non-musicians from processing the input in the global, holistic way necessary to appreciate phrase structure and to anticipate phrase boundaries, while for musicians the detection of the off-key notes required little effort and normal holistic processing of the melodies remained unimpaired. Hence, it remains to be investigated whether under different task conditions, non-musicians would exhibit similar electrophysiological correlates for phrase processing as musicians, indicating similar underlying brain processes.

1.3. The influence of culture

Music is culturally tuned. Musical styles from different cultures differ in many respects, such as pitch structure, rhythmic and metric rules, and timbre of main instruments, to name but a few. The implicit exposure to music of a particular cultural tint shapes the musical perception of individuals in that environment. This enculturation starts in infancy, when basic universal features are acquired first, forming the basis for the later acquisition of system specific aspects (Hannon and Trainor, 2007). Behavioral research by Lynch et al. (1990) found that, in contrast to adults, 6-year-old American children were equally sensitive to a mistuned tone in a non-Western scale context as in a Western one. Similarly, Trainor and Trehub (1992) reported that, although adults from a diatonic tonal environment find non-diatonic changes easier to detect than diatonic ones, infants find both types of change equally easy to detect. Jusczyk and Krumhansl (1993) suggested that the basic grouping principles appear to be quite reliable and robust in infancy. These basic grouping principles might facilitate the acquisition of the particular intervalic structures of a musical culture. Cross-cultural behavioral studies (Castellano et al., 1984; Kessler et al., 1984) showed that listeners unfamiliar with a particular musical style tend to use these cues for identifying the structurally important tones.