NeuroImage 159 (2017) 185–194

Contents lists available at ScienceDirect

NeuroImage

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

Trading of dynamic interaural time and level difference cues and its effect on the auditory motion-onset response measured with electroencephalography

Christian F. Altmann[a,*], Ryuhei Ueda[b], Benoit Bucher[b], Shigeto Furukawa[c], Kentaro Ono[a,d], Makio Kashino[c], Tatsuya Mima[e], Hidenao Fukuyama[a]

[a] Human Brain Research Center, Graduate School of Medicine, Kyoto University, Kyoto 606-8507, Japan
[b] Department of Psychology, Graduate School of Letters, Kyoto University, Kyoto 606-8501, Japan
[c] NTT Communication Science Laboratories, NTT Corporation, 3-1 Morinosato Wakamiya, Atsugi, Kanagawa 243-0198, Japan
[d] Center of KANSEI Innovation, Hiroshima University, Hiroshima 734-0037, Japan
[e] Graduate School of Core Ethics and Frontier Science, Ritsumeikan University, Kyoto 603-8577, Japan

A R T I C L E I N F O

Keywords:
Auditory evoked potentials
Auditory motion
Cue integration
Electroencephalography
Motion-onset response
Spatial hearing

A B S T R A C T

Interaural time (ITD) and level differences (ILD) constitute the two main cues for sound localization in the horizontal plane. Despite extensive research in animal models and humans, the mechanism of how these two cues are integrated into a unified percept is still far from clear. In this study, our aim was to test with human electroencephalography (EEG) whether integration of dynamic ITD and ILD cues is reflected in the so-called motion-onset response (MOR), an evoked potential elicited by moving sound sources.

To this end, ITD and ILD trajectories were determined individually by cue trading psychophysics. We then measured EEG while subjects were presented with either static click-trains or click-trains that contained a dynamic portion at the end. The dynamic part was created by combining ITD with ILD either congruently to elicit the percept of a right/leftward moving sound, or incongruently to elicit the percept of a static sound.

In two experiments that differed in the method to derive individual dynamic cue trading stimuli, we observed an MOR with at least a change-N1 (cN1) component for both the congruent and incongruent conditions at about 160–190 ms after motion-onset. A significant change-P2 (cP2) component for both the congruent and incongruent ITD/ILD combination was found only in the second experiment peaking at about 250 ms after motion onset.

In sum, this study shows that a sound which – by a combination of counter-balanced ITD and ILD cues – induces a static percept can still elicit a motion-onset response, indicative of independent ITD and ILD processing at the level of the MOR – a component that has been proposed to be, at least partly, generated in non-primary auditory cortex.

1. Introduction

Tracking of moving objects in our environment is critical for our survival. Think of the threat of an approaching predator or the necessity to capture prey. Even in our modern times in which hunting and being hunted are rather rare occasions, the perception of moving objects helps us in many instances such as navigating urban traffic. When visual contact cannot be made, such as in a dark night, in a dense forest, or when traffic and buildings block our view, we have to rely on the sense of hearing, even though hearing's capacity to resolve motion is inferior to vision (Carlile and Best, 2002; see, e.g., Middlebrooks, 2015 for a review).

Like localization of static sounds, motion perception in the horizontal plane depends on changes in interaural time differences (ITD) and interaural level differences (ILD). The “duplex theory” of sound localization posits that the two cues dominate localization in different frequency ranges (Strutt, 1907): while ITD mainly mediates localization for lower frequencies below about 1.5 kHz, ILD is more important in the

Abbreviations: cN1, change-N1; cP2, change-P2; EEG, electroencephalography; ERP, event-related potential; fMRI, functional magnetic resonance imaging; ILD, interaural level difference; ITD, interaural time difference; MEG, magnetoencephalography; MMN, mismatch negativity; MOR, motion-onset-response; PET, positron emission tomography; SL, sensation level; 2AFC, two-alternative-forced-choice.

* Corresponding author. Human Brain Research Center, Graduate School of Medicine, Kyoto University, 54 Shogoin Kawaracho, Sakyo-ku, Kyoto 606-8507, Japan.
E-mail address: christian.altmann.6z@kyoto-u.ac.jp (C.F. Altmann).

http://dx.doi.org/10.1016/j.neuroimage.2017.07.055
Received 6 April 2017; Received in revised form 15 June 2017; Accepted 25 July 2017
Available online 27 July 2017

higher frequency ranges above about 4 kHz (Stevens and Newman, 1936; Mills, 1958). Consistently, physiological descriptions of the binaural pathway show a segregation of low and high frequency processing in the superior olivary complex (see, e.g., Grothe et al., 2010 for a review).

In real life conditions, sound sources give rise to both ITD and ILD cues for sound localization, and at a psychophysical level, we need to integrate them to attain the percept of a single object. In fact, the two cues can – to some extent – even be combined to counter-balance each other: i.e., a sound with left-lateralized ITD and right-lateralized ILD can be perceived as originating from midline. This phenomenon is called ITD/ILD cue trading (Shaxby and Gage, 1932; Hafter and Jeffress, 1968; Stecker, 2010). In a previous study, we tested with electroencephalography (EEG) whether the mismatch negativity (MMN) reflects ITD and ILD in a manner suggesting independent or integrated processing (Altmann et al., 2014): we found significant MMN when after a midline standard sound a deviant with individually counter-balanced ITD/ILD cues was presented. In this study, the percept of sound location did not change from standard to deviant sound, only the underlying ITD/ILD cues, therefore our EEG data suggested independent processing of ITD/ILD cues at the level of the MMN. However, ITD/ILD processing at the level of the MMN turned out to be more complicated than that, as the MMN responses to combined ITD/ILD cues did not exhibit linear additivity, in particular in the lower frequency range, indicating some degree of cue integration.

Our previous study investigated ITD/ILD cue trading for static stimuli at the level of the MMN. This led us to the question, how far ITD and ILD cues for dynamic, i.e., moving stimuli are processed separately. Interestingly, EEG has been shown to be sensitive to sound motion in event-related potential (ERP, Halliday and Callaway, 1978; see Ahveninen et al., 2014 for a review) and MMN studies (Altman et al., 2005, 2010; Altmann et al., 2013). To distinguish the ERP in response to the motion onset from the sound energy onset, several studies have employed sound stimuli that consisted of a sustained static part followed by a dynamic part (Bidet-Caulet and Bertrand, 2005; Krumbholz et al., 2007; Getzmann, 2009). The onset of the static period elicits a sound energy response and the motion onset results in a clearly distinguishable ERP: the motion-onset response (MOR). It has been described to consist of a small positive deflection, cP1 (change-P1) at about 76 ms after motion onset, followed by a prominent negative cN1 (change-N1) peaking at about 140 ms, and two positive deflections (change-P2), cP2a at 228, and cP2b at 322 ms after motion onset (Krumbholz et al., 2007). Its cortical generators have been proposed to lie within higher non-primary auditory cortex (Krumbholz et al., 2007). This has been corroborated by sensitivity to moving compared to static sounds in non-primary but not primary auditory cortex in a human functional magnetic resonance imaging (fMRI) study (Krumbholz et al., 2005). Depending on the hemifield (left or right) in which the motion occurs and depending on the motion direction, the MOR exhibits pronounced hemispheric asymmetry (Krumbholz et al., 2007; Getzmann, 2009).

