Neural representation of a melodic motif: Effects of polyphonic contexts

Madeline Huberth a,⇑, Takako Fujioka a,b

a Center for Computer Research in Music and Acoustics, Department of Music, Stanford University, United States

b Stanford Neurosciences Institute, Stanford University, United States

Abstract

In music, a melodic motif is often played repeatedly in different pitch ranges and at different times. Event-related potential (ERP) studies have shown that the mismatch negativity (MMN) reflects memory trace processing that encodes two separate melodic lines (“voices”) with different motifs. Here we investigated whether a single motif presented in two voices is encoded as a single entity or two separate entities, and whether motifs overlapping in time impede or enhance encoding strength. Electroencephalogram (EEG) from 11 musically-trained participants was recorded while they passively listened to sequences of 5-note motifs where the 5th note either descended (standard) or ascended (deviant) relative to the previous note (20% deviant rate). Motifs were presented either in one pitch range, or alternated between two pitch ranges, creating an “upper” and a “lower” voice. Further, motifs were either temporally isolated (silence in between), or temporally concurrent with two tones overlapping. When motifs were temporally isolated, MMN amplitude in the one-pitch-range condition was similar to that in the two-pitch-range upper voice. In contrast, no MMN, but P3a, was observed in the two-pitch-range lower voice. When motifs were temporally concurrent and presented in two pitch ranges, MMN exhibited a more posterior distribution in the upper voice, but again, was absent in the lower voice. These results suggest that motifs presented in two separate voices are not encoded entirely independently, but hierarchically, causing asymmetry between the upper and lower voice encoding even when no simultaneous pitches are presented.

1. Introduction

Listeners of music must parse multiple simultaneous sounds and connect them to form perceptual objects such as melody, harmony, and rhythm. In contrast to music that features a single dominant melody against a background texture, some music is structured with two or more lines of independent melody that occur concurrently. The resulting musical texture, called polyphony, requires the listener to simultaneously follow each melodic line, or voice, as it unfolds in time while integrating simultaneous notes as harmony. In some polyphonic musical forms, such as the fugue, a single melodic motif (i.e., a short succession of notes that has its own identity based mainly on melodic contour), is repeatedly played in the different voices in different pitch ranges (Walker, 2001). While recognizing each motif may seem an easy task, it is often the case that the next motif starts before the previous motif concludes. In this situation, listeners have to not only track the two motifs at the same time, but also each with a distinct time stamp in reference to the entry point. This type of music can be played even in a single instrument, as evidenced in the rich repertoire of canons composed for keyboard instruments, for which voices are differentiated by pitch, but not much by timbre. How our brain works for such a complex challenge entails important questions about our auditory perceptual and cognitive functions. The automatic process of separating auditory streams in the acoustic environment and integrating them into perceptual objects is referred to as auditory scene analysis (Bregman, 1990). Generally, sounds that have similar acoustic features are perceived as one object, coming from one source, whereas sounds with acoustically different features are perceived as segregated objects, coming from different sources. Polyphonic music gives us an interesting opportunity to study interactions between stream segregation and memory encoding where both functions are presumably active and possibly interacting with each other when processing multiple concurrent voices.

Neural correlates of auditory stream segregation have been examined through event-related potentials (ERP), and in particular, by using the mismatch negativity (MMN) response. The MMN is typically observed as a negative voltage deflection in the difference wave obtained by subtracting the ERPs to frequent “standard” stimuli from that of infrequent “deviant” stimuli (Näätänen, 1990).
Oddball sequences can be made naturally resembling dense and intricate musical textures.

In music, a melody can be considered a single auditory stream, requiring sequential integration of pitch and duration information. With regards to melody encoding, the MMN has been shown to reflect detection of contour (general direction of pitch change) and interval (precise pitch distance) changes in melodic patterns, as found in electroencephalography (EEG) and magnetoencephalography (MEG) studies (Fujioka, Trainor, Ross, Kakigi, & Pantev, 2004; Tervaniemi, Rytkönen, Schröger, Ilmoniemi, & Näätänen, 2001; Trainor, McDonald, & Alain, 2002). While both nonmusicians (Trainor et al., 2002) and musicians (Tervaniemi et al., 2001) detect contour changes as reflected in MMN, MMN is larger in musicians than nonmusicians in response to contour changes (Fujioka et al., 2004), suggesting that musical training enhances the ability to automatically register abstract changes in the relative pitch structure of melodies.

In addition to the MMN, the P3a response is often elicited in response to deviants in a train of frequent standards, even if the listeners’ attention is not directed to the sound. The P3a response is thought to reflect an early attention process resulting from a representational change in working memory operating mainly in the frontal lobe (Polich, 2007), and has been found to exhibit multiple subcomponents. When found with a frontocentral distribution and a ~230 ms latency, it is thought to represent inadvertent capture of attention, while a later subcomponent, peaking at ~315 ms and right-frontally dominant, reflects the actual orienting of attention (Escera, Alho, Winkler, & Näätänen, 1998). Although P3a amplitude may not relate directly to MMN amplitude (Horváth, Winkler, & Bendixen, 2008), it is also hypothesized that the MMN not only reflects processing at a preattentive level, but that the process underlying MMN also triggers the switching of attention to potentially important events in the unattended auditory environment (Giard, Perrin, Perrier, & Bouchet, 1990; Näätänen et al., 1978) as if the process related to P3a can take the output of the MMN-related process as its input (Escera et al., 1998).

Although task manipulations can affect the magnitude of the P3a (Comerchero & Polich, 1999; Katayama & Polich, 1999), a stimulus can elicit a P3a even if it is ignored (Schwent, Hillyard, & Galambos, 1976). Particularly in response to music, the P3a response appears to be quite robust, occurring in response to harmonic and melodic deviant stimuli in both active (Janata, 1995; Trainor, Desjardins, & Rockel, 1999) and passive paradigms (Seppänen, Pesonen, & Tervaniemi, 2012). While some studies report that its amplitude in response to musical stimuli grows larger with musical training (Putkinen, Tervaniemi, Saarikivi, Ojala, & Huotilainen, 2014), others find no such training effects (Trainor et al., 1999).
دریافت فوری متن کامل مقاله

امکان دانلود نسخه تمام متن مقالات انگلیسی
امکان دانلود نسخه ترجمه شده مقالات
پذیرش سفارش ترجمه تخصصی
امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
امکان دانلود رایگان ۲ صفحه اول هر مقاله
امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
دانلود فوری مقاله پس از پرداخت آنلاین
پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات