Auditory motion: perception and cortical response

Der Fakultät für Biowissenschaften, Pharmazie und Psychologie
Der Universität Leipzig
eingereichte

DISSERTATION

zur Erlangung des akademischen Grades

Doctor rerum naturalium
(Dr. rer. nat.)

vorgelegt
von Master of Arts **Mikaella Sarrou**
geboren am 15.09.1987 in Nikosia, Zypern

Leipzig, den 29.03.2019

Statement of authorship

I assure that this dissertation has been composed without prohibited help and without other materials than stated in the text. Directly or indirectly adopted thoughts and ideas from other people have been specifically indicated and acknowledged. I have not engaged help from a graduation consultant and nobody had ever received direct or indirect valuable considerations for work correlated to the present dissertation. This dissertation has never been submitted before for the award of any other degree in any other tertiary institution.

Selbständigkeitserklärung

Ich versichere hiermit, dass die vorliegende Arbeit ohne unzulässige Hilfe und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt wurde und dass die aus fremden Quellen direkt oder indirekt übernommenen Gedanken in der Arbeit als solche kenntlich gemacht worden sind. Ich habe nicht die Hilfe eines Promotionsberaters in Anspruch genommen und es hat von mir auch nie jemand unmittelbar noch mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen. Die vorgelegte Arbeit wurde noch nie in gleicher oder in ähnlicher Form einer anderen wissenschaftlichen Einrichtung zum Zwecke einer Promotion oder eines anderen Prüfungsverfahrens vorgelegt.

Leipzig, den

Mikaella Sarrou

Mikaella Sarrou

Haeckelstrasse 1, 39104 Magdeburg, Germany
+49(0)15120546819
smikaella@hotmail.com

EDUCATION / ACADEMIC CAREER

2017 – 2020	**Postdoctoral researcher** University of Magdeburg, Magdeburg, Germany
2012 – Now	**Doctoral Candidate** University of Leipzig, Leipzig, Germany
2016	**Visiting Doctoral Researcher** (6 months) International Laboratory for Brain, Music and Sound Research (BRAMS), Montreal, Canada
2009 – 2011	**MA Psychology** (passed with very good) University of Cyprus, Nicosia, Cyprus
2005 – 2009	**BA Psychology** (passed with very good) University of Cyprus, Nicosia, Cyprus

WORK EXPERIENCE

2009 – 2012	Callcenter Agent MTN Cyprus Ltd, Nicosia, Cyprus
2010 – 2011	Teacher Assistant University of Cyprus, Nicosia, Cyprus
2006 – 2008	Callcenter Agent RAI Consultants, Nicosia, Cyprus

LANGUAGES

Greek	Fluent (Native language)
English	Fluent (Workplace language)
German	B2

PUBLICATIONS

Sarrou, M., Trapeau, R., Rübsamen, R., & Schönwiesner, M. Integration of dynamic binaural and spectral cues. (*Under preparation*)

Sarrou, M., Schmitz, P. M., Richter, N., & Rübsamen, R. (2018) Sound frequency affects the auditory motion-onset response in humans. Experimental Brain Research. https://doi.org/10.1007/s00221-018-5329-9

Avraamides, M. N., **Sarrou, M.**, & Kelly, J. W. (2013). Cross-sensory reference frame transfer in spatial memory: The case of proprioceptive learning. *Memory & Cognition* 42(3).

TEACHING EXPERIENCE

Lectures

Sarrou, M. (2011). *Visual and auditory pathways*. Introduction to Neurobiology, University of Cyprus, Nicosia, Cyprus.

Sarrou, M. (2013). *Auditory Motion Perception*. Practical course of Biology, University of Leipzig, Leipzig, Germany.

Sarrou, M. (2014). *Ethics in Scientific Research*. Practical course of Biology, University of Leipzig, Leipzig, Germany.

Practical Courses

4.2013	Practical course of Biology, University of Leipzig, Leipzig, Germany.
10.2013	Practical course of Biology, University of Leipzig, Leipzig, Germany.
4.2014	Practical course of Biology, University of Leipzig, Leipzig, Germany.
10.2014	Practical course of Biology, University of Leipzig, Leipzig, Germany.
4.2015	Practical course of Biology, University of Leipzig, Leipzig, Germany.
10.2015	Practical course of Biology, University of Leipzig, Leipzig, Germany.
4.2016	Practical course of Biology, University of Leipzig, Leipzig, Germany.
10.2017-2.2018	Experimental Psychology, University of Magdeburg, Magdeburg, Germany.

Seminars

4.2018-7.2018	General Psychology I/II, University of Magdeburg, Magdeburg, Germany.

STUDENT SUPERVISION

Schmitz, Pia, MSc (January 2014 - September 2014)
Hofmann, Felix, MSc (January 2013 - September 2013)

AWARDS

6.2016	Six-month scholarship Erasmus-Mundus on Auditory Cognitive Neuroscience
10.2012	Three-year PhD scholarship International Max Planck Research School for Neuroscience on Communication

POSTERS

Sarrou, M., Trapeau R., Rübsamen, R, & Schönwiesner, M. (2017). *Evidence of integration of dynamic binaural and spectral cues: a study on motion-direction adaptation.* Presented at the Final Erasmus-Mundus Symposium on Auditory Cognitive Neuroscience, Leipzig, Germany.

Sarrou, M., Richter, N., & Rübsamen, R. (2015). *Features of acoustic stimuli modulate late auditory responses during an Electroencephalogram.* Presented at ESCOP 2015, Paphos, Cyprus.

Sarrou, M., Richter, N., & Rübsamen, R. (2015). *Features of acoustic stimuli modulate late auditory responses during an Electroencephalogram: Preliminary data.* Presented at the 5. IMPRS NeuroCom Summer School, Leipzig, Germany.

Sarrou, M., Richter, N., & Rübsamen, R. (2015). *Modulation effects on the Change-Onset Response*. Presented at ARO 2015, Baltimore, Maryland, USA.

Sarrou, M., Richter, N., & Rübsamen, R. (2014). *Modulation effects on the auditory Change-Onset Response*. Presented at the 4. IMPRS NeuroCom Summer School, UCL, London, England.

CONFERENCES/WORKSHOPS

4.2017	Final Erasmus-Mundus Symposium on Auditory Cognitive Neuroscience, Leipzig, Germany
2-3.2016	Scientific Writing, RAL, Leipzig, Germany
3.2016	Self-management for the writing process, RAL, Leipzig, Germany
2.2016	Grant Proposal Writing, RAL, Leipzig, Germany
9.2015	19. Conference of the European Society for Cognitive Psychology, Paphos, Cyprus
9.2015	7. Mismatch Negativity Conference, Leipzig, Germany
7.2015	5. IMPRS NeuroCom Summer School, Leipzig, Germany
4.2015	Grant Proposal Writing H2020, Leipzig, Germany
2.2015	38. Annual MidWinter Meeting of the Association for Research in Otolaryngology, Baltimore, Maryland, USA
11.2014	Career Planning workshop, Leipzig, Germany
9.2014	Data visualization workshop: Well-formed information, Leipzig, Germany
4.2014	From the PhD-Odyssey to Project Management, Leipzig, Germany
4.2014	The Erasmus Mundus Symposium on Auditory Cognitive Neuroscience, Leipzig, Germany
2.2014	Symposium on cutting-edge methods in EEG research for the study of cognition, Berlin School of Mind and Brain, Berlin,

	Germany
7.2014	4. IMPRS NeuroCom Summer School, London, England
9.2013	EEG Analysis with EEGlab, Delmenhorst, Germany
9.2013	Intentional Inhibition: from motor suppression to self-control, UCL, London, England
7.2013	3. IMPRS NeuroCom Summer School, Leipzig, Germany

Πάντα ῥεῖ

Ηράκλειτος

Everything flows

Heracleitus

Acknowledgments

During the academic career noone walks alone; all important knowledge and scientific progress is built on the shoulders of giants. Therefore, I would like to thank the people and institutions that have helped me navigate in this journey.

First and foremost, I would like to thank my supervisor Prof. Dr. Rudolf Rübsamen. Thank you for giving me the opportunity to work with you and for supporting me until the end. Your motivation and trust has been invaluable to me.

I would like to thank Prof. Dr. Marc Schönwiesner for giving me the opportunity to conduct research at the BRAMS institute in Montreal and for supporting my doctoral work when you became professor at the University of Leipzig. You have an outstandingly positive attitude towards life and work that inspires me. I look forward to collaborating with you in the future as well.

My sincere thanks go to the two postdocs whose support and advice was the foundation of my experiments, Dr. Nicole Hamm from the University of Leipzig and Dr. Régis Trapeau from the BRAMS institute. You were both my guides and words cannot express how beneficial your help has been to me.

I am very grateful to the IMPRS NeuroCom programme of the Max-Planck-Institute for Human Cognitive and Brain Sciences. It has provided me with not only knowledge but also the opportunity to meet outstanding scientists by attending many events and conferences.

I would like to thank Prof. Dr. Erich Schröger for being in my committee and for generously offering your time and knowledge in reviewing my EEG data.

I would like to thank all the members of the Neurobiological lab of the University of Leipzig. The free-field group Fabian Schrumpf, Jan Bennemann, Pia Marlena Schmitz, Dr. Claudia Freigang, Marc Stöhr, Dr. Kristina Schmiedchen and Felix Hoffmann as well as the neurobiological group Dr. Tamara Radulovic, Dr. Christian Keine and Dr. Sasa Jovanovich thank you for many many meetings of me demanding your help and attention. Each and every one of you has helped me in several stages of my research. Our conversations, drinks and barbeques we had together have been imprinted in my mind and I will carry our memories with me forever.

I would like to give special thanks to Dr. Tamara Radulovic. Thank you for being not only an excellent colleague but also a great roommate. It goes without saying that our relationship shaped my experience in Leipzig and I would be forever greatful. Thank you for your support, advice and many much needed hugs when I needed them the most.

Last but not least, a big thank you to my family: my mum Elenitsa, my dad Antonis, my sister Eliana, my borther Marios, and of course Jim and Elizabeth. I would have never made it here if it wasn’t for you. You have encouradged me to go abroad and acquire as much knowledge as possible. You have motivated me to follow my passion and taught me to trust me gut. Thank you for believing in me and for being my compass when I felt like all hope was gone. Finally, thank you Christian. Thank you for making me laugh, for being my best friend, for supporting me and for having faith in me. It is not an accident that the last four years have been the most productive in work I have ever had. You are my favourite way to recharge and my life is much more interesting because of you.

Leipzig, 16th of July 2018

Contents

List of publications

Sarrou M, Schmitz PM, Hamm N, Rübsamen R (2018) Sound frequency affects the auditory motion-onset response in humans. Experimental Brain Research. Accepted.

Author contribution statement

Mikaella Sarrou

Auditory motion: perception and cortical response

Author contribution statement:

Title: Sound frequency affects the auditory motion-onset response in humans.

Journal: Experimental Brain Research

Authors: Sarrou M, Schmitz PM, Hamm N, Rübsamen R

Sarrou M (First author):

- Conceptualiuzation of research design
- Organization of experiments
- Data collection
- Analysis of results
- Writing of manuscript

Schmitz PM (Author 2):

- Data collection
- Contribution in writing of manuscript

Hamm N (Author 3):

- Conceptualization of research design
- Contribution in writing of manuscript

Rübsamen R (Senior author):

- Conceptualization of research design
- Contribution in writing of manuscript

Mikaella Sarrou

Rudolf Rübsamen

Bibliographische Darstellung

Mikaella Sarrou

Auditory motion: perception and cortical response

Fakultät für Lebenswissenschaften

Universität Leipzig

Dissertation

157 Seiten, 63 Literaturangaben, 38 Abbildungen, 8 Tabellen

The localization of sound sources in humans is based on the binaural cues, interaural time and level differences (ITDs, ILDs) and the spectral cues (Blauert 1997). The ITDs relate to the timing of sound arrival at the two ears. The ILDs refer to the difference of sound pressure-level between the two ears. The spectral cues are based on the shape of the pinna’s folds and they are useful for sound source localization in elevation, but they also help in azimuthal localization (Schnupp et al. 2012). The contribution of the spectral cues on the azimuthal localization arises from the fact that due to the symmetrical position of the ears on the head, the binaural cues vary symmetrically as a function of spatial location (King et al. 2001). The role of these cues in localizing sounds in our environment is well established. But their role in acoustic motion localization is not yet clear and the debate over whether these sensors exist is still ongoing. The current dissertation adds to the literature by supporting motion-specific mechanisms in the brain. My experiments demonstrated that a cortical response in humans elicited specifically from acoustic motion is modulated differently based on the frequency of the acoustic stimulus. These results were present even when the experimental paradigm used to elicit the motion response was altered. Further, the experiments showed that the faster the acoustic motion is, the stronger the cortical response to the motion gets. This is perhaps the sudden angular shifts in location from fast acoustic motion provide stronger motion evidence to the cortical system. When these evidence is abundant, then the cortical response to acoustic motion, is divided in two cumulative parts: one that represents the recognition and another that represents the evaluation of the motion. Lastly, a behavioural study using an adaptation paradigm, showed that the detection of motion-direction is based on the integration of interaural and spectral cues.

Abbreviations

AEP	Auditory evoked potentials
aMAE	Auditory motion aftereffect
AVCN	Anteroventral cochlear nucleus
BT	Brief time
CLat	Central to lateral
CLatC	Central to lateral and then to central
EEG	electroencephalogram
EOR	Energy-onset response
ERP	Event-related potential
fMRI	Functional magnetic resonance imaging
FV	Fast velocity
hEOG	Horizontal electrooculogram
HighF / hf	High frequency
HR	Hit rates
HRTF	Head-related transfer function
ILD	Interaural level difference
ISI	Interstimulus interval
ITD	Interaural time difference
LatC	Lateral to central
LatCLat	Lateral to central and then lateral
LH	Left hemisphere
LowF / lf	Low frequency
LSO	Lateral superior olive
LT	Long time
M	Means
MEG	Magnetoencephalography
MMN	Mismatch negativity
MNTB	Medial nucleus of the trapezoid body
MoMo	Moving-moving
MOR	Motion-onset response

MoSt	Moving-stationary
MSO	Medial superior olive
P2P	Peak-to-peak
PET	Positron-emission tomography
RH	Right hemisphere
rm	Repeated measures
RT	Reaction time
SCR	Spatial-change response
SEM	Standard-error of the mean
SL	Sensation level
SPL	Sound-pressure level
STG	Superior temporal gyrus
StMo	Stationary-moving
StSt	Stationary-stationary
SV	Slow velocity
vEOG	Vertical electrooculogram

Summary

The localization of sound sources in humans is based on the binaural cues, interaural time and level differences (ITDs, ILDs) and the spectral cues (Blauert 1997). The ITDs relate to the timing of sound arrival at the two ears. For example, a sound located at the right side will arrive at the right ear earlier than at the left ear. The ILDs refer to the difference of sound pressure-level between the two ears. In the example mentioned above, if the sound located at the right has short wavelength then it will arrive at the right ear with higher sound-pressure than at the left ear. This is because a sound with short wavelength cannot bypass the head. In other words, the head creates an obstacle that diffracts the waves and that is why the sound arriving at the ear closer to the sound source will receive the sound with higher sound-pressure. Due to the association of each of the binaural cues with the wavelength of a sound, Rayleigh (1907) proposed the 'duplex theory' of sound source localization suggesting that on the azimuth, the ITDs is the main localization cue for low frequency sounds and the ILDs is the main localization cue for high frequency sounds. The spectral cues are based on the shape of the pinna's folds and they are very useful for sound source localization in elevation but they also help in azimuthal localization (Schnupp et al. 2012). The contribution of the spectral cues on the azimuthal localization arises from the fact that due to the symmetrical position of the ears on the head, the binaural cues vary symmetrically as a function of spatial location (King et al. 2001). Whereas the ITDs have a very symmetrical distribution, the ILDs become more symmetrical the higher the sound frequency is. This way, there are certain locations within the left-frontal and left-posterior hemifield, as well as the right-frontal and the right-posterior hemifield that share the same binaural cues, which makes the binaural cues ambiguous and so the auditory system cannot depend solely on these for sound source localization. To resolve this ambiguity, our auditory system uses the spectral cues that help to disambiguate frontal-back confusion (King et al. 2001, Schnupp et al. 2012). The role of these cues in localizing sounds in our environment is well established. But their role in acoustic motion localization is not yet clear. This is the topic of the current thesis.

The auditory localization cues are processed on the subcortical and cortical level. The ITDs and ILDs are processed from different neurons along the auditory pathway (Schnupp et al. 2012). Their parallel processing stages seem to converge at the inferior colliculus as evidence shows from cat experiments (Chase and Young 2005). But in humans, an electroencephalographic (EEG) study measuring the mismatch negativity (MMN; Schröger

1996) and a study using magnetoencephalographie (MEG; Salminen et al. 2005) showed that these cues are not integrated.

One of the models of the spatial representation of sound sources is Jeffress' place code (1948). This model suggests that each location of the azimuthal space is encoded differently, thus the name 'place code'. Evidence in support of this model comes from studies on the cat (Yin and Chan 1990). However, arguments against this model come from studies in gerbils whose results showed that their subcortical neurons respond maximally to locations that are outside the physiological range based on the size of their heads (Pecka et al. 2008). An alternative model of auditory spatial encoding is the hemifield code (von Bekesy 1960). This model proposes that subcortical neurons are separated into two populations, one tuned to the left hemifield and another tuned to the right. Thus, the receptive field of the neurons is wide and the estimation of the sound source location is derived from the balance of activity of these two populations. Evidence from human studies support this model. Salminen and colleagues (2009) employed an adaptation paradigm during MEG recording. They presented sets of adaptor and probe stimuli that either had the same or different spatial location. Their results showed that the response to the probe was more reduced when the adaptor was located at the far left location and not when the adaptor and probe shared the exact same location. Also, an EEG study on auditory motion showed that sounds that move from central to lateral locations elicit higher amplitudes than when the move in the opposite direction (Magezi and Krumbholz 2010). The authors concluded that these results are based on the movement of the sound source towards the location of the maximal activity of the neurons (also in Salminen et al. 2012).

The ability to detect moving objects is well-embedded into our nature. Whereas it enriches predators and prey with the skills to survive, in everyday life it enables us to interact with our environment. For example, the task of crossing a street (without traffic signs) safely is based on the encoding of visual and auditory features of moving vehicles. In the visual modality, the capability of the system to encode motion is based on motion-specific neurons (Mather 2011). In the auditory modality, the debate over whether these sensors exist is still ongoing.

One theory on how the auditory system encodes motion is the ‘snapshot’ theory (Chandler and Grantham 1991, Grantham 1986). In a series of experiments, Grantham (1986) showed that auditory perception was not affected by features of motion such as velocity, but it was more sensitive on distance as a spatial cue. Thus, what he suggested is that the encoding of auditory motion is based on the mechanisms that encode stationary sounds. In other words, when a sound is moving it activates the neurons that correspond to the points that are located along the trajectory of that sound but in a serial manner. This way, the perception of auditory motion is based on ‘snapshots’ instead of processing motion as a complete feature. This mechanism of auditory motion processing corroborates with Jeffress’ place code (1948). Animal studies on monkeys (Ahissar et al. 1992) and owls (Wagner et al. 1994) showed that neurons responded similarly to moving and stationary sounds. Evidence against this theory come from a recent behavioural study that introduced velocity changes within acoustic motion and showed that participants were able to detect them (Locke et al. 2016). The authors concluded that if ‘snapshot’ theory would be true, then these detections of velocity change would not occur.

Another theory of auditory motion was evolved that supports the motion-specific mechanisms in the brain (Warren et al. 2002, Docummun et al. 2004, Poirier et al. 20017). A human study using functional magnetic resonance imaging (fMRI) and positron-emission tomography (PET) showed evidence of a motion-specific cortical network that includes the planum temporale and the parietotemporal operculum (Warren et al. 2002). The authors suggested that these areas are part of a posterior processing stream that is responsible for analysis of auditory moving objects. Moreover, a recent primate fMRI study provided evidence of motion-specificity in the activity of the posterior belt and parabelt regions of the primary auditory cortex (Poirier et al. 2017). The authors contrasted cortical response to auditory motion with stationary and spectrotemporal sounds and found that the aforementioned cortical areas were only activated by moving sounds.

All in all, the neuronal mechanism underlying auditory motion perception has been vaguely described. However, there is an increasing number of evidence that show that specialized motion areas and mechanisms exist in the cortex. To study how exactly these mechanisms function, it is important to know which aspects of the stimulus paradigm affect the response.

Study 1. In this study, I focused on eliciting the cortical motion-onset response (MOR) in the freefield. This specific response is measured with EEG and it is elicited when a sound motion follows a stationary sound without any temporal gaps between them. The stationary part serves as an adaptive sound and the onset of motion provides a release-of-adaptation, which gives rise to the MOR. One of the focus was to investigate the effect on the MOR when the initial part is moving in space instead of being stationary. In addition, a secondary focus was the effect of the stimuli frequency on the MOR. I hypothesized that, due to the adaptation provided by the initial stimulus part, the motion response would be smaller after moving than after stationary adaptation. Also, I expected that the effects of frequency would follow the literature and since the motion response is a late response, the amplitude would be smaller after the high frequency than low frequency stimulus presentation. The results showed that the current paradigm did not elicit the MOR. Comparison of the current experimental settings with those used previously in the literature showed that the MOR is strongly depended on the adaptation time provided by the first part of the stimuli.

Study 2. In this study, the stimulus characteristics were adapted after failing to elicit the response in the previous study. In addition, I employed an active instead of a passive paradigm, since data from the literature show that the motion response is strongly dependent on the allocation of attention on auditory motion. Thus, in this study, the elicitation of the MOR was successful. The current study examines the modulation of the MOR based on the frequency-range of sound stimuli. Higher amplitude on the motion response was expected after the presentation of stimuli with high frequency spectrum. Also, I studied the effects of hemifield presentation and the direction of motion on the MOR. The results showed that the early part of the motion response (cN1) was modulated by the frequency range of the sounds with stronger amplitudes elicited by stimuli with high frequency range.

Study 3. This study is focused on analysis from data collected in the previous study. The focus, however, is on the effects of the stimulus paradigm on the MOR. I hypothesized that after the adaptation provided by an initial moving part, lower amplitude was expected in comparison to the stimuli with an initial stationary part. These responses were also analysed based on the effects of stimulus frequency. The results showed that the stimulus paradigm with the initial moving part elicited a response that resembles the MOR but has lower amplitude. In addition, the effects of stimulus frequency evident from the previous analysis apply here as well, with high frequency stimuli eliciting higher MOR amplitude than low frequency stimuli.

Study 4. This study examined further the effects of stimuli characteristics on the MOR. Since the latency of the MOR in the previous study was a bit later than what is usually reported in the literature, the focus here was to test the effects of motion velocity and adaptation duration on the MOR. The results showed that faster velocity elicited higher amplitudes on the peak-to-peak comparison. Separate analysis on the MOR components, showed that this effect was based on higher cN1 amplitude. A separate analysis between the electrodes over the left and right hemisphere, showed that the peak-to-peak amplitude was stronger on the electrodes over the right hemisphere. Lastly, the strong adaptation created by the long duration of the initial stationary part provided abundant evidence of auditory motion, which led to the separation of the cP2 into its constituent parts.

Study 5. This behavioural study focused on the effect of motion adaptation on the rear field to the presentation of motion in the frontal field. Thus, the presentation of adaptors and probes within the left-frontal and left-rear fields aimed at locations that share the same ITDs and ILDs. The disambiguation of auditory localization of motion is based on how these interaural cues interact with the spectral cues. A moving probe was presented in the left hemifield, following an adaptor that spanned either the same trajectory or a trajectory located in the opposite field (frontal/ rear). Participants had to indicate the direction of the probe. The results showed that performance was worse when adaptor and probe were sharing the same binaural cues, even if they were in different hemifields and their direction was opposite. But the magnitude of the adaptation effect when the pair was in different hemifields was smaller, thus showing motion-direction detection depends on the integration of interaural and spectral cues.

Zusammenfassung

Die Lokalisierung von Geräuschquellen beim Menschen basiert auf binauralen Merkmalen, Laufzeitdifferenzen (ITDs) und Pegeldifferenzen (ILDs) sowie spektralen Merkmalen (Blauert 1997). Die ITDs beziehen sich auf den Zeitpunkt der Ankunft des Tones an den Ohren. So wird zum Beispiel ein Geräusch, das sich auf der rechten Seite des Hörers befindet, das rechte Ohr früher erreichen als das linke. Die ILDs beziehen sich auf den Unterschied der Schalldruckintensität zwischen den Ohren. So können Geräusche mit kurzer Wellenlänge den Kopf nicht durchdringen. Mit anderen Worten stellt der Kopf ein Hindernis dar, das Schalldruckwellen ablenkt, weshalb das näher an der Geräuschquelle liegende Ohr stärkeren Schalldruck empfängt. Basierend auf der Tatsache, dass jedes binaurale Merkmal auch einer bestimmten Wellenlänge zugeordnet werden kann schlug Rayleigh (1907) die „duplex theory" zur Lokalisierung von Geräuschquellen vor. Dieser Theorie zufolge dienen die ITDs als Hauptmerkmale zur Lokalisierung von Niedrigfrequenztönen auf dem Azimut und die ILDs zur Lokalisierung von Hochfrequenztönen. Spektrale Merkmale basieren auf der Form der Ohrmuschel und sind in erster Linie hilfreich für die Höhenlokalisierung von Geräuschquellen, sind zudem aber auch bei der azimutalen Lokalisierung von Nutzen (Schnupp et al. 2012). Die Beteiligung der spektralen Merkmale bei der azimutalen Lokalisierung beruht auf der Tatsache, dass durch die symmetrische Anordnung der Ohren die binauralen Merkmale in Abhängigkeit der räumlichen Verortung der Schallquelle variieren. (King et al. 2001). Während die ITDs eine sehr symmetrische Verteilung aufweisen werden die ILDs hingegen zunehmend symmetrischer je höher die Frequenz des Tones wird. Daher gibt es bestimmte Positionen innerhalb des links-frontalen und links-posterioren (und dementsprechend auch auf der rechten Seite) Halbfeldes in denen die gleichen binauralen Merkmale vorhanden sind. Somit sind binaurale Merkmale mehrdeutig, weshalb das auditorische System bei der Lokalisierung von Geräuschquellen nicht ausschließlich auf diese zurückgreifen kann. Um diese Mehrdeutigkeit aufzulösen, verwendet das auditorische System zusätzlich spektrale Merkmale (King et al. 2001, Schnupp et al. 2012). Die Rolle dieser Merkmale bei der Geräuschlokalisierung in unserer Umwelt ist bekannt. Ihre Rolle bei der Lokalisierung von akustischer Bewegung hingegen ist unklar. Das ist das Thema der vorliegenden Arbeit.

Die auditorischen Merkmale werden auf subkortikaler und kortikaler Ebene verarbeitet. Die ITDs und ILDs werden von verschiedenen Neuronen entlang des auditorischen Pfades verarbeitet (Schnupp et al. 2012). Basierend auf Experimenten mit Katzen scheinen diese parallelen Verarbeitungsstufen im inferioren Colliculus zusammenzulaufen (Chase und Young 2005). Beim Menschen jedoch scheinen Befunde anhand von Elektroenzephalographie (EEG) zur sogenannten mismatch negativity (MMN; Schröger 1996) diese Annahme nicht zu bestätigen. Zusätzlich zeigte eine Studie unter Verwendung von Magnetenzephalographie (MEG; Salminen et al. 2005), dass diese Merkmale nicht integriert werden.

Eines der Modelle zur räumlichen Repräsentation von Geräuschquellen ist Jeffress‘ place code (1948). Dieses Modell behauptet, dass jede Position im azimutalen Raum unterschiedlich kodiert ist. Daher auch der Name „place code“. Evidenz zur Unterstützung dieses Modelles leistet eine Studie an Katzen (Yin und Chan 1990). Dennoch gibt es Argumente gegen dieses Modell basierend auf Studien an Gerbils. Diese Studien zeigen, dass ihre subkortikalen Neuronen am stärksten auf Positionen reagieren, die außerhalb der physiologischen Bandbreite basierend auf der Größe ihres Kopfes liegen (Pecka et al. 2008). Ein alternatives Model zur auditorisch-räumlichen Kodierung ist der sogenannte hemifield code (Bekesy 1960). Dieses Modell geht davon aus, dass die subkortikalen Neuronen in zwei Gruppen aufgeteilt sind. Eine Gruppe reagiert auf Geräusche innerhalb des linken Halbfeldes, die andere auf Geräusche des rechten Halbfeldes. Das rezeptive Feld dieser Neuronen ist weitreichend und die Bestimmung von Geräuschquellen leitet sich aus der Balance zwischen der Aktivität dieser zwei Gruppen ab. Studien am Menschen bieten Evidenz für diese Annahme. Salminen und Kollegen (2009) verwendeten ein MEG-Adaptionsparadigma. Sie präsentierten Versuchspersonen Paarevon verschiedenen Reizen, die entweder identische oder unterschiedliche Positionen hatten. Die Ergebnisse zeigten, dass das MEG-Signal in Antwort auf den Testreiz niedriger war, wenn der Adapterreiz sich links-außen befand und wenn sich der Adapterreiz und der Testreiz nicht an derselben Position befanden. Eine weitere EEG-Studie zur auditorischen Bewegung zeigt, dass sich von zentral nach lateral bewegende Geräusche eine höhere Amplitude hervorrufen, als Geräusche, die sich in die entgegengesetzte Richtung bewegen (Magezi und Krumbholz 2010). Die Autoren schlossen hieraus, dass diese Ergebnisse auf der Bewegung von Geräuschquellen zur Position der maximalen Aktivität der Neuronen beruhen (so auch Salminen et al. 2012).

Die Fähigkeit bewegte Objekte zur erkennen ist in uns veranlagt. Sie verschafft Räubern und Beutetieren die Fähigkeit zu überleben. Im Alltag ermöglicht uns diese Fähigkeit mit unserer Umwelt zu interagieren. Beispielsweise basiert die Aufgabe eine Straße sicher zu überqueren auf der erfolgreichen Kodierung von visuellen und auditorischen Merkmalen sich bewegender

Fahrzeuge. Im visuellen System gibt es bestimmte Neurone, die ausschließlich bewegte Reize kodieren (Mather 2011). Im auditorischen System hingegen ist die Existenz solcher Neurone weiterhin umstritten.

Eine Theorie, wie das auditorische System Bewegung kodiert, ist die sogenannte „snapshot“ Theorie (Chandler und Grantham 1991, Grantham 1986). In einer Serie von Experimenten zeigte Grantham (1986), dass auditorische Wahrnehmung nicht von bestimmten Bewegungsmerkmalen, wie etwa Schnelligkeit, beeinflusst wird, sondern vielmehr von der Distanz als räumliches Merkmal. Hieraus zieht Grantham den Schluss, dass die Kodierung von auditorischer Bewegung auf denselben Mechanismen wie die Kodierung unbewegter Geräusche basiert. In anderen Worten, wenn sich ein Geräusch bewegt aktiviert es die korrespondierenden Neuronen in einer Sequenz entlang des Bahnverlaufes. Die Wahrnehmung von auditorischer Bewegung beruht also auf „Schnappschüssen“ (engl. snapshots) anstatt einer durchgehenden Verarbeitung der Bewegung. Dieser Mechanismus der auditorischen Bewegungsverarbeitung bekräftigt Jeffress‘ place code (1948). Tierstudien an Affen (Ahissar et al. 1992) sowie Eulen (Wagner et al. 1994) zeigen, dass es Neuronen gibt, die eine ähnliche Aktivierung bei unbewegten und bewegenden Geräuschen zeigen. Befunde gegen diese Theorie hingegen stammen aus einer Verhaltensstudie, die demonstrierte, dass Probanden Geschwindigkeitsveränderungen im Verlauf einer auditorischen Bewegung entdecken konnten (Locke et al., 2016). Die Autoren argumentieren hierzu, dass Probanden die Änderung der Geschwindigkeit nicht hätten feststellen können, wenn die snapshot Theorie zutreffend wäre.

Eine weitere Theorie zur auditorischen Bewegung geht davon aus, dass es bewegungsspezifische Mechanismen gibt (Warren et al. 2002, Docummun et al. 2004, Poirier et al. 20017). Eine Studie am Menschen unter Verwendung von funktioneller Magnetresonanztomographie (fMRT) und Positronen-Emissions-Tomographie (PET) liefert Beweise eines bewegungsspezifischen kortikalen Netzwerkes. Zwei Bestandteile dieses Netzwerkes sind das planum temporale und das parietotemporale operculum (Warren et al. 2002). Die Autoren kommen zu dem Schluss, dass diese Bereiche Teil eines posterioren Verarbeitungsflusses sind, der für die Analyse von sich bewegenden Objekten verantwortlich ist. Zudem zeigte kürzlich eine fMRT-Studie an Primaten, dass sich in den posterior belt und parabelt Regionen bewegungsspezifische Aktivität beobachten lassen (Poirier et al. 2017). Die Autoren vergleichen die kortikale Aktivität bei auditorischer Bewegung mit unbewegten und spektrotemporalen Geräuschen. Sie fanden heraus, dass die vorgenannten kortikalen Regionen nur durch sich bewegende Geräusche aktiviert werden.

Zusammenfassend kann gesagt werden, dass die neuronalen Mechanismen, die die Wahrnehmung von auditorischer Bewegung unterstützen, bislang nur sehr vage beschrieben wurden. Dennoch gibt es eine steigende Zahl von Befunden, die auf die Existenz bewegungsspezifischer Areale im Kortex hinweisen. Um mehr über die Funktionsweise dieser Mechanismen in Erfahrung zu bringen, ist es wichtig zu verstehen wie einzelne Stimulusmerkmale das durch auditorische Bewegung ausgelöste EEG-Signal beeinflussen.

Studie 1. In dieser Studie habe ich mich darauf konzentriert, die sogenannte kortikale motion-onset response (MOR) im Freifeldlabor herauszuarbeiten. Die spezifische Reaktion wird mittels Elektroenzephalogramm erfasst und es wird ausgelöst, wenn auf ein unbewegtes Geräusch unmittelbar im Anschluss ein sich bewegendes Geräusch folgt. Der unbewegte Teil des Geräusches dient als ein adaptiver Reiz und der Beginn der Bewegung schafft eine sogenannte release-of-adaptation, was die MOR auslöst. Ein Schwerpunkt lag auf der Untersuchung von Veränderungen in der MOR, wenn der erste Teil des Reizes nicht ortsgebunden sondern beweglich ist. Zusätzlich lag ein zweiter Schwerpunkt auf der Beeinflussung der MOR durch die Frequenz der Stimuli. Die Hypothese hierzu lautete, dass aufgrund der Adaption auf den ersten Teiles des Reizes, die Reaktion auf Bewegung im zweiten Teil geringer ist, als wenn der erste Teil des Reizes unbewegt ist. Zudem wurde erwartet, dass die Effekte der Frequenz denen aus der bestehenden Literatur entsprechen. Die Amplitude sollte demzufolge geringer sein, wenn der Stimulus hochfrequent und nicht niederfrequent ist. Die Ergebnisse zeigten, dass das hier verwendete Paradigma keine MOR auslöst. Ein Vergleich der Ergebnisse dieser Studie mit den Ergebnissen bestehender Studien in der Literatur zeigt, dass die MOR stark von der Adaptionszeit des ersten Teiles des Stimulus abhängig ist.

Studie 2. In dieser Studie wurden die Merkmale des Stimulus angepasst nachdem die vorherige Studie nicht dazu führte, dass die MOR ausgelöst wurde. Zusätzlich wurde ein aktives anstatt eines passiven Paradigmas genutzt, da Ergebnisse aus der Literatur zeigten, dass die Reaktion auf Bewegung stark von der Aufmerksamkeit des Probanden auf die auditorische Bewegung abhängt. In dieser Studie konnte eine MOR beobachtet werden. Diese Studie untersucht die

Modulation der MOR basierend auf der Frequenz des Reizes. Eine höhere Amplitude der MOR wurde erwartet, nachdem hochfrequente Reize präsentiert wurden. Zudem wurden die Einflüsse der Halbfeldpräsentation und der Bewegungsrichtung auf die MOR untersucht. Die Ergebnisse zeigen, dass der erste Teil der MOR (cN1) durch verschiedene Frequenzen moduliert wurde. Reize mit hoher Frequenz führen zu einer starken Amplitude der cN1.

Studie 3. Diese Studie fokussiert sich auf weitergehende Analysen der Daten aus den vorherigen Studien. Der Schwerpunkt lag jedoch auf dem Einfluss des Stimulusparadigmas auf die MOR. Die Hypothese hierzu lautete, dass die Adaption des ersten Teiles des sich bewegenden Stimulus zu einer niedrigeren Amplitude führt, als wenn der erste Teil des Stimulus ortsgebunden ist. Die MOR wird außerdem in Bezug auf Frequenzänderungen untersucht. Die Ergebnisse zeigen, dass das Stimulusparadigma mit einem bewegenden ersten Teil dazu führt, dass eine Reaktion ausgelöst wird, die der MOR ähnelt aber eine geringere Amplitude hat. Zusätzlich bestätigten sich die Ergebnisse der vorherigen Studie dahingehend, dass hochfrequente Reize zu einer höheren MOR Amplitude führen als niederfrequente Reize.

Studie 4. Diese Studie untersuchte weitergehend die Effekte der Reizmerkmale auf die MOR. Nachdem die Latenz der MOR in der vorangegangenen Studie etwas länger war als in der Literatur beschrieben, lag der Fokus hier auf den Effekten der Bewegungsgeschwindigkeit und Adaptionsdauer auf die MOR. Die Ergebnisse zeigten, dass Reize mit hoher Geschwindigkeit eine höhere MOR-Amplitude auslösen. Eine weitere Analyse der MOR-Komponenten zeigte, dass dieser Effekt auf einer höheren cN1-Amplitude basiert. Eine vergleichende Analyse zwischen den Elektroden der linken und der rechten Hemisphäre zeigte, dass die Amplitude für die Elektroden der rechten Hemisphäre stärker ausgeprägt war. Letztlich zeigte sich, dass eine starke Adaption, hervorgerufen durch eine lange Dauer des ersten, ruhenden Teiles des Reizes, dazu führt, dass sich die cP2 Komponente in zwei Abschnitte aufteilt.

Studie 5. Diese Verhaltensstudie fokussierte sich auf den Effekt der Bewegungsadaption des hinteren Halbfeldes mit anschließender Präsentation von bewegenden Reiz im vorderen Halbfeld. Dazu wurden Adapter- und Testreize im linken frontalen und linken posterioren Feld an Positionen präsentiert, die gleiche ITDs und ILDs aufweisen. Die Lokalisierung von auditorischer Bewegung basiert auf der Interaktion zwischen interauralen Merkmalen und spektralen Merkmalen. Im Anschluss an einen bewegenden Adapterreiz im linken Halbfeld wurde ein Testreiz präsentiert, der sich entweder auf der gleichen Position des Bahnverlaufes oder auf dem entgegengesetzten Punkt des Bahnverlaufes befindet (frontal/posterior). Die Probanden mussten die Richtung des Testreizes identifizieren. Die Resultate zeigten, dass die Leistung der Probanden schlechter war wenn der Adapter- und Testreiz die gleichen binauralen Merkmale aufwiesen, auch wenn sich diese in verschiedenen Halbfeldern befanden und ihre Bewegungsrichtung entgegengesetzt war. Das Ausmaß des Adaptionseffekts war geringer, wenn sich Adapter- und Testreize in unterschiedlichen Halbfeldern befanden. Dies weist darauf hin, dass das Erkennen der Bewegungsrichtung von der Integration von interauralen und spektralen Merkmalen abhängt.

Chapter 1. Introduction

1.1 Spatial hearing: from the environment to the cortex

The localization of sound sources in humans is based on the binaural cues, interaural time and level differences (ITDs, ILDs) and the spectral cues (Blauert 1997) (Fig. 1). The former are based on the location of ears at the sides of the head and the latter on the shape of the pinna. The ITDs relate to the timing of sound arrival at the two ears. For example, a sound located at the right side will arrive at the right ear earlier than at the left ear. The ILDs refer to the difference of sound pressure-level between the two ears. In the example mentioned above, if the sound located at the right has short wavelength then it will arrive at the right ear with higher sound-pressure than at the left ear. This is because a sound with short wavelength cannot bypass the head. In other words, the head creates an obstacle that diffracts the waves and that is why the sound arriving at the ear closer to the sound source will receive the sound with higher sound-pressure. Due to the association of each of the binaural cues with the wavelength of a sound, Rayleigh (1907) proposed the 'duplex theory' of sound source localization suggesting that on the azimuth, the ITDs is the main localization cue for low frequency sounds and the ILDs is the main localization cue for high frequency sounds. The spectral cues are based on the shape of the pinna's folds and they are very useful for sound source localization in elevation but they also help in azimuthal localization (Schnupp et al. 2012). The contribution of the spectral cues on the azimuthal localization arises from the fact that due to the symmetrical position of the ears on the head, the binaural cues vary symmetrically as a function of spatial location (King et al. 2001). Whereas the ITDs have a very symmetrical distribution, the ILDs become more symmetrical the higher the sound frequency is. This way, there are certain locations within the left-frontal and left-posterior hemifield, as well as the right-frontal and the right-posterior hemifield that share the same binaural cues, which makes the binaural cues ambiguous and so the auditory system cannot depend solely on these for sound source localization. In order to resolve this ambiguity, our auditory system uses the spectral cues that help to disambiguate frontal-back confusion (King et al. 2001, Schnupp et al. 2012).

The localization cues of the auditory system are processed on the subcortical and cortical level (Fig. 1). The ITDs are processed in the neurons of the medial superior olive (MSO) of the brainstem, which receive excitatory inputs from spherical bushy cells located at the ipsilateral and contralateral anteroventral cochlear nucleus (AVCN) (Schnupp et al. 2012). The ILDs are processed in the neurons of the lateral superior olive (LSO), which receive ipsilateral excitatory input from the bushy cells in AVCN. Also, they receive inhibitory glycinergic input from the ipsilateral medial nucleus of the trapezoid body (MNTB), which received excitatory input from globular bushy cells from the contralateral AVCN (Schnupp et al. 2012). These parallel processing stages seem to converge at the inferior colliculus as evidence shows from cat experiments (Chase and Young 2005). But in humans, an electroencephalographic (EEG) study measuring the mismatch negativity (MMN) showed that these cues are not integrated (Schröger 1996). This study’s results showed that the MMN to deviants that had both ITD and ILD characteristics were equal to the sum of the MMN resulting from deviants that had only-ITD and only-ILD. In addition, a recent magnetoencephalographic (MEG) study applied an adaptation paradigm during which the adaptors and probes had either ITD or ILD characteristics (Salminen et al. 2005). Their results showed that the adaptation was location-specific and not localization cue-specific. The authors concluded that the binaural cues remain separated at the cortical level.

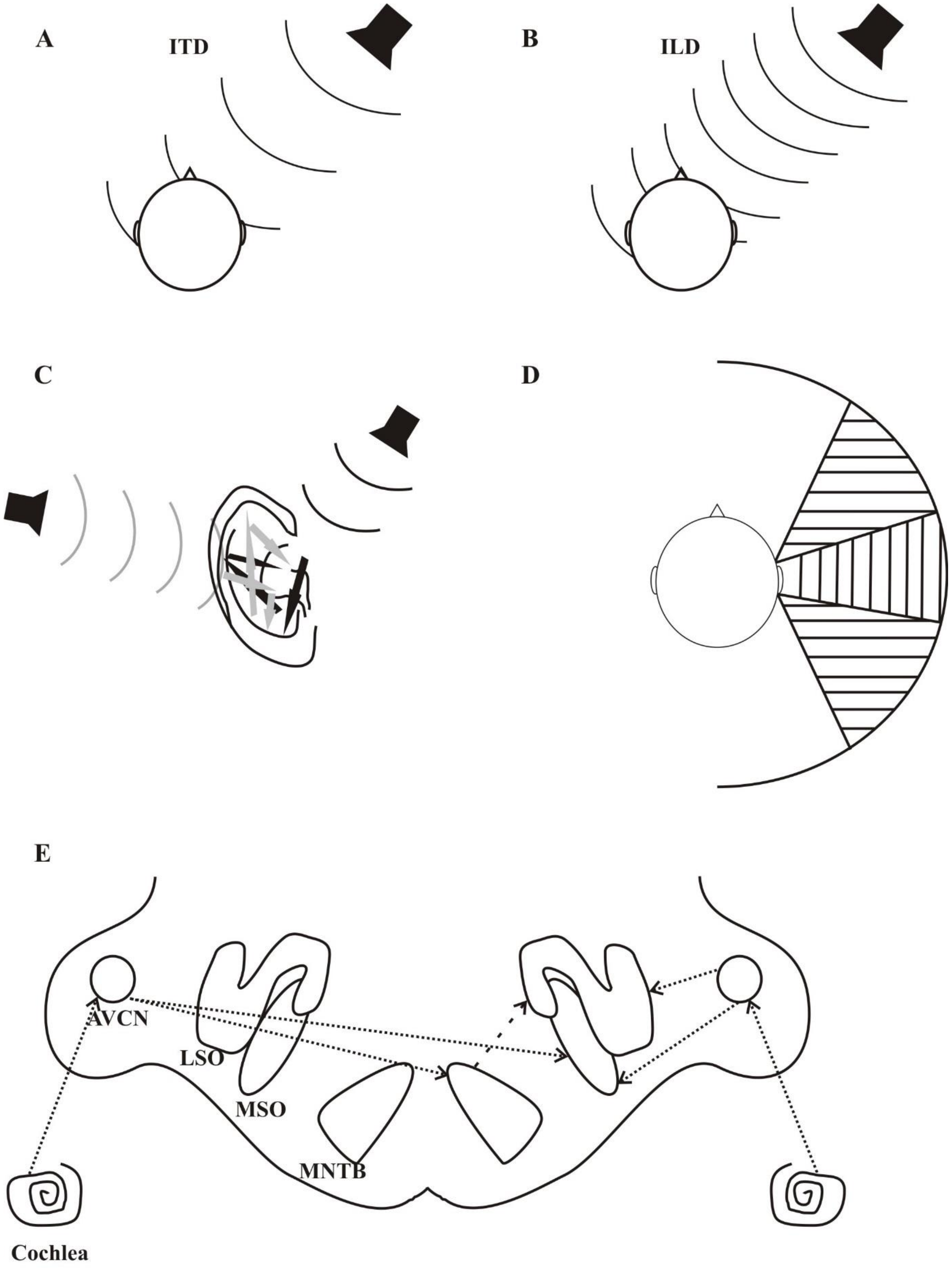


Fig. 1 Graphic representation of the auditory localization cues used humans use and the processing pathway they follow in the brainstem. **(A)** The interaural time differences are the basis for the processing of low frequency sounds because the wavelength of the sounds is large and can bypass the head. Thus, the sound will

arrive earlier at the right ear than the left. **(B)** The interaural level differences are the basis for the processing of high frequency stimuli because the wavelength of the sounds is short and the head creates an obstacle that deflects the sound waves. Thus, the sound will arrive weaker (lower amplitude) at the right ear than the left. **(C)** The spectral cues are the basis for localization of sounds in elevation (front speaker and black arrows) and they are used to disambiguate locations on the azimuth that have the same ITDs/ILDs (posterior speaker and grey arrows). **(D)** Locations on the azimuth between the frontal and rear hemifields share the same ITD and ILD cues. Here, the horizontal lines how some of these locations. **(E)** If the sound that arrives at the right ear carries mainly ITDs (like in A), then the information moves from the ipsilateral and contralateral AVCN to the ipsilateral MSO with excitatory projections. If the sound that arrives at the right ear carries mainly ITDs (like in A), then the information moves from the ipsilateral AVCN to the LSO with excitatory projections and from the contralateral AVCN with excitatory projections to the ipsilateral MNTB and then, from the ipsilateral MNTB to the ipsilateral LSO with inhibitory projections. Dotted arrows show excitatory projections. Dashed arrows show inhibitory projections. ITD: interaural time differences, ILD: interaural level differences, AVCN: anteroventral cochlea nucleus, LSO: lateral superior olive, MSO: medial superior olive, MTNB: medial nucleus of the trapezoid body.

Moreover, a growing number of evidence shows that the representation of auditory localization cues is better in one cortical hemisphere than the other (Krumbholz et al. 2005; Krumbholz et al. 2007; Palomäki et al. 2005; Zatorre and Penhune 2001). Palomäki and colleagues (2005) used MEG to study the processing of different types of sounds and found that the N1m was stronger at the right hemisphere as a response to sounds that contained stronger spatial cues. The authors concluded that the right hemisphere is dominant for auditory spatial processing. Also, lesion studies (Zatorre and Penhune 2001) found that impairment of the right auditory cortex impacts sound source localization on both hemifields, while impairment of the left auditory cortex has very low impact on sound source localization. Moreover, an auditory motion study using EEG, found that there is a pattern of contralaterality in the signal based on the motion direction, which is stronger for leftward than rightward motion (Krumbholz et al. 2007). In addition, a study using functional magnetic resonance imaging (fMRI) showed that the right hemisphere is activated from sounds moving within both hemifields, while the left hemisphere is activated from sounds moving within the right hemifield (Krumbholz et al. 2005).

1.2 Models of auditory spatial representation

One of the models of the spatial representation of sound sources is Jeffress' place code (1948). Jeffress provided a theory on how the MSO is processing the ITDs. He suggested that the ITDs are encoded in the temporal structure of the action potentials that leave from the left and right

AVCN to reach the MSO. The MSO neurons act as coincidence detectors comparing the time of arrival of the action potentials from the left and right ears. This model suggests that each ITD will elicit difference activity of the MSO neurons; in other words, each location of the auditory azimuthal space is encoded differently, thus the name ‘place code’. Evidence in support of this model comes from studies on the cat (Yin and Chan 1990) that show results consistent with what the theory proposes. However, arguments against this model come from studies in gerbils whose results showed that their MSO neurons response maximally to locations that are outside the physiological range based on the size of their heads (Pecka et al. 2008). In addition, the location of the MSO in the brainstem makes it difficult to approach it to study its physiology (Schnupp et al. 2012).

An alternative model of auditory spatial encoding is the hemifield code (von Bekesy 1960). This model proposes an activity pattern from the MSO neurons that separates them into two populations, one tuned to the left hemifield and another tuned to the right. Thus, the receptive field of the neurons is wider and the estimation of the sound source location is derived from the balance of activity of these two populations in the MSO. In this way, the activity of the neurons is maximal to locations coinciding to the maximal ITD/ILD, meaning the far left or far right azimuthal locations. Animal studies that showed that the neurons in MSO are maximally activated by ITDs that are disproportional to the size of the animal’s head, like the gerbil study mentioned previously, support this model (Pecka et al. 2008). Moreover, evidence from human studies support this model. Salminen and colleagues (2009) employed an adaptation paradigm during MEG recording. They presented sets of adaptor and probe stimuli that either had the same or different spatial location. The idea behind this paradigm is that the response to the probe will be more reduced, the more similar the characteristics of the adaptor and probe are. However, their results showed that the response to the probe was more reduced when the adaptor was located at the far left location and not when the adaptor and probe shared the exact same location. Furthermore, an EEG study on auditory motion showed that sounds that move from central to lateral locations elicit higher amplitudes than when the move in the opposite direction (Magezi and Krumbholz 2010). The authors concluded that these results are based on the movement of the sound source towards the location of the maximal activity of the neurons (also in Salminen et al. 2012).

All in all, research so far provides stronger evidence from human studies in support of the hemifield code.

1.3 Acoustic motion

The ability to detect moving objects is well-embedded into our nature. Whereas it enriches predators and prey with the skills to survive, in everyday life it enables us to interact with our environment. For example, the task of crossing a street (without traffic signs) safely is based on the encoding of visual and auditory features of moving vehicles. In the visual modality, the capability of the system to encode motion is based on motion-specific neurons (Mather 2011). In the auditory modality, the debate over whether these sensors exist is still ongoing.

One theory on how the auditory system encodes motion is the ‘snapshot’ theory (Chandler and Grantham 1991, Grantham 1986). In a series of experiments, Grantham (1986) showed that auditory perception was not affected by features of motion such as velocity, but it was more sensitive on distance as a spatial cue. Thus, what he suggested is that the encoding of auditory motion is based on the mechanisms that encode stationary sounds. In other words, when a sound is moving it activates the neurons that correspond to the points that are located along the trajectory of that sound but in a serial manner. Thus, the brain calculates the sound’s different locations during its duration (Middlebrooks and Green 1991). This way, the perception of auditory motion is based on ‘snapshots’ instead of processing motion as a complete feature. This mechanism of auditory motion processing corroborates with Jeffress’ place code (1948) for representing the auditory space in the cortex, which was described in the previous section. An animal study employing single-unit recordings on monkeys showed that neurons responded similarly to moving and stationary sounds (Ahissar et al. 1992) and similar evidence were derived from a study with owls (Wagner et al. 1994). Despite these results, both aforementioned studies recorded from neurons that were direction selective. Although both studies refer to these neurons as ‘motion-direction-sensitive’, their results could be explained as contralaterality selectivity from neurons and not direction. Evidence against this theory come from a recent behavioural study that introduced velocity changes within acoustic motion and showed that participants were able to detect them (Locke et al. 2016). The authors concluded that if ‘snapshot’ theory would be true, then these detections of velocity change would not occur.

Due to the issues of ‘snapshot’ theory finding support, another theory of auditory motion was evolved that supports the motion-specific mechanisms in the brain. Support for this theory comes from several recent studies. A human study using fMRI and positron-emission tomography (PET) showed evidence of a motion-specific cortical network that includes the planum temporale and the parietotemporal operculum (Warren et al. 2002). The authors suggested that these areas are part of a posterior processing stream that is responsible for analysis of auditory moving objects. Moreover, a recent primate fMRI study provided evidence of motion-specificity in the activity of the posterior belt and parabelt regions of the primary auditory cortex (Poirier et al. 2017). The authors contrasted cortical response to auditory motion with stationary and spectrotemporal sounds and found that the aforementioned cortical areas were only activated by moving sounds. Also, in support of the motion-specific mechanisms of the auditory system is a lesion study showing that a patient was specifically unable to detect auditory motion following surgery on the right anterior temporal lobe and the superior temporal gyrus (STG) (Ducommun et al. 2004).

Chapter 2. Methodology

2.1 Free-field stimulus presentation

Free-field laboratories offer the possibility of presenting sounds in experimental sessions with natural percept (Getzmann and Lewald 2010). This is a big advantage in comparison to other methods for sound presentation such as the employment of interaural differences (ITD and ILD) and head-related transfer function (HRTF) because the two former methods create an intracranial sense for sound source position. In the current thesis, all experiments were carried out in free-field laboratories. This gave the opportunity for studying the ITDs and ILDs as cues that are used for localization of different sound frequencies (low and high frequency, respectively; Chapter 3, Chapter 4, Chapter 5).

Free-field laboratory at the Department of Neurobiology, University of Leipzig, Leipzig, Germany

This laboratory is a semi-anechoic chamber (40 m^2; Industrial Acoustics Company, Niederkrüchten, Germany) with foam wedges attenuating sound waves and thus controlling echoes (Fig. 1). In the middle of this room, forty-seven loudspeakers (Visaton, FRS8 4 Ohm, Haan, Germany) are arranged in an azimuthal semi-circular plane spanning from 98° left to 98° right in the frontal hemifield, with 2.35 m radius. A comfortable chair is placed in the middle of this trajectory for the participants to sit during the experiments. The angular distance between two loudspeakers is 4.3°. Each loudspeaker was equilibrated individually. For this, the transmission spectrum was measured using the a Bruel & Kjaer measuring amplifier (B&K 2610), a microphone (B&K 2669, pre-amplifier B&K 4190) and a real-time signal processor (RP 2.1, System3, Tucker Davis Technologies, TDT, Alachua, Florida, USA). For each loudspeaker, a calibration file was generated in Matlab 6.1 (The MathWorks Inc., Natick, Massachusetts, USA) and was later used to ensure flat spectra of the acoustic stimuli across the frequency range tested. The loudspeakers are covered with a sound transparent black curtain to prevent participants having visual location information of the loudspeakers' positions. Stimuli were generated digitally using real-time TDT processors (RX8, System3) and controlled with custom-made scripts in MATLAB 7.5 (R2007b, The MathWorks Inc., Natick, Massachusetts, USA). An infrared camera was installed in this chamber to oversee participants during testing.

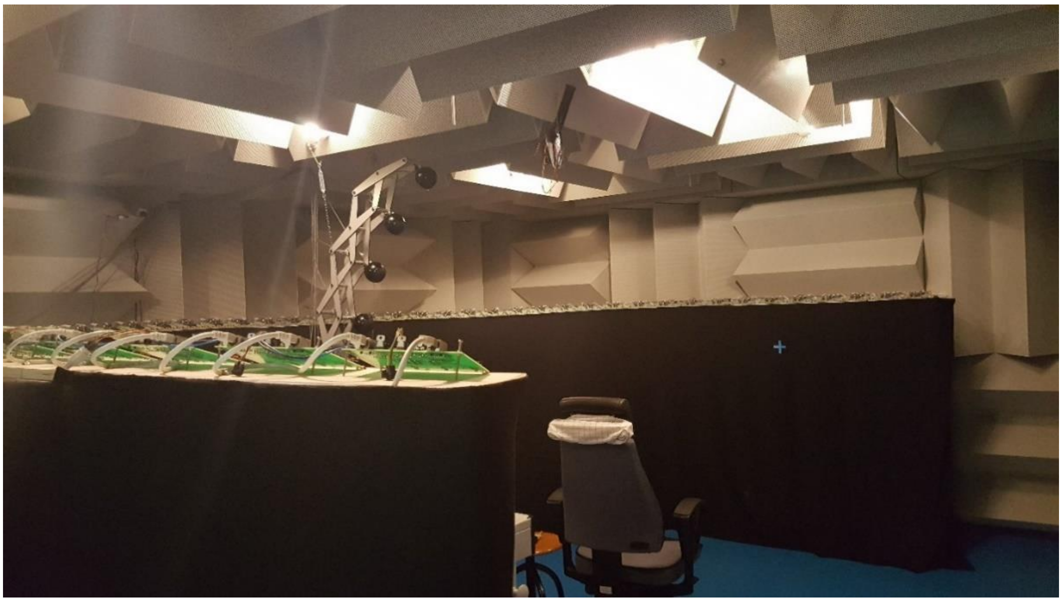


Fig. 1 The free-field laboratory in Leipzig. The black curtain covers the 47 loudspeakers. The chair is placed in the middle to the semi-circle and it can rotate to face any direction. Photograph taken by Mikaella Sarrou.

Free-field laboratory at the International Laboratory for Brain, Music, and Sound Research (BRAMS), Montreal, Canada

This laboratory is also a semi-anechoic room (2.5 X 5.5 X 2.5 m) and it also has foam wedges for sound attenuation (Fig. 2). On the side of this chamber, eighty loudspeakers form a spherical array (Orb Audio, New York, NY, USA) with 1.8m diameter. A comfortable chair with a neck rest was located in the centre of this array. An acoustically transparent black curtain was placed in front of the loudspeakers to avoid localization of sounds based on visual cues (not shown in Fig. 2). A laser pointer and an electromagnetic head-tracking sensor (Polhemus Fastrak, Colchester, VT) were attached to a headband work by the participants to confirm their 0° -to-the-front head orientation. The stimuli were generated and controlled by customized Matlab scripts (The MathWorks Inc., Natick, Massachusetts, USA) and presented to the speakers using TDT System 3 hardware (Tucker Davis Technologies, Alachua, FL, USA). An intercom system was placed to allow communication between the experimenter and the participants during experimental sessions.

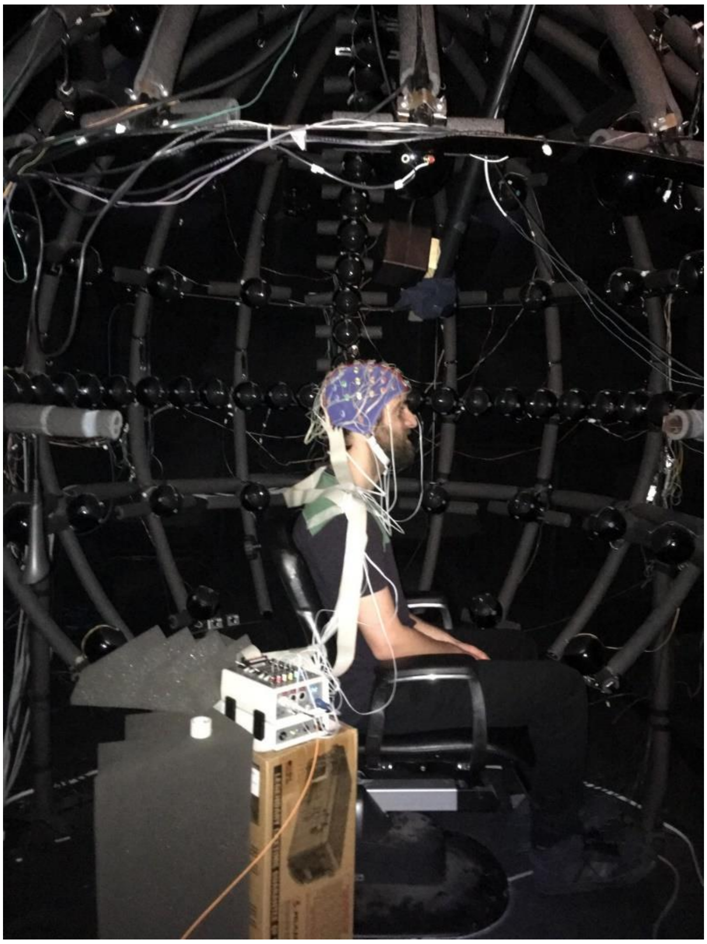


Fig. 2 The free-field laboratory in Montreal. Here the black curtain was removed to allow better view of the spherical speaker array. The chair is placed in the middle of the array and can rotate. Photograph taken by Mikaella Sarrou.

2.2 Electroencephalogram

The electroencephalogram (EEG) is a widely used non-invasive method for recording neuronal potentials with electrodes placed at the scalp. Its use mandates the application of an electrolyte gel between the electrodes and the scalp in order to reduce the impedance of the skin and hair. The EEG is sensitive to the time domain, meaning it can record activity to the level of milliseconds. Therefore, it is the ideal method for studying auditory events because their processing all the way up to the cortex is very rapid.

The activity recorded with EEG represents postsynaptic activity of the neurons (Luck 2014). The neurons communicate with each other with the changes in the homeostasis of the ionic elements that are located in and around the neurons. This change causes an action potential to travel down the cell and then to be transferred to the subsequent cells. This activity can be elicited in symphony from a large number of neurons a response to the same stimulus and thus, it results as an electrical activity large enough in amplitude to be detected by the EEG.

2.3 Event-related potential technique

The event-related potential technique (ERP) is a specific technique that shows the brain activity as a response to a repetitive stimulus presentation and it is often applied with the use of EEG. The idea behind this technique is that the presentation of a stimulus will elicit neuronal responses that overlap with their ongoing activity (Luck 2014). Thus, by averaging the responses to the same stimulus from a large number of stimulus repetitions and different people, it is possible to extract the activity that represents only the response to the stimulus, since the ongoing neuronal activity is canceled out. This stimulus representation is the ERP.

The current study uses the motion-onset response (MOR), an ERP that represents the response to the onset of acoustic motion. This ERP requires the presentation of a specific stimulus that has an initial stationary part that starts to move to another location after a brief delay. The initial stationary part elicits the responses of the auditory system that are called together energy-onset (EOR; Chait et al. 2004). The EOR consists of one positive deflection, P1, a second negative deflection, N1 and a third positive deflection P2. The MOR consists of similar to the EOR peaks: a positive deflection, cP1, a negative deflection, cN1 and a third positive deflection, cP2. In the chapters that follow, I describe with which sound motion characteristics each peak is associated with. In addition, I show that under certain experimental designs, the cP1 might not be elicited and the cP2 could be presented in the ERP with two peaks, cP2a and cP2b, showing that certain stimuli characteristics lead this peaks into being separated into its constituent parts.

2.4 Difference waves

Difference wave is a technique used to isolate ERP components (Luck 2014). It requires the subtraction of the ERP of one condition from the ERP of another condition. For example, in the current thesis difference waves are used to isolate the MOR. In Chapter 4, the trials to conditions without an acoustic moving part were subtracted from those with an acoustic moving part. Thus, the responses to the initial stationary part of the stimulus, as well as responses to other stimulus characteristics (such as amplitude) were canceled out, leaving the waveforms only with the MOR. This is an important technique because it removes all possible neuronal activity that could have affected the signal.

2.5 Stimulus adaptation

The stimulus-specific adaptation is a paradigm used widely in research to study sensory effects. With the words of Mollon (1974) “If any one of our senses is exposed for some time to a stimulus that is unchanging in a particular attribute (such as colour, direction of movement, temperature), then our subsequent perception of that attribute, or our capacity to detect it, is briefly altered.” Thus, this paradigm is composed of the presentation of adaptor and probe paired stimuli: the adaptor is repeatedly presented and afterwards, the response to the probe stimuli is measure using either behavioural or electrophysiological measures or both. The underlying mechanism of our altered response following adaptation is based on the repeated activation of the neurons that respond to the adaptors. These neurons fire with each adaptor repetition following an exponential decline of the firing rate with increasing number of adaptor repetitions. Therefore, the response to the probe will be reduced, the more similar characteristics the adaptor and probe have. This paradigm has been used to study visual and auditory motion aftereffects (Carlile and Leung 2016, Dong et al. 2000, Grantham 1989, Mollon 1974). The usual application to study the behavioural effects of the auditory motion aftereffect was to present an auditory stimulus with a certain direction repeatedly and then ask participants to indicate the direction of the probe stimulus. The participants would reply that the probe stimulus moved in the opposite direction, even if it was stationary. Moreover, in the cortical responses, the adaptation paradigm was used to study the encoding of locations in the auditory cortex (Salminen et al. 2009; 2012). In Salminen’s (2009) study, the participants were exposed to the repetitive presentation of stationary auditory stimuli and the response to probe stationary stimuli was measured using magnetoencephalography. The results of this study showed that the response to the probes was reduced when the adaptor was presented at the location in which the neuronal populations are strongly activated, which is in accordance with the population rate code. The adaptation paradigm was also used to study the direction-specificity of the MOR using EEG (Grzeschik et al. 2013). In this study, the participants were exposed to repetitive moving sounds that either had the same or opposite direction with the subsequent probe stimuli. The MOR elicited by the probe stimuli had lower amplitude following the adaptor that had the same direction as the probe and the authors interpreted this as evidence that neurons encoding direction contribute to the signal of the motion-onset response. Taken together, this paradigm is very useful to investigate which stimuli attributes contribute to a response. Thus, in Chapter 7 this paradigm is employed to study the interaction of the binaural and spectral cue for auditory motion direction detection.

2.6 Research questions and aims of the studies

As it is outlined in the introduction, the neuronal mechanism underlying auditory motion perception has been vaguely described. However, there is an increasing number of evidence that show that specialized motion areas and mechanisms exist in the cortex. In order to study, how exactly these mechanisms function, it is important to know which aspects of the stimulus paradigm affect the response. This study focused on examining the motion-specific as well as the auditory-specific effects of the stimuli on the MOR.

Study 1. In this study, I focused on eliciting the MOR in the freefield. Since this response, is elicited by the specific stimulus paradigm, the focus was to elicit it in a passive paradigm. In addition, the effects of stimulus frequency on the response's amplitude and latency were studied. Moreover, another focus was to test how dependent the response is on the adaptation provided by the initial stimulus part. I hypothesized that, due to the adaptation provided by the initial stimulus part, the motion response would be smaller after moving than after stationary adaptation. Also, I expected that the effects of frequency would follow the literature and since the motion response is a late response, the amplitude would be smaller after the high frequency than low frequency stimulus presentation.

Study 2. In this study, the stimulus characteristics were adapted after failing to elicit the response in the previous study. In addition, I employed an active instead of a passive paradigm, since data from the literature show that the motion response is strongly dependent on the allocation of attention on auditory motion. Thus, in this study, the elicitation of the MOR was successful. The analysis in the current study aimed at examining the effects of stimulus frequency on the MOR. Higher amplitude on the motion response was expected after the presentation of stimuli with high frequency spectrum. Also, I studied the effects of hemifield presentation and the direction of motion on the MOR.

Study 3. This study is focused on analysis from data collected in the previous study. The focus, however, is on the effects of the stimulus paradigm on the MOR. I hypothesized that after the adaptation provided by an initial moving part, lower amplitude was expected in comparison to the stimuli with an initial stationary part.

Study 4. This study examined further the effects of stimuli characteristics on the MOR. Since the latency of the motion response in the previous study was a bit later than what is usually reported in the literature, the focus here was to test the effects of motion velocity and adaptation duration on the motion response.

Study 5. This behavioural study focused on the effect of motion adaptation on the rear field to the presentation of motion in the frontal field. Thus, the presentation of adaptors and probes within the left-frontal and left-rear fields aimed at locations that share the same interaural time and level differences and the disambiguation of auditory location and motion is based on how these interaural cues interact with the spectral cues.

Chapter 3. *Study 1*: Motion-onset response: A passive paradigm – evidence for the dependence of the response on the stimuli characteristics.

Abstract

The current experiment studies the stimuli characteristics on the elicitation of the motion-onset response (MOR). This specific response is measured with electroencephalogram and it is elicited when a sound motion follows a stationary sound without any temporal gaps between them. The stationary part serves as an adaptive sound and the onset of motion provides a release-of-adaptation, which gives rise to the MOR. Here, one of the focus was to investigate the effect on the MOR when the initial part is moving in space instead of being stationary. In addition, a secondary focus was the effect of the stimuli frequency on the MOR. The results showed that the current paradigm did not elicit the MOR. Comparison of the current experimental settings with those used previously in the literature showed that the MOR is strongly depended on the adaptation time provided by the first part of the stimuli. Lastly, an analysis of the stimuli characteristics and task paradigm that might affect the MOR is provided in the discussion.

3.1 Introduction

The ability to detect and recognize sources of sounds in the environment is crucial for survival. This ability gets more complicated when these sources move, in other words change location in space. The mechanism of moving sound source localization in the brain has been a matter of debate and a growing number of evidence exists in support of specialized cortical areas (Poirier et al. 2017) or cortical mechanisms (Krumbholz et al. 2007; Magezi et al. 2013) in the detection of sound motion.

In the past decade, a growing number of studies employing electroencephalogram (EEG) have investigated the cortical response to auditory motion, the motion-onset response (MOR) (Getzmann 2009; 2011; Getzmann and Lewald 2010; 2012; Grzeschik et al. 2013; 2016; Krumbholz et al. 2007). This is a specific response that is elicited with the presentation of a moving sound after the presentation of a stationary sound. The mechanism underlying this response is an adaptation method. The initial stationary part of the sound activates neurons that are selective to the location in which the sound is presented. This is evident in the EEG via the initial auditory response with the apparent peaks, P1, N1 and P2 that are together named the energy-onset response (EOR) (Chait et al. 2004). The introduction of motion activates a different set of neurons that is believed to represent the activity of true motion detectors. This activation of the motion detectors is evident with the MOR. The MOR resembles the EOR in that it is triphasic with an initial positive (cP1), subsequent negative (cN1) and a second positive peak (cP2). The overall amplitude of the MOR, however, is lower than the one of the EOR. Since this response is relatively new, many studies have investigated how it is modulated by stimuli characteristics and by experimental settings. Getzmann (2009) presented sounds to the participants via earphones with the paradigm of the delayed motion-onset, as described above, and varied the velocity of the moving sounds by changing their duration. His results showed that high velocities elicited MOR with earlier latency and higher amplitude. Another study (Getzmann and Lewald 2010), compared the MOR elicited from four types of auditory motion: freefield presentation, head-related transfer function (HRTF), interaural time differences and interaural level differences. The results showed that freefield motion presentation elicits stronger MOR amplitude and that freefield and HRTF motion elicited earlier MOR and faster reaction times. The authors concluded that the motion presented in the freefield and HRTF have more natural motion percept and this affects the MOR. Moreover, other studies investigated which neuronal populations contribute to the MOR and how this is represented in this response. In a study using leftward and rightward moving stimuli within each hemifield, the results showed that the different peaks of the MOR show a pattern of contralaterality based on different motion characteristics (Getzmann 2011). The cN1 had higher amplitude in the hemisphere contralateral to the hemifield of stimuli presentation and the cP2 had higher amplitude to the hemisphere contralateral to the direction of motion. Based on these evidence, the author concluded that the early part represents the location of motion and the second represents the direction of motion. In addition, Grzeschik and colleagues (2013) investigated the contribution of direction-specific neurons on the MOR. In this study, the researchers presented moving sounds as adaptors before the test stimuli that elicit the MOR. These adaptors either had the same or different direction as the test stimuli. The results showed that the test stimuli elicited stronger MOR when the preceding moving stimuli had a different direction. The authors concluded that some of the neurons that contribute to the MOR encode the direction of motion. In a later study, the crossmodal effect of directionality was investigated between visual adaptors on the auditory MOR (Grzeschik et al. 2016). The results showed that the visual adaptors had an effect on the auditory MOR, but this was not specific to direction.

What is missing from the literature is the frequency-specificity of the MOR. The effects of the stimulus frequency on the components of the EOR have been previously documented. In an article investigating the magnetoencephalographic (MEG) response to tones of different frequency ranges, the data showed that the N1m elicited from low frequency tones had longer latency and larger amplitude than the N1m elicited from middle frequency or high frequency tones (Jacobson et al. 1992). In support of the previous results, an EEG study showed that the P150 had decreasing frontocentral positivity with increasing frequency (Verkindt et al. 1994). Furthermore, Wunderlich and Cone-Wesson (2001) showed that the mismatch negativity (MMN) decreases in area amplitude with increasing frequency. The last three studies show the frequency-specific effects on the early cortical responses and they highlight the stronger response of the cortical activity to the low frequency stimuli. Despite these results, a different EEG study investigating stimulus frequency as a parameter that affects the P3 found different results (Sugg and Polich 1995). This study employed an oddball paradigm using variations of stimulus frequency for standard and deviant sounds. The P3 elicited from standard sounds had higher amplitude if these sounds had high instead of low frequency spectrum. The authors concluded that since stimulus frequency affects the P3, it is important to consider it when designing experiments since it might lead to different conclusions.

Therefore, in the current study we wanted to examine the effects of stimulus frequency on the MOR by comparing the MOR elicited from low frequency noise stimuli to the MOR elicited from high frequency stimuli. In addition, we wanted to study how the modulation of the MOR would change if the initial part of the specific stimulus paradigm was a moving one instead of a stationary one. We hypothesized that the MOR would differ between the two stimulus frequency ranges. Also, on the basis that the MOR is elicited from true cortical motion detectors, the initial moving part would lead to lower MOR amplitude than the stationary part.

3.2 Method

Participants

A total of 15 healthy volunteers participated in the study (8 females) with a mean age of 24.2 years. All participants were informed about the scope of the experiment and had signed a written consent form for their participation. The experimental procedures were in agreement with the Declaration of Helsinki (World Medical Association 2000).

Setup

The experiment was conducted in complete darkness in a semi-anechoic chamber (40 m^2; Industrial Acoustics Company, Niederkrüchten, Germany) in which 47 loudspeakers (Visaton, FRS8 4 Ohm, Haan, Germany) with radius 2.35 m are arranged in an azimuthal, semi-circular plane spanning from 98° left to 98° right in the frontal hemifield. The distance between the loudspeakers was 4.3°. Each loudspeaker was equilibrated individually. For this, the transmission spectrum was measured using the a Bruel & Kjaer measuring amplifier (B&K 2610), a microphone (B&K 2669, pre-amplifier B&K 4190) and a real-time signal processor (RP 2.1, System3, Tucker Davis Technologies, TDT, Alachua, Florida, USA). For each loudspeaker, a calibration file was generated in Matlab 6.1 (The MathWorks Inc., Natick, Massachusetts, USA) and was later used to ensure flat spectra of the acoustic stimuli across the frequency range tested. The loudspeakers were covered with a sound transparent black curtain to prevent participants having visual location information of the loudspeakers’ positions. The generation of the stimuli was digital using real-time TDT processors (RX8, System3) and controlled with custom-made scripts in MATLAB 7.5 (R2007b, The MathWorks Inc., Natick, Massachusetts, USA). An infrared camera was installed in this chamber to oversee participants during testing.

Stimuli

The stimuli were 2 Gaussian noise signals, one at a low frequency range (200-800 Hz) and one at high frequency range (1500-6000 Hz). The intensity was set on 40 dB sensation level (SL; above individual hearing threshold) with a rowing level +/- 6dB. For each participant, the individual hearing threshold for the low- and high-frequency noise bands was measured at the beginning of each session using a yes/no (heard/not heard) paradigm for stationary and moving sound stimuli.

Acoustic stimuli had a duration of 500ms separated in two 150ms-components (without any temporal gap between them), either stationary or moving. Four different signal combinations were used: (i) stationary-stationary (StSt), (ii) stationary-moving (StMo), (iii) moving-stationary (MoSt), (iv) moving-moving (MoMo). For stationary stimulus components a single speaker was activated at a particular azimuthal position. Moving sounds were created by sequential activation of adjacent speakers with linear cross-fading adjusting the intensity of the output signal voltage of neighbouring speakers. Stationary sound stimuli were presented in one of the following locations (negative numbers show locations left from midline): -40°, -10°, +10°, +40°. The velocity of the moving sounds was invariantly set at 120°/sec. Sound presentation started at 10° or 40° in either hemifield. For StSt the sound source position remained unchanged. For moving sounds, either being the first and/or the second component of the 500ms stimulus, the movement starting at +/-10° was laterally directed to +/-40° and the movement starting at +/-40° was frontally directed to +/-10°. This way sound stimuli being stationary or moving (and the latter moving either inward or outward) remained within the same hemifield.

Altogether 32 stimulus conditions were explored: 4 signal combinations (StSt, Stmo, MoSt, MoMo), 2 frequency ranges (lf, hf), both hemifields (left, right). Each stimulus combination was presented 60 times in a pseudo-random manner resulting in 1920 stimulus presentations. The interstimulus interval (ISI) varied between 350-450 ms the end of the previous stimulus to the beginning of the next stimulus. Short breaks were inserted every about 12 minutes.

Procedure

Participants were asked to sit in a comfortable chair in the middle of a semi-circular loudspeaker array in the semi-anechoic chamber. In order to keep the head in a steady position, they were instructed to keep their eyes on a fixation cross during the audiogram and the experiment. Both of these sessions were contacted in complete darkness.

The heard/not-heard audiogram was combined with a simple staircase paradigm used to determine participants' individual hearing threshold for stationary and moving sounds of low and high frequency. The stationary sounds were presented at 0° in front of the participants and the moving sounds spanned a trajectory from -30° (left from midline) to +30°, each lasting 1000ms, with an initial intensity of 62 dB SPL. Participants were instructed to press a left button on a response box to indicate that they have detected a sound (*heard* response) and the right button when they did not detect the sound (*not-heard* response). The stimulus' intensity was decreased by 5 dB for each *heard* response and increased by 5 dB for each *not-heard* response. The average of the stationary and moving noise stimuli of low frequency range defined the SL of the low frequency noise stimuli; the average of the stationary and moving noise stimuli of high frequency range defined the SL of the high frequency noise stimuli.

During the experiment, the participants were instructed to watch a subtitled-movie without sound of their choice and minimize head movements. Therefore, the current experiment employed a passive experimental design.

For the current analysis, all stimuli were used to calculate difference waves to investigate the elicitation of the motion-onset response and its modulation based on stimuli frequency. The stimuli were averaged for type and frequency level, resulting in 480 stimuli in each type (StSt, StMo, MoSt, MoMo) with 240 stimuli for each frequency level (low, high) (Table 1).

Table. 1 Table showing the stimuli conditions used in the current analysis. The ‘Frequency’ column refers to the frequency spectra of the stimuli, low (200-800 Hz) and high (1500-6000 Hz).

Type	**Frequency**
StSt	Low
StSt	High
StMo	Low
StMo	High
MoSt	Low
MoSt	High
MoMo	Low
MoMo	High

This was the initial experiment for the investigation of the MOR and thus, the focus was how the different stimuli characteristics affect this response. In subsequent, experiments the focus was also how the laterality and hemifield presentation affect this response.

Data Recording and Analysis

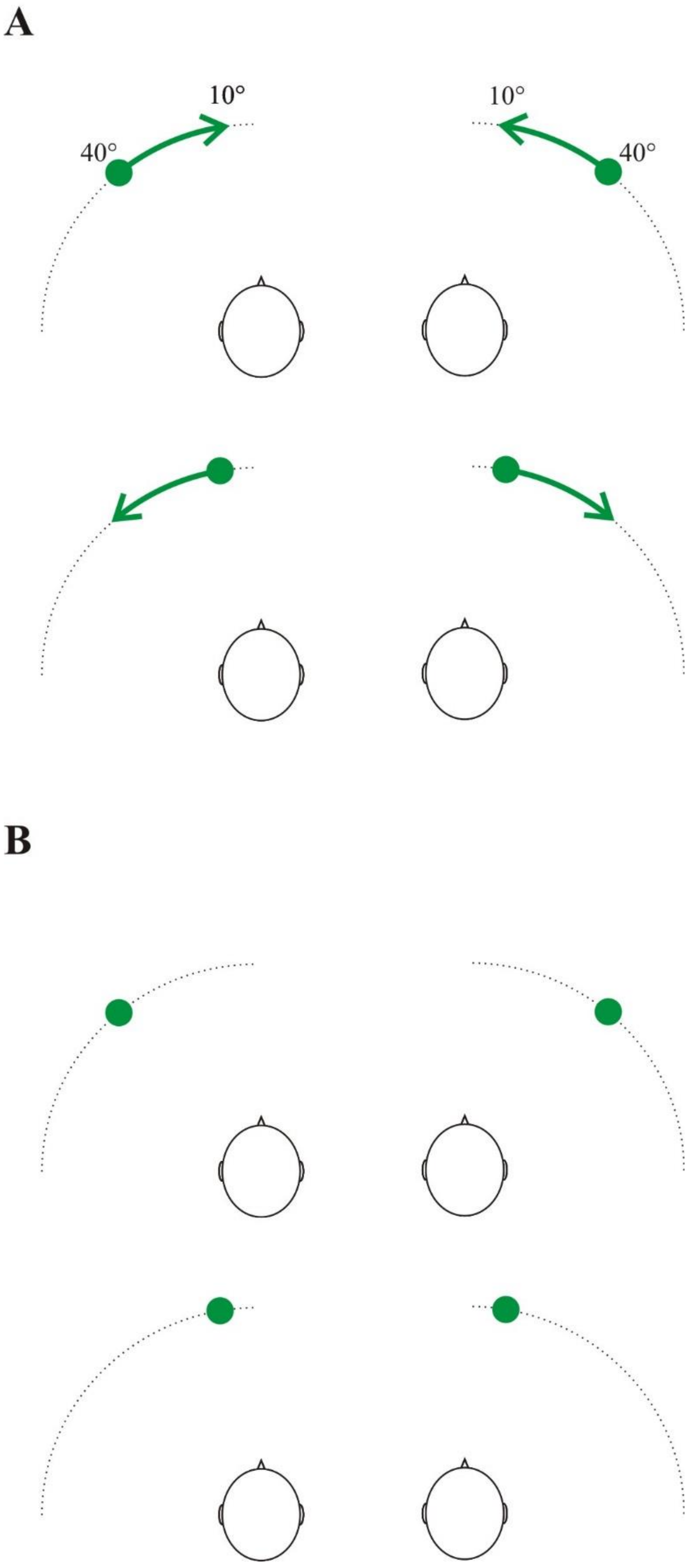
The EEG data were recorded using actiCAP Standard-2 (Brain Products GmbH, Gilching, Germany) at a sampling rate of 500 Hz. The recording consisted of 64 Ag/AgCl active electrodes placed on the scalp according to the international 10-20 system (Towle et al. 1993). The horizontal EOG (hEOG) was recorded by placing 2 electrodes next to the eyes and the vertical EOG (vEOG) was recorded by placing 2 electrodes above and below the left eye; both placements were at the level of the retina. The ground electrode was placed in the middle of the forehead just above the nation and the online reference electrode was placed at the tip of the nose. Two additional electrodes were placed on the mastoids. Impedances were kept below 5 kΩ.

The EEG data were pre-processed with Brain Vision Analyzer 2.0 (Brain Products GmbH, Gilching, Germany). The data were bandpass filtered (IIR, cut-off frequencies 0.1 40Hz; slope 48 dB/octave) and re-referenced to the average of 60 electrodes (58 scalp channels and 2 mastoids; implicit reference included in the re-reference). Then they underwent ocular correction using the Gratton and Coles procedure (Gratton et al. 1983). The epochs were set at -200ms to 700ms according to stimulus onset (200ms baseline period). Those which exceeded an amplitude of ±100 μV were excluded from further analysis. They were then baseline corrected for the period of 200ms before motion onset and averaged between the trials.

Statistical analysis

The statistical analysis was done using IBM SPSS Statistics for Windows, version 22.0 (IBM Corporation, Armonk, New York, USA). Visual inspection of the difference waves showed no prominent motion response. In this case, we decided to extract the mean activity of the area in rectified values separately for each stimulus part: 0-250ms for the initial part, 250-500ms for the second part for the Cz electrode.

Repeated-measures ANOVA (rm ANOVA) were carried out at the extracted area amplitude in a 2X2 design, with the factors history (stationary, motion) and frequency range (low, high), with p values set at .05 (Fig. 1). Multiple comparisons were reported with Bonferroni correction.


A
10°
40°
10°
40°
B

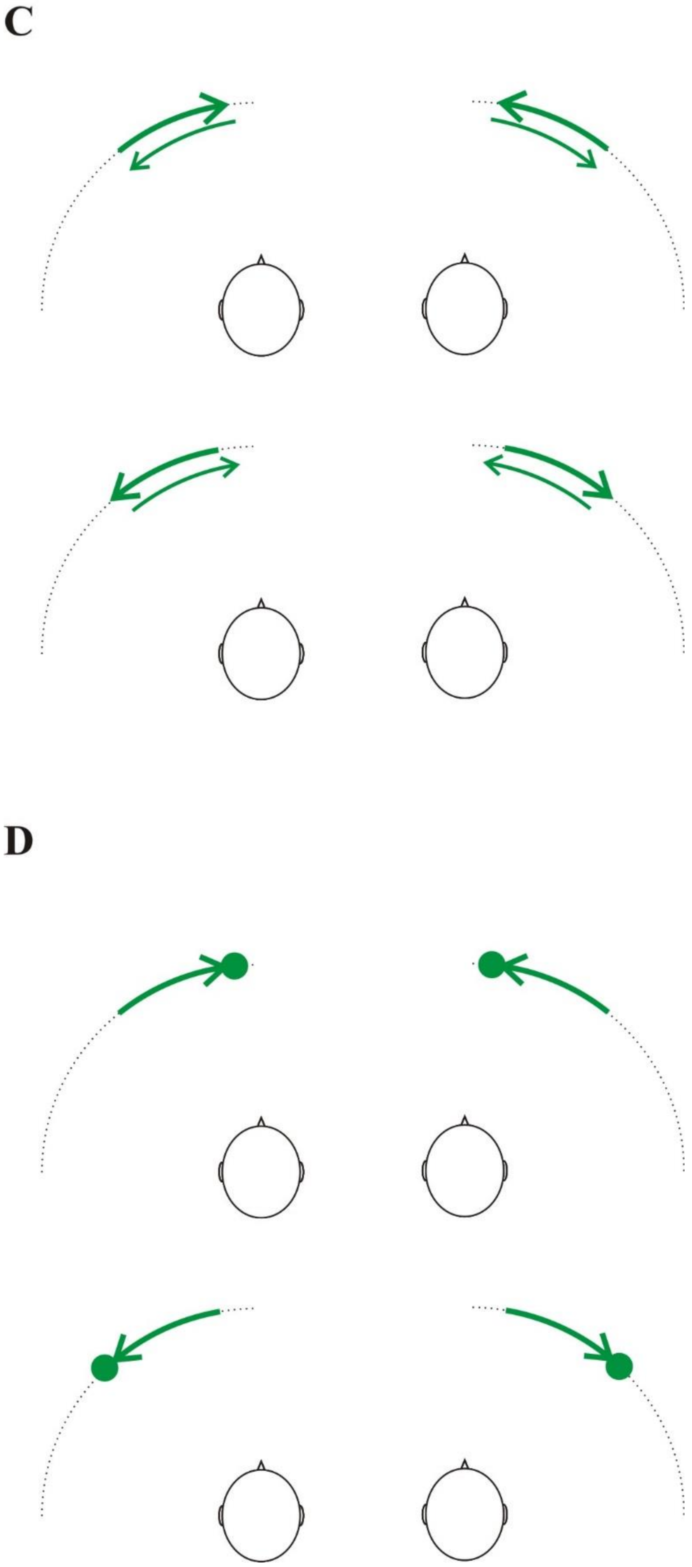


Fig. 1 Graphic representation of stimuli conditions. In each hemifield (left and right column), there were 2 types of StMo (A), StSt (B), MoMo (C) and MoSt (D) stimuli: one started from a lateral position and another near-central position. (A) Stimuli presented in the StMo condition. They started at the positions indicated by

the green dots, where they remained for 250ms. At 250ms the stimuli started to move to the other location within the same hemifield and the motion lasted 250ms (green arrows). (B) Stimuli presented for the StSt condition. There were 2 stimuli in each hemifield, one for each location, in which they remained for 500ms. (C) Stimuli presented for the MoMo condition. They started at the lateral (top row) or central (bottom row) location and moved to the other location within the same hemifield within 250ms. They then returned to the initial location within 250ms. (D) Stimuli used for the MoSt condition. They either started at a lateral (top row) or central position (bottom row) and they moved to the other location within the same hemifield within 250ms. They remained at the second location for 250ms. All of the stimuli illustrated here were presented in lf (200-800 Hz) and hf (1500-6000 Hz).

3.3 Results

The onset of the stimuli elicited the common triphasic response of the auditory system, P1, N1, and P2, referred to as the energy-onset response (EOR) (Chait et al. 2004) (Fig. 1). The conditions with the low frequency spectrum sounds seem to have lower peak-to-peak amplitude differences than those elicited by the high frequency conditions. The onset of motion within the stimuli did not elicit a prominent motion response for neither the StMo, nor the MoMo conditions. The presence of a motion response would be evident by another triphasic set of components that resembles the EOR but has lower amplitude. On the contrary, after the initial 250ms phase of the stimuli, signal has a short sustained negativity and later on, moves towards the baseline, in all conditions.

Four difference waves were calculated, 2 for each frequency range, with the following subtractions: StMo minus StSt, MoMo minus MoSt. The conditions had no apparent differences between them and thus, no differences appeared on the difference waves as well. Therefore, a decision was made to extract rectified area values of the Cz electrode for the first part of the stimuli conditions, 0-250ms and for the second part 250-500ms and to analyse them in a 2X2 repeated-measures Anova, with the factors frequency (low, high) and history (stationary, motion). This way, if any differences were detected, we could have some evidence on which factor affects the signal the most. The analyses did not yield any significance in neither of the two time frames values. For the initial time frame, 0-250ms, the following mean area values were calculated: StMoL minus StStL M= .513, StMoH minus StStH M= .404, MoMoL minus MoStL

M= .394, MoMoH minus MoStH M= .38. There was a tendency for higher area amplitude resulting from the stationary than the moving part, p=.434, and a tendency for higher area amplitude resulting from the low than the high frequency conditions, p=.4. For the subsequent time frame, 250-500ms, the following mean area values were calculated: StMoL minus StStL M= .654, StMoH minus StStH M= .428, MoMoL minus MoStL M= .491, MoMoH minus MoStH M= .375. There was a pattern of higher area amplitude resulting from stationary than moving sounds, p=.336, and a tendency for higher area amplitude resulting from low than high frequency conditions, p=.174.

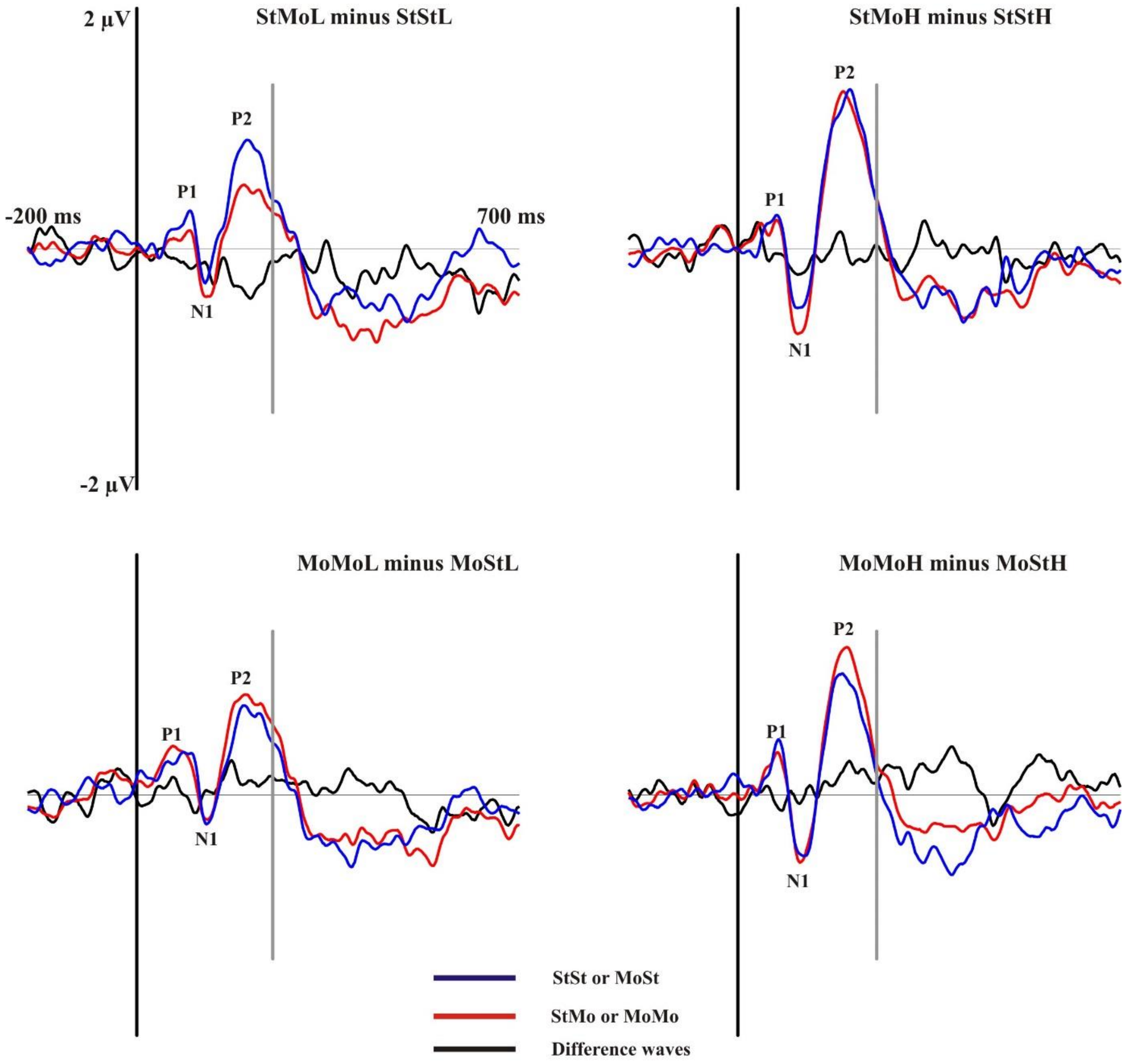


Fig. 2 Overlay of raw waveforms used to calculate difference waves and the resulting difference waves. The conditions StSt and MoSt are coloured in blue. The

conditions StMo and MoMo are coloured in red. The difference waves resulting from each subtraction are coloured in black. Vertical black lines show the onset of the stimuli. Vertical grey lines show the 250ms mark when the change occurred within the stimuli. The left column represents responses from the stimuli presented at low frequency (L; 200-800 Hz) and the right column represents the responses from the stimuli presented at high frequency (H; 1500-6000 Hz).

3.4 Discussion

The purpose of the current study was to investigate the effects of stimulus frequency on the EEG response to auditory motion, the MOR. In addition, since the elicitation of the MOR is based on the specific stimulus that has the initial stationary part acting as an adaptor, extra stimuli were added that had a moving sound as the initial part. The results showed that there was no MOR elicited from any of the conditions neither in the raw waveforms nor in the subsequently calculated difference waves. In addition, the statistical processing on the values extracted from the difference waves showed that there were no differences between the conditions for neither the first nor the second part of the stimuli. No differences were anticipated as a response to the first part of the stimuli. Yet the lack of the MOR as a response to the second part of the stimuli points to several methodological issues in the paradigm used to elicit the MOR in the current study and these issues will be outlined below.

To begin with, an obvious error that one can see from the graphs is that the duration of the initial part of the stimuli was not long enough to allow auditory responses to be elicited and then to be followed by the signal’s return to baseline. The acoustic stimuli elicit cortical responses that are evident with the P1, N1 and P2, and then usually the signal has a brief sustained negativity and then it returns to baseline. In the current study, the initial part of the stimulus had a very brief duration (250ms) and this duration was not long enough to allow the signal to return back to baseline, before introducing the onset of motion. The onset of motion started while the P2 component was elicited. In an attempt to minimize the experimental time, the duration of the stimuli was much reduced, overlooking the usual duration used previous studies. These durations are longer varying from 700-1000ms (700ms: Getzmann 2009; 2011; Getzmann & Lewald 2010; 2011; 2012; Kreitewolf et al. 2011, 1000ms: Getzmann & Lewald 2014; Grzeschik et al. 2013; 2016; Krumbholz et al. 2007). Thus, the offset duration of the P2 overlapped with the motion onset and the normal MOR elicitation. In addition, the duration of the initial stationary sound does not serve only the purpose of eliciting the EOR. It also functions as an adaptor that engages the activity of the neuronal populations corresponding to the sound so that, the onset of motion in the stimuli activates neurons responding to motion. The current duration then is very brief to allow for the adaptation to take place.

Moreover, the duration of motion part is also brief in comparison to the literature. The usual range is 330-500ms (330ms: Getzmann 2011, Grzeschik et al. 2016, Kreitewolf et al. 2011; 500: Getzmann & Lewald 2010; 2011; 2012; 2014, Grzeschik et al. 2013). The study of Getzmann (2009), however, on the velocity effects on the MOR used the following durations: 250ms, 500ms and 1000ms. Although, in the latter paper the results show that the motion duration of 250ms led to the elicitation of the MOR, this is based on the fast velocity rather than the duration of the motion. In relation to this, the current velocity of the sound was set to 120°/s. This velocity is in line with those usually used in the literature and faster than the velocity used in a previous study that showed MOR elicitation (60°/s: Grzeschik et al. 2013).

Another factor that could have contributed to the current data is the task. The current study employed a passive paradigm during which the participants were instructed to watch a silent movie with subtitles and pay no attention to the ongoing sounds. However, attentional factors significantly affect the MOR (Kreitewolf et al. 2011). Although there is no study comparing the MOR during passive and active experimental tasks, the aforementioned study compared the MOR elicited between a task that asked participants to focus on the motion trajectory and a task that asked participants to focus on the pitch of the sounds. Their results showed that there was an increase of MOR amplitude when the task was motion-specific and the authors interpreted it as evidence that motion processing involves attentional processing as well.

All things considered, it seems that what plays a major role in the MOR elicitation is the duration of the initial part of the stimulus. It seems that the MOR is heavily dependent on the sufficient adaptation that this part provides. This point is addressed in the following studies.

Chapter 4. *Study 2*: Sound frequency affects the auditory motion-onset response in humans.

Abstract

The current study examines the modulation of the motion-onset response based on the frequency-range of sound stimuli. Delayed motion-onset and stationary stimuli were presented in a free-field by sequentially activating loudspeakers on an azimuthal plane keeping the natural percept of externalized sound presentation. The sounds were presented in low- or high-frequency ranges and had different motion direction within each hemifield. Difference waves were calculated by contrasting the moving and stationary sounds to isolate the motion-onset responses. Analyses carried out at the peak amplitudes and latencies on the difference waves showed that the early part of the motion response (cN1) was modulated by the frequency range of the sounds with stronger amplitudes elicited by stimuli with high frequency range. Subsequent post hoc analysis of the normalized amplitude of the motion response confirmed the previous finding by excluding the possibility that the frequency range had an overall effect on the waveform, and showing that this effect was instead limited to the motion response. These results support the idea of a modular organization of the motion-onset response with the processing of primary sound motion characteristics being reflected in the early part of the response. Also, the study highlights the importance of specificity in auditory stimulus design.

4.1 Introduction

Localizing moving sound sources is vital for navigation in, and interaction with, our immediate environment. On the azimuthal level, this ability is based on the dynamic aspects of the interaural differences in time of sound arrival at the two ears (ITD) and sound-pressure level (ILD) between the two ears (Middlebrooks and Green 1991; Wightman and Kistler 1993). The ITDs and ILDs change as a function of spatial location and ILDs also change as a function of stimulus frequency (King et al. 2001). Therefore, in behavioural studies, both are related to the perception of different levels of sound frequency through the duplex theory (Rayleigh 1907). Localization of low frequency sounds is predominantly based on ITD; localization of high frequency sounds is predominantly based on ILD (Stevens and Newman 1936). In electrophysiological studies, these interaural differences are represented in the physiological responses at the brainstem and are thought to have parallel but separate processing pathways (Grothe et al. 2010; McAlpine et al. 2001; Tollin 2013). Evidence from cat studies showed that these cues merge at the central nucleus of the inferior colliculus (Chase and Young 2005). However, a study in humans provided evidence that the processing of ITDs and ILDs is not integrated, even at the level of the cortex (Schröger 1996).

Auditory motion studies often employ slow transitions of ITD and ILD to create the sensation of a moving sound (ITD: Krumbholz et al. 2007; ILD: Poirier et al. 2005; both: Getzmann and Lewald 2010; Griffiths et al. 1998) presented through headphones and they unavoidably create an intracranial sound percept. Although some of these studies do not focus on how auditory motion is created, interpreting the results nevertheless requires details of the cue/s that were used. For example, Magezi and colleagues (2013) argue that simulating motion using only ITDs ensures that there is no contribution from non-motion neurons in the electrophysiological signal, such as those responding to sound level or frequency changes, which appear when simulating motion using ILDs or head-related transfer function (HRTF), respectively. Thus, in their study, they used ITD transitions to present auditory motion and their results showed that auditory ILD-motion processing is not direction-sensitive. These results, however, could have been differently interpreted if ILDs were also included. Also, in a behavioural study (Dong et al. 2000) on auditory motion aftereffect (aMAE), the researchers studied the effects of adaptation when the adaptor and the probe stimuli had the same or different frequency range.

The data showed that adaptor stimuli with a low frequencyrange resulted in stronger adaptation of high frequency probe stimuli than when the spectra of the adaptor and probe stimuli were reversed. This result shows that stimuli with a low frequency range (which are processed based on ITDs according to the duplex theory) adapt more strongly than stimuli with high a frequency range (which are processed based on ILDs) than vice versa. If high frequency moving stimuli would also activate non-motion related neurons, such as neurons responding to sound level changes, then these stimuli would activate a larger set of neurons, and thus the adaptation effect would be stronger on the stimuli with a low frequencyrange, because the latter activate a smaller number of neurons. This contradicts the argument from Magezi et al. (2013).

Earlier studies also point out the benefits of using natural cues with externalized sound presentation (Getzmann and Lewald 2010). In the aforementioned study, free-field motion resulted in faster reaction times, as well as earlier and stronger motion-onset cortical responses (MOR), in comparison to ITD, ILD and HRTF motion. Furthermore, the bandwidth of the acoustic stimuli affects cortical responses (Langers et al. 2014). In the latter study, broadband stimuli excited a bigger range of neurons with different characteristic frequencies than narrowband stimuli, and thus resulted in a stronger signal. However, the authors concluded that because broadband stimuli recruit neuronal populations with different characteristic frequencies, this leads to neuroimaging results that are not consistent with stimulus repetition, thus highlighting the consistent results with narrowband stimulus presentation.

The electrophysiological response to the onset of auditory motion is the MOR, which is a transient response elicited with the onset of motion in an otherwise stationary sound (Krumbholz et al. 2007). It has a unique morphology with distinct negative deflection (cN1), followed by a positive deflection (cP2), sometimes preceded by an earlier positive deflection (cP1). The early and late parts of MOR (cN1 and cP2, respectively) represent distinct levels of motion processing: the early part is sensitive to the hemifield in which the motion occurs and is lateralized to the contralateral hemisphere, whereas the latter part is sensitive to the motion direction and is lateralized to the hemisphere contralateral to motion direction (Getzmann 2011). MOR components are modulated by sound motion characteristics and attention (Velocity: Getzmann 2009; attention: Kreitewolf et al. 2011). Thus, the data illustrate the complexity of auditory motion processing and highlight the significance of the specificity of the stimuli and experimental design.

To further investigate the modular organization of the auditory motion response, and to determine whether it is frequency-specific, the current study examines the modulation of MOR when presenting sounds of different frequency ranges in the free-field triggering externalized motion percepts. White noise stimuli have been used in most studies using free-field signal presentation to quantify MOR (Getzmann 2009, 2011; Getzmann and Lewald 2010, 2011, 2012, 2014; Grzeschik et al. 2013, 2016; Kreitewolf et al. 2011). This impedes the exploration of the extent to which the underlying frequency-specific central auditory processing affects the respective auditory cortical responses. Therefore, on the current study, we used two narrow-band Gaussian noise stimuli (high frequency and low frequency) as acoustic stimuli in an experiment employing a delayed motion-onset sound paradigm.

4.2 Method

Participants

A total of 14 healthy volunteers participated in the study (10 females) with a mean age of 24.9 years (13 right-handed, 1 ambidextrous) in exchange for a small monetary compensation. All participants were informed about the scope of the experiment and had signed a written consent form for their participation. The experimental procedures were in agreement with the Declaration of Helsinki (World Medical Association 2000) and were approved by the local Ethics Committee of the University of Leipzig.

Setup

The experiment was conducted in complete darkness in a semi-anechoic chamber (40 m^2; Industrial Acoustics Company, Niederkrüchten, Germany) in which 47 loudspeakers (Visaton, FRS8 4 Ohm, Haan, Germany) with radius 2.35 m were arranged in an azimuthal, semi-circular plane spanning from 98° left to 98° right in the frontal hemifield. Sound-attenuating sponge wedges (size 0.5 m) were attached to the four walls and the ceiling of this chamber. The distance between the loudspeakers was 4.3°. Each loudspeaker was calibrated individually. For this, the transmission range was measured using a Bruel and Kjaer measuring amplifier (B&K 2610), a microphone (B&K 2669, pre-amplifier B&K 4190) and a real-time signal processor (RP 2.1, System3, Tucker Davis Technologies, TDT, Alachua, Florida, USA). For each loudspeaker, a calibration file was generated in Matlab 6.1 (The MathWorks Inc., Natick, Massachusetts, USA) and was later used to ensure flat spectra of the acoustic stimuli across the frequency range tested. The loudspeakers were covered with a sound transparent black curtain to prevent participants from having visual location information of the loudspeakers' positions. Stimuli were generated digitally using real-time TDT processors (RX8, System3) and controlled with custom-made scripts in MATLAB 7.5 (R2007b, The MathWorks Inc., Natick, Massachusetts, USA). An infrared camera was installed in this chamber to oversee participants during testing.

Stimuli

The stimuli were 2 Gaussian noise signals, one with a low-to-moderate frequency range (300-1200 Hz) and one with a high frequency range (2000-8000 Hz) and they were identical during the entire session for all participants. The intensity was set to 50 dB sensation level (SL; above individual hearing threshold) with a rowing level of +/- 3dB. For each participant, the individual hearing threshold for the low- and high-frequency noise ranges was measured at the beginning of each session using a yes/no (heard/not heard) paradigm for stationary and moving sound stimuli.

Acoustic stimuli had a duration of 1000 ms separated in two 500 ms-components (without any temporal gap between them), either stationary or moving. Four different signal combinations were used: (i) stationary-stationary (StSt), (ii) stationary-moving (StMo), (iii) moving-stationary (MoSt), and (iv) moving-moving (MoMo). For stationary stimulus components a single speaker was activated at a particular azimuthal position. Moving sounds were created by the sequential activation of adjacent speakers, and we used linear cross-fading to adjust the intensity of the output signal voltage of neighbouring speakers. Stationary sound stimuli were presented in one of the following locations (negative numbers represent locations left of the midline): -60°, -10°, +10°, +60°. The velocity of the moving sounds was set to a constant 100°/sec. Sound presentation started at 10° or 60° in either hemifield. For StSt the sound source position remained unchanged. For moving sounds, either being the first and/or the second component of the 1000 ms stimulus, the movement starting at +/-10° was laterally directed to +/-60° and the movement starting at +/-60° was frontally directed to +/-10°. This way stationary and moving sound stimuli (the latter moving either inward or outward) remained within the same hemifield.

Altogether 32 stimulus conditions were explored: 4 signal combinations (StSt, StMo, MoSt, MoMo), 2 frequency ranges (low: lf, high: hf), 2 hemifields (left, right) and 2 locations (lateral: away from the midline, central: near the midline). Each stimulus combination was presented 60 times in a pseudo-random manner resulting in 1920 stimulus presentations. The interstimulus interval (ISI) varied between 900-1000 ms starting from the time point at which the participant gave a response (see below). Because of the long duration needed to collect a complete dataset from a single participant, data acquisition was divided into 2 sessions with an overall recording time of 121.6 – 128 min. 960 stimuli were presented within each session. Short breaks were inserted about every 12 minutes in each session.

Procedure

Participants were asked to sit in a comfortable chair in the middle of a semi-circular loudspeaker array in the semi-anechoic chamber. In order to keep the head in a steady position, they were instructed to keep their eyes on a fixation cross during the audiogram and the experiment. Both of these sessions were conducted in complete darkness.

The heard/not-heard audiogram was combined with a simple staircase paradigm used to determine each participant’s individual hearing threshold for stationary and moving sounds of low and high frequency. The stationary sounds were presented at 0° in front of the participants and the moving sounds spanned a trajectory from -30° (left of the midline) to +30°, each lasting 1000 ms, with an initial intensity of 62 dB SPL. Participants were instructed to press the left button on a response box to indicate that they detected a sound (*heard* response) and the right button if they did not detect the sound (*not-heard* response). Stimulus’ intensity was decreased by 5 dB for each *heard* response and increased by 5 dB for each *not-heard* response. The SL of the low frequency noise stimuli was defined as the average SL of the stationary and moving noise stimuli in the low frequency range; the SL of the high frequency noise stimuli was defined as the average of the stationary and moving noise stimuli in the high frequency range.

During the experiment, the participants were instructed to press the right button on a response box when the presented sound stimulus included a change in modality (stationary-to-moving or moving-to-stationary) and the left button if no such change occurred (stationary-stationary or moving-moving). The participants were asked to reply as quick and accurately as possible. Reaction time (RT) was quantified as the time from the offset of the stimulus to the button press, only for the correct responses. Participants were specifically instructed to respond after the offset of the stimulus to avoid overlap between the components of interest and the response potentials that appear over the motor cortex with the press of a button (Smulders and Miller 2012). To ensure that participants understood the response process, responses were monitored during the demonstration and throughout the experiment. If a participant did not follow this instruction, feedback was given during the breaks. If a response was not given within a 1000 ms time window, the trial was considered missed. Before starting the data acquisition, participants were instructed to familiarize themselves with the task in a demonstration block.

For the current study, 960 stimuli corresponding to 16 categories were analysed (Table 1). The 480 StMo stimuli were paired with their corresponding StSt stimuli to create difference waves.

Table 1 Table showing the stimulus conditions used in the current analysis (StMo- and StSt-stimuli) and those that are not (MoSt- and MoMo-stimuli). The stimulus name is provided in the first column. The ‘Frequency’ column refers to the frequency spectra of the stimuli, low (300-1200 Hz) and high (2000-8000 Hz). The ‘Laterality’ column refers the initial position of the stimuli before the arrow and the end position of the stimuli after the arrow (L: lateral, C: central). In the StSt condition, the position was not changed, whereas in the MoMo condition there were two movements.

Stimulus Name	Type	Frequency	Laterality	Hemifield	Analysed in this chapter
StMo_LowF_LatC_LH	Stationary-Moving	Low	L → C	Left	Yes
StMo_LowF_CLat_LH	"	"	C → L	"	"
StMo_LowF_LatC_RH	"	"	L → C	Right	"
StMo_LowF_CLat_RH	"	"	C → L	"	"
StMo_HighF_LatC_LH	"	High	L → C	Left	"

StMo_HighF_CLat_LH	"	"	C → L	"	"
StMo_HighF_LatC_RH	"	"	L → C	Right	"
StMo_HighF_CLat_RH	"	"	C → L	"	"
StSt_LowF_Lat_LH	Stationary-Stationary	Low	L	Left	"
StSt_LowF_C_LH	"	"	C	"	"
StSt_LowF_Lat_RH	"	"	L	Right	"
StSt_LowF_C_RH	"	"	C	"	"
StSt_HighF_Lat_LH	"	High	L	Left	"
StSt_HighF_C_LH	"	"	C	"	"
StSt_HighF_Lat_RH	"	"	L	Right	"
StSt_HighF_C_RH	"	"	C	"	"
MoSt_LowF_LatC_LH	Moving-Stationary	Low	L→C	Left	No
MoSt_LowF_CLat_LH	"	"	C→L	"	"
MoSt_LowF_LatC_RH	"	"	L→C	Right	"
MoSt_LowF_CLat_RH	"	"	C→L	"	"
MoSt_HighF_LatC_LH	"	High	L→C	Left	"
MoSt_HighF_CLat_LH	"	"	C→L	"	"
MoSt_HighF_LatC_RH	"	"	L→C	Right	"
MoSt_HighF_CLat_RH	"	"	C→L	"	"
MoMo_LowF_LatCLat_LH	Moving-Moving	Low	L→C→L	Left	"
MoMo_LowF_CLatC_LH	"	"	C→L→C	"	"
MoMo_LowF_LatCLat_RH	"	"	L→C→L	Right	"
MoMo_LowF_CLatC_RH	"	"	C→L→C	"	"
MoMo_HighF_LatCLat_LH	"	High	L→C→L	Left	"
MoMo_HighF_CLatC_LH	"	"	C→L→C	"	"
MoMo_HighF_LatCLat_RH	"	"	L→C→L	Right	"
MoMo_HighF_CLatC_RH	"	"	C→L→C	"	"

Half of the data from this experiment are presented in the current chapter with the focus on MOR with an initial stationary phase (StMo) and its comparison to sounds that have no movement (StSt). This way, in the current chapter we focus on MOR as it is used in other studies with the mechanism of release-of-adaptation following an initial stationary phase (Krumbholz et al. 2007). The other half of the data are presented in the next chapter that will focus on MOR with an initial moving phase (MoMo) and its comparison to sounds that have movement only in the beginning (MoSt). Thus, we will investigate the effects on MOR when the initial adaptation is a moving stimulus (MoMo) and whether frequency-specific effects are evident.

Data Recording and Analysis

The EEG data were recorded using actiCAP Standard-2 (Brain Products GmbH, Gilching, Germany) at a sampling rate of 500 Hz. The recording consisted of 58 Ag/AgCl active electrodes placed on the scalp according to the international 10-20 system (Towle et al. 1993). The horizontal EOG (hEOG) was recorded by placing 2 electrodes next to the eyes and the vertical EOG (vEOG) was recorded by placing 2 electrodes above and below the left eye; both placements were at the level of the retina. The ground electrode was placed in the middle of the forehead just above the nose and the online reference electrode was placed at the tip of the nose. Two additional electrodes were placed on the mastoids. Impedances were kept below 5 kΩ. Data were amplified by NeuroScan Synamps amplifier and filtered online with 0.1-100 Hz bandpass filter.

The EEG data were pre-processed with Brain Vision Analyzer 2.0 (Brain Products GmbH, Gilching, Germany). The data were digitally band-pass filtered (IIR, cut-off frequencies 0.1 40Hz; slope 48 dB/octave) and re-referenced to the average of 60 electrodes (58 scalp channels and 2 mastoids; implicit reference included in the re-reference). Then they underwent ocular correction using the Gratton and Coles procedure (Gratton et al. 1983). The epochs were set at -200 ms to 1100 ms according to stimulus onset (200 ms baseline period). Those which exceeded an amplitude of ±100 μV were excluded from further analysis. They were then baseline corrected for the period of 200 ms before motion onset and averaged between the trials.

Statistical analysis

The data recorded for the StSt and StMo conditions were analysed in a 2X2X2 design: Frequency range (lf, hf), hemifield presentation (left, right), location of stationary part (lateral, central;Fig. 1).

The statistical analysis was done using IBM SPSS Statistics for Windows, version 22.0 (IBM Corporation, Armonk, New York, USA). For the EEG results, the presence of MOR was calculated by creating difference waves between the StSt and StMo conditions: StMo minus StSt. To define time windows for extracting the components, first a visual inspection of the scalp distribution maps per grand average of each condition was conducted to detect the electrodes of maximum MOR activity. These electrodes were F1, Fz, F2, FC1, FC2, and Cz. Then, the maximum local negativity and positivity (for the cN1 and the cP2, respectively) of these electrodes was calculated from the grand averages of all conditions, and their peak latency was extracted and used to calculate an overall time average for the cN1 and the cP2. Using this average, we defined 95 ms-width time windows, which we used to extract component information from difference waves of individual data. The cN1 peaks were defined as the maximal local negativity between 170-265 ms after motion onset and the cP2 peaks were defined as the maximal local positivity between 326-421 ms after motion onset on the difference waves for a subset of electrodes where MOR was present for most conditions: F1, Fz, F2, FC1, FC2, Cz. The latency of these peaks was also extracted and analysed. Both extractions were done automatically.

6X2X2X2 repeated-measures ANOVAs (rm ANOVA) were carried out on the extracted values (latency and amplitude separately for cN1 and cP2) with the factors: electrode (F1, Fz, F2, FC1, FC2, Cz), frequency range (low, high), starting position (lateral, central) and hemifield of presentation (left, right), with p values set at 0.05. For all ANOVAs, if the assumption of sphericity was violated, then Greenhouse-Geisser correction was applied. For all ANOVAs, multiple comparisons were reported with Bonferroni correction.

To confirm that any amplitude differences observed on MOR based on the frequency range did not result as a general effect of the frequency range on the whole waveform, post hoc tests were carried out on the individual data. Firstly, the N1 and the P2 components were detected on the StSt and StMo stimuli as the maximal local negativity and positivity, respectively, within 0-500 ms after stimulus onset, and the peak-to-peak amplitude information was extracted (referred to as N1-P2). Then, they were entered in a 6X2X2X2 rm ANOVA with the factors condition (StSt, StMo), hemifield (left, right), starting position (lateral, central), frequency (low, high) and electrode (Fz, FC1, FC2, Cz, F1, F2) with p values set at 0.05. This test confirmed that there were no significant differences between the two conditions and so their N1-P2 information was averaged. We also calculated, the peak-to-peak data between the cN1 and the cP2 components isolated with the difference waves of the StMo stimuli (referred to as cN1-cP2). Then, the averaged N1-P2 data of the stationary stimulus part were used to normalize the cN1-cP2 data of the moving stimulus part by dividing the latter with the former: for example, cN1-cP2 / N1-P2 (the normalized data are referred to as NcN1-cP2). Thus, the resulting amplitude information of the motion responses was not influenced by the overall waveform response to the stimuli. The normalized amplitudes were then entered in a 6X2X2X2 rm ANOVA with factors hemifield (left, right), starting position (lateral, central), frequency (low, high) and electrode (Fz, FC1, FC2, Cz, F1, F2) with p values set at 0.05. If the assumption of sphericity was violated, then Greenhouse-Geisser correction was applied. Multiple comparisons were reported with Bonferroni correction.

The recorded behavioural measures were reaction times (RT) and hit rates (HR). RT of correct responses to StMo were averaged for each participant with the factors frequency range (low-high), initial position (lateral-central) and hemifield of presentation (left-right) and entered in a rm ANOVA for statistical analysis. Incorrect and missed responses were excluded from RT analysis. HR was analysed separately for correct, incorrect and missed responses for each stimulus of the StMo. rm ANOVAs were performed on HR with the factors hemifield (left-right), position (lateral-central), frequency (low, high) and type of response (correct, wrong, missed) for statistical analysis.

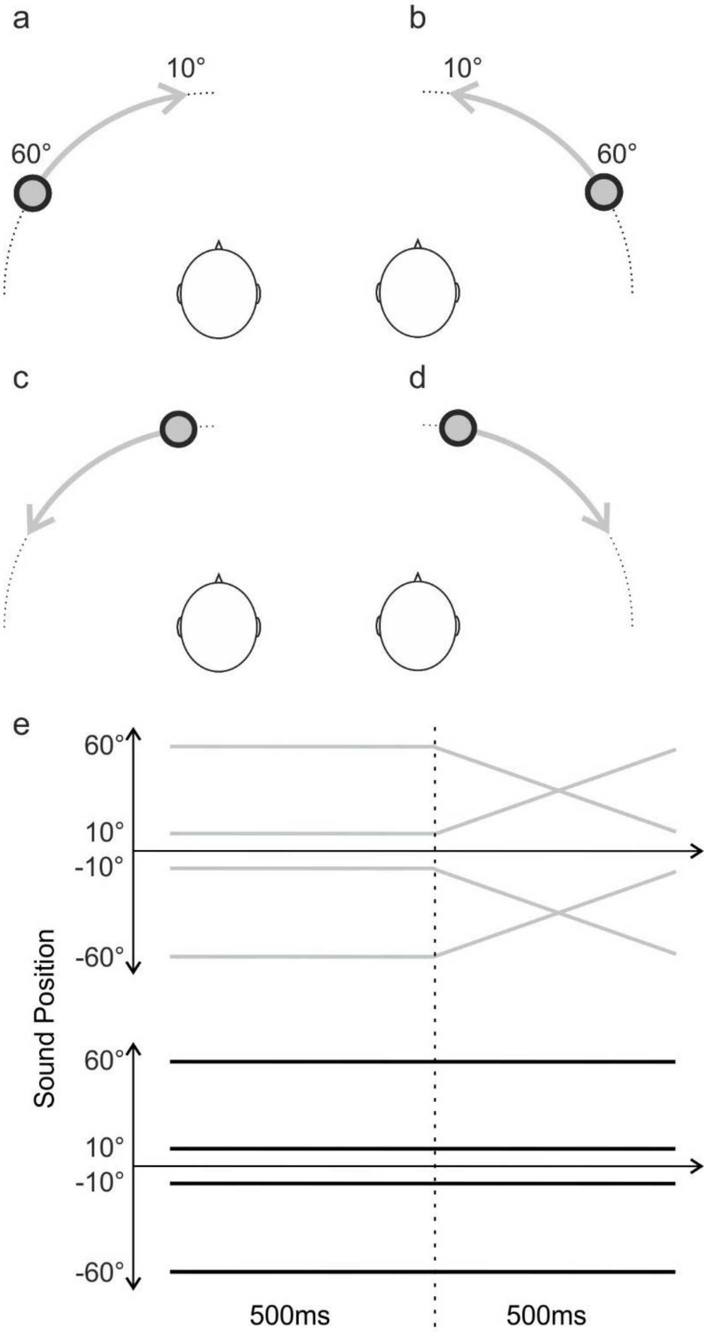


Fig. 1 *a -d*: Graphic representation of stimulus conditions. In each hemifield, there were 2 types of StMo (grey trajectories) stimuli: one started from a near-central position (grey circles at *b* and *d*) and the other one at a lateral position (grey circles at *a* and *c*), where they remained for 500 ms. At 500 ms the stimuli started to move to the other location within the same hemifield and the motion lasted 500 ms (grey arrows). The black circles correspond to the positions of the StSt stimuli used to create the difference waves. In each StSt condition, the sound was played for 1000 ms in one position. All stimuli illustrated here were presented at low (300-1200 Hz) and high (2000-8000 Hz) frequency. Left hemifield: *a*, *b*. Right hemifield: *c*, *d*. *e*: Schematic of the stimulus time course. Each stimulus (StMo and StSt) consisted of a 500-ms stationary sound. The StMo stimuli (grey lines in the upper illustration) then travelled for 500 ms to the other location within the same hemifield. In the left hemifield, the sound moved from -10° to -60° and vice versa. In the right hemifield,

the sound moved from 10° to 60° and vice versa. The StSt stimuli (black thick lines in the lower illustration) remained for another 500 ms at the same location. In the left hemifield, the locations were either -10° or -60° and in the right hemifield, the locations were either 10° or 60°.

4.3 Results

Behavioural Results

The behavioural rm ANOVA on RT focused on the time the participants needed to identify whether a change in the type of sound within a stimulus occurred or not. Means (M) RTs are expressed in seconds (s). Participants were faster to respond to high-frequency noise stimuli than to low-frequency noise stimuli, $p < 0.05$, $F(1,13) = 6.125$ ($M=0.327$ and $M=0.365$, respectively, with absolute p value $p=0.028$) (Fig. 2). Also, frequency had a different effect on participants' RT depending on the hemifield the sounds were presented. For sounds presented in the left hemifield, the analysis yielded $M=0.334$ and $M=0.333$ for lf and hf, respectively. For sound stimuli presented in the right hemifield, participants were significantly faster for hf than for lf stimuli, ($M=0.321$ and $M=0.395$, respectively). This is supported by the interaction between sound frequency and hemifield, $p<0.05$, $F(1,13)=21.765$ (absolute p value $p=0.001$).

The analysis of HR focused on participants' accuracy differences between the different conditions. Participants responded overall correctly to all stimuli, $p<0.05$, $F(2, 26) =145$ ($M=52.3$, $M=3.7$, $M=4.02$ for correct, incorrect and missed responses respectively, with absolute p value $p=0.001$).

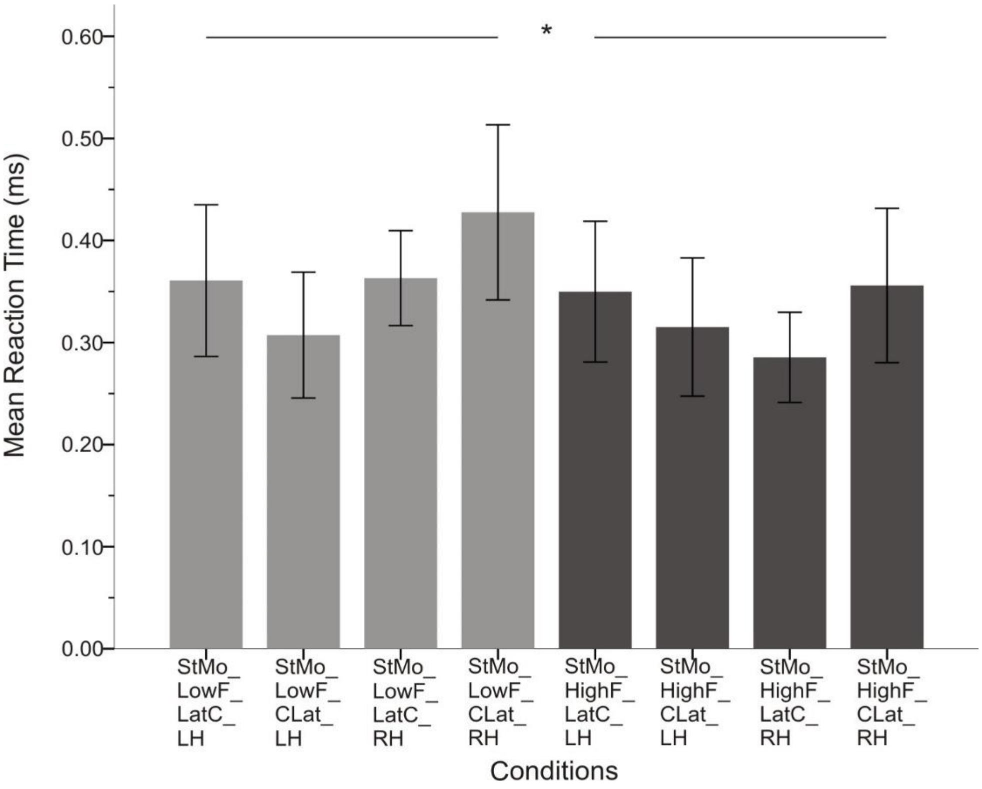


Fig. 2 Faster RTs as a response to high frequency stimuli, p<0.05. Presented here are the mean RTs with the error bars indicating the standard error of the mean (SEM). The grey bars show the low frequency conditions and the black bars show the high frequency conditions. From left to right, the respective conditions of the grey bars are: lateral position left hemifield (StMo_LowF_LatC_LH), central position left hemifield (StMo_LowF_CLat_LH), lateral position right hemifield (StMo_LowF_LatC_RH) and central position right hemifield (StMo_LowF_LatC_LH); and the respective conditions of the black bars are: lateral position left hemifield (StMo_HighF_LatC_LH), central position left hemifield (StMo_HighF_CLat_LH), lateral position right hemifield (StMo_HighF_LatC_RH) and central position right hemifield (StMo_HighF_LatC_LH).

Electrophysiological results

In all different stimulus conditions, the onset of the signal elicited the typical triphasic AEP between 50 ms and 200 ms (P1, N1, and P2; Fig. 3 and Supplementary material Fig. 1) referred to as the energy-onset response (EOR; Chait et al. 2004). In StSt conditions, in which the sound remained in one position for 1000 ms, the signal went back to baseline where it remained until the end of the signal that gave rise to offset effects. In StMo conditions, however, in which there was onset of motion at 500 ms, the usual response MOR was elicited and it is evident in the raw data as well as the difference waves. The MOR started with a characteristic negative deflection (cN1), followed by a positive deflection (cP2; Krumbholz et al. 2007; Fig. 3 and Fig. 4). The cP1, which sometimes precede the cN1, was not observed under the current experimental conditions (Krumbholz et al. 2007). At 1000 ms post onset, a small positivity was evident, which is characteristic of the offset of the sound and varies with the preceding potentials and stimulus duration (Hillyard and Picton 1978). The difference waves for StMo minus StSt (Fig. 4) showed that the EOR for this response remained around the midline, showing no evidence of differences between conditions, whereas MOR was prominent.

In Fig. 4A the difference waves at the Fz electrode are shown for 2 of the conditions of the low- and high-frequency ranges, respectively, with the moving part directed from a lateral-to-central position in the left hemifield (LowF_LatC_LH and HighF_LatC_LH, respectively). The scalp maps (Supplementary material 1h) show the signal distribution for the cN1 and the cP2. A detailed inspection of the topography of the difference waves revealed that the cN1 and cP2 of MOR were evident at frontal sites as well as frontocentral, central and anterior-frontal sites, in line with previous literature (Getmann and Lewald 2012; Grzeschik et al. 2010, 2013; Krumbholz et al. 2007). This led to the selection of 6 electrodes for further analysis: F1, Fz, F2, FC1, FC2 and Cz.

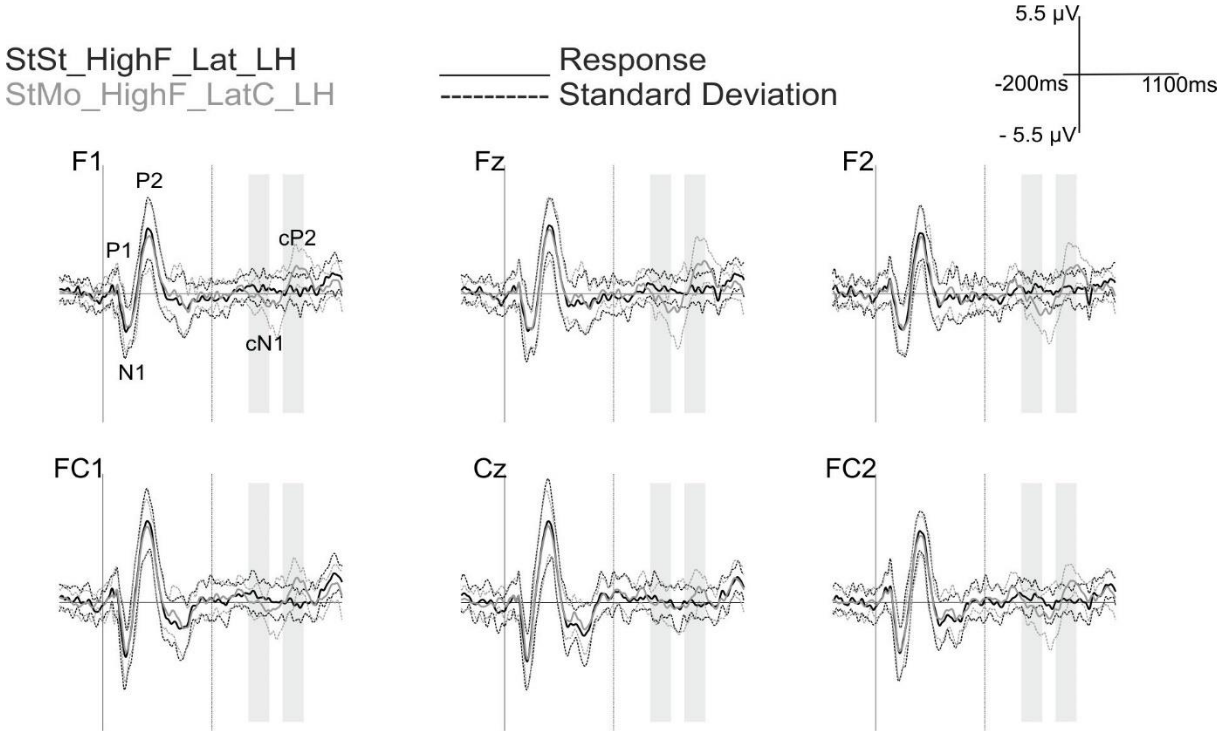


Fig. 3 Grand averages of raw waveforms. Raw waveforms of hf StSt presented at the lateral position in the left hemifield (black line) overlaid with hf StMo presented at a lateral position in the left hemifield (grey line). Solid lines show the waveforms and dashed lines show the standard deviation. The early peaks (grey arrows) represent the initial reaction of the system to the stimuli, with the complex P1-N1-P2 and are similar in both conditions. With the onset of motion (vertical dotted line), there is the usual MOR with prominent peaks cN1 and cP2 (highlighted with grey boxes). For the grand averages of the other conditions, please see Supplementary Material 1.

Using the difference waves, we chose a timeframe for the cN1 and cP2 components (a description for how we chose the time frame is given in the Methods). Then, we extracted the peak amplitudes and peak latencies as the maximal local negativity and positivity, respectively, and analysed them with a rm ANOVA. Here amplitudes are in µV. The interactions between all possible combinations of the two factors are reported with Bonferroni correction. Overall, conditions with high-frequency noise stimuli elicited higher amplitudes than low-frequency stimuli at the deflections of MOR (Fig. 4a-d). At the cN1, this effect reached significance, $p<0.05$, $F(1,13) = 8.6$ (high-frequency stimuli: $M=-1.418$; low-frequency stimuli: $M=-1.173$, with absolute p value $p=0.012$). However, at the cP2, there was a tendency toward the same effect, but this did not reach significance, $p=0.26$ (high-frequency stimuli: $M=1.351$; low-frequency stimuli: $M=1.165$; Fig. 5a). Also, stimuli starting at a central position elicited overall higher amplitudes at the cN1, $p<0.05$, $F(1,13) =4.73$ (with absolute p value $p=0.049$), than those starting at a lateral position (starting at a central position: $M=-1.477$; starting at a lateral position: $M=-1.114$), confirming previous data that showed more prominent responses from outward- than inward-moving sound stimuli (Magezi and Krumbholz 2010; Supplementary Material Fig. 2).

The analysis of the latencies of MOR did not yield the same frequency effect. However, cN1 was elicited faster after sounds were presented in the right than in the left hemifield, $p<0.05$, $F(1,13) =8.75$ ($M=710.274$, $M=718.571$, respectively, with absolute p value $p=0.011$). Also, cP2 was elicited faster when sounds were presented at a central position and moved to a lateral position than vice versa, $p<0.05$, $F(1,13) =6.245$ ($M=868.810$, $M=878.298$ respectively, with absolute p value $p=0.027$).

The post hoc analysis of the N1-P2 between the StSt and StMo stimuli showed that there was no significant difference between conditions, p=0.469. The results of the rm ANOVA on the NcN1-cP2 confirmed the results of the rm ANOVA analysis on the non-normalized data extracted from the difference waves, showing a significant frequency effect, p<0.05, F(1, 13) =9.24 (M=0.44, M=0.56, for low and high frequency, respectively, with absolute p value p=0.009; Fig. 5b). In addition, the NcN1-cP2 was higher for sounds starting at a central position is comparison to those starting at a lateral position, p <0.05, F(1, 13) =13.6 (M=0.4, M=0.62, for lateral and central starting position, respectively, with absolute p value p=0.003). Moreover, there was a weak but significant effect of the electrodes, p<0.05, Greenhouse-Geisser, F(5, 65) =3.5 (M=0.55, M=0.45, M=0.5, M=0.4, M=0.54, M=0.6, for Fz, FC1, FC2, Cz, F1 and F2, respectively, with absolute p value p=0.045), showing a propagation of the signal from frontal electrodes towards frontocentral and central electrodes. Also, there was an interaction between the hemifield of sound presentation and the electrodes, p< 0.05, Greenhouse-Geisser, F(2.7, 35) =3.54 (in the left hemifield: M=0.54, M=0.45, M=0.5, M=0.4, M=0.54, M=0.63; in the right hemifield: M=0.55, M=0.46, M=0.5, M=0.4, M=0.54, M=0.53; for Fz, FC1, FC2, Cz, F1 and F2, respectively, with absolute p value p=0.028).

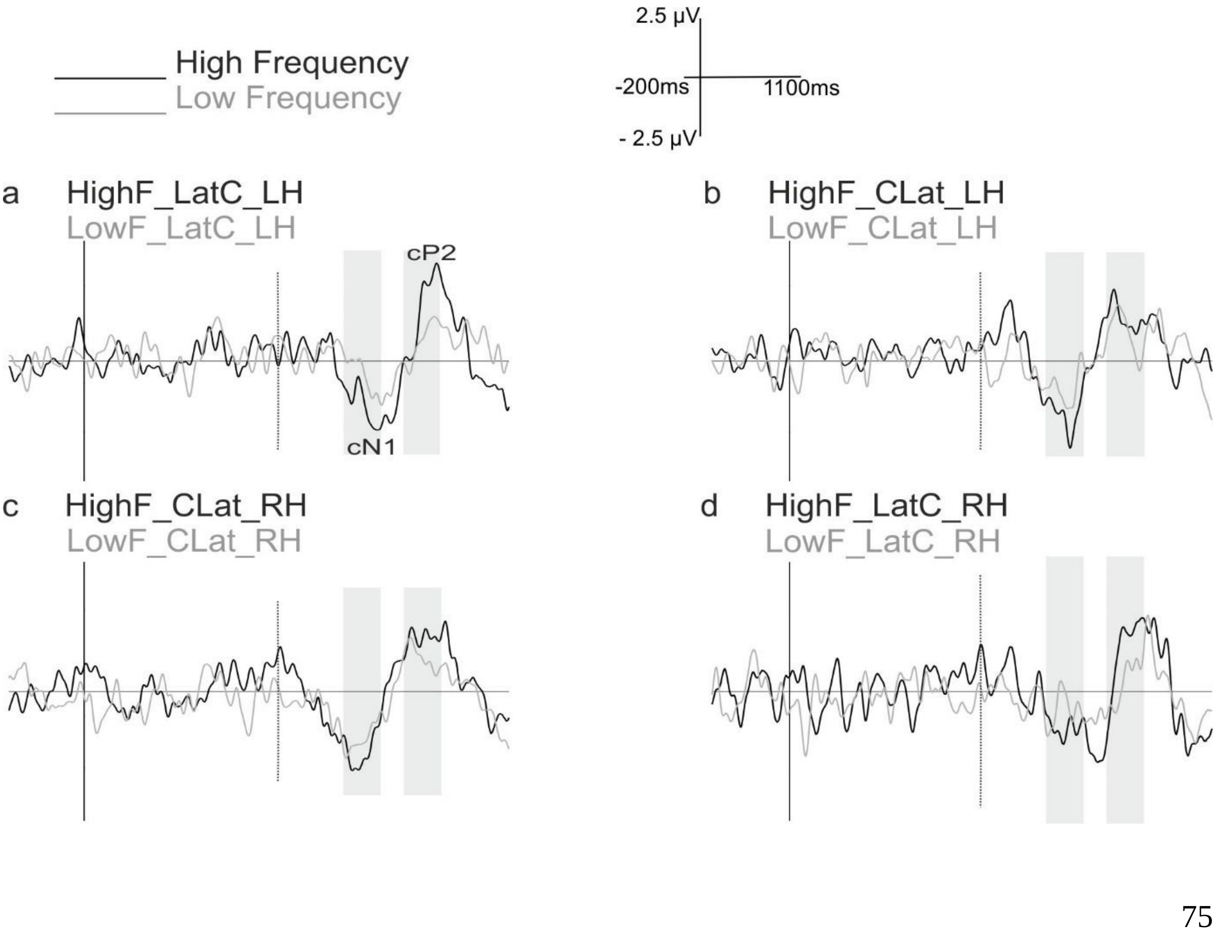