The aim of this current study was to test whether ITD and ILD are processed separately at the level of the MOR. To this end, we created stimuli with dynamic ITD and ILD cues that did not elicit the percept of a moving, but rather that of a static, intracranial image. We measured EEG responses to these stimuli while listeners were watching a silent movie and compared them with MORs to sounds that elicited the percept of a moving sound by a congruent combination of the same ITD and ILD cues. We hypothesized that independent processing of ITD and ILD at the level of the MOR would be reflected by similar ERPs for congruent and incongruent ITD/ILD cue combinations, with both ERPs being clearly discernible from that elicited by a static stimulus. On the other hand, integrated ITD/ILD processing should result in differences between the ERPs elicited by congruent and incongruent ITD/ILD cue combinations. We tested this in two experiments which mainly differed in the way in which the ILDs that counter-balanced the ITD cues were determined: in Experiment 1, participants psychophysically counter-balanced a range of several static ITDs with static ILD cues and we used these values to construct the dynamic ITD and ILD motion trajectories. However, it is possible that the static conditions would not reflect cue trading under dynamic conditions well, in particular given the observation that performance in a minimum audible movement angle task improves for a continuous motion trajectory versus presentation of the endpoints only (Perrott and Marlborough, 1989). Therefore, in Experiment 2, participants psychophysically counter-balanced the ITD and ILD of dynamic click trains. In addition, in Experiment 2 we also included presentation of click trains with non-zero ITD paired with zero ILD cues and vice versa, in order to test whether elicited motion-onset responses to combined cue changes are equal to a linear sum of the single cue changes.

2. Material and methods

2.1. Participants

Experiment 1 was conducted on 24 healthy, normal hearing participants. Data of two participants were discarded because artifact-rejection resulted in dismissal of more than 50% of these participants' data. The remaining sample of 22 consisted of 11 male and 11 female participants with an average age of 23.8 years (range: 20–38). All participants were right-handed as determined by self-report. Experiment 2 was conducted on 20 healthy, normal hearing participants. Of these, 5 participated in Experiment 1, with a period of at least 8 months between the two experiments. In Experiment 2, data of two participants were discarded. The analyzed sample of 18 participants consisted of 11 males and 7 females, average age was 24.7 years (range: 21–39), 16 participants were right and two left-handed. Prior to the experiments, all subjects were informed of the aims and risks of the experiment and gave written informed consent. The experiments were performed in accordance with the ethical standards laid down in the declaration of Helsinki of 1964 and the guidelines approved by the local ethics committee of the Graduate School of Medicine and Faculty of Medicine, Kyoto University.

2.2. Experimental stimuli and apparatus

The sound stimuli (sampling rate: 96.0 kHz) consisted of click trains with a duration of 850 ms and a click rate of 100 Hz (see Fig. 1A). The individual clicks were monopolar with an on-duration of 100 μs. Sound pressure was adjusted to 45 dB above the individual sensation level (SL). The first 700 ms of the sound stimuli contained no ITD or ILD cue inducing the percept of a static midline-centered sound source. Following that, the click train continued for 150 ms either being static (in the static condition) or with dynamic ITD and ILD cues. In congruent conditions, ITD and ILD were paired congruently to elicit the percept of a sound source moving to the left or right. In incongruent conditions, ITD and ILD were paired incongruently in a counter-balanced manner to elicit the percept of a static sound source. The ILD was implemented by increasing the level in one ear and decreasing it in the other relative to the center level (45 dB SL), so that the average for the two ears on a decibel scale was the center level. The ITD trajectories were always from 0 μs to ±210 μs and provided fine-structure rather than an envelope ITD cue (for a recent MEG experiment testing sensitivity of auditory cortex for envelope ITD, see Salminen et al., 2015a). The ILDs were determined individually by a psychophysical ITD/ILD cue trading procedure conducted before the EEG experiment. Experiment 2 was conducted in a similar manner, with the additional inclusion of trials with stimuli containing non-zero ITD cues paired with zero ILD cues and vice versa.

Etymotic Research ER4 in-ear-headphones (Etymotic Research Inc., Elk Grove Village, IL, USA) were used to deliver the sound stimuli from a PC via a USB audio interface (M-Audio Fast-Track Pro, M-Audio Inc., Irwindale, CA, USA) using the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007), running in a Matlab environment (The Mathworks Inc., Natick, MA, USA). During all experiments, subjects were seated in a single-walled acoustic booth (AT-66, RION Co., Ltd., Kokubunji, Japan).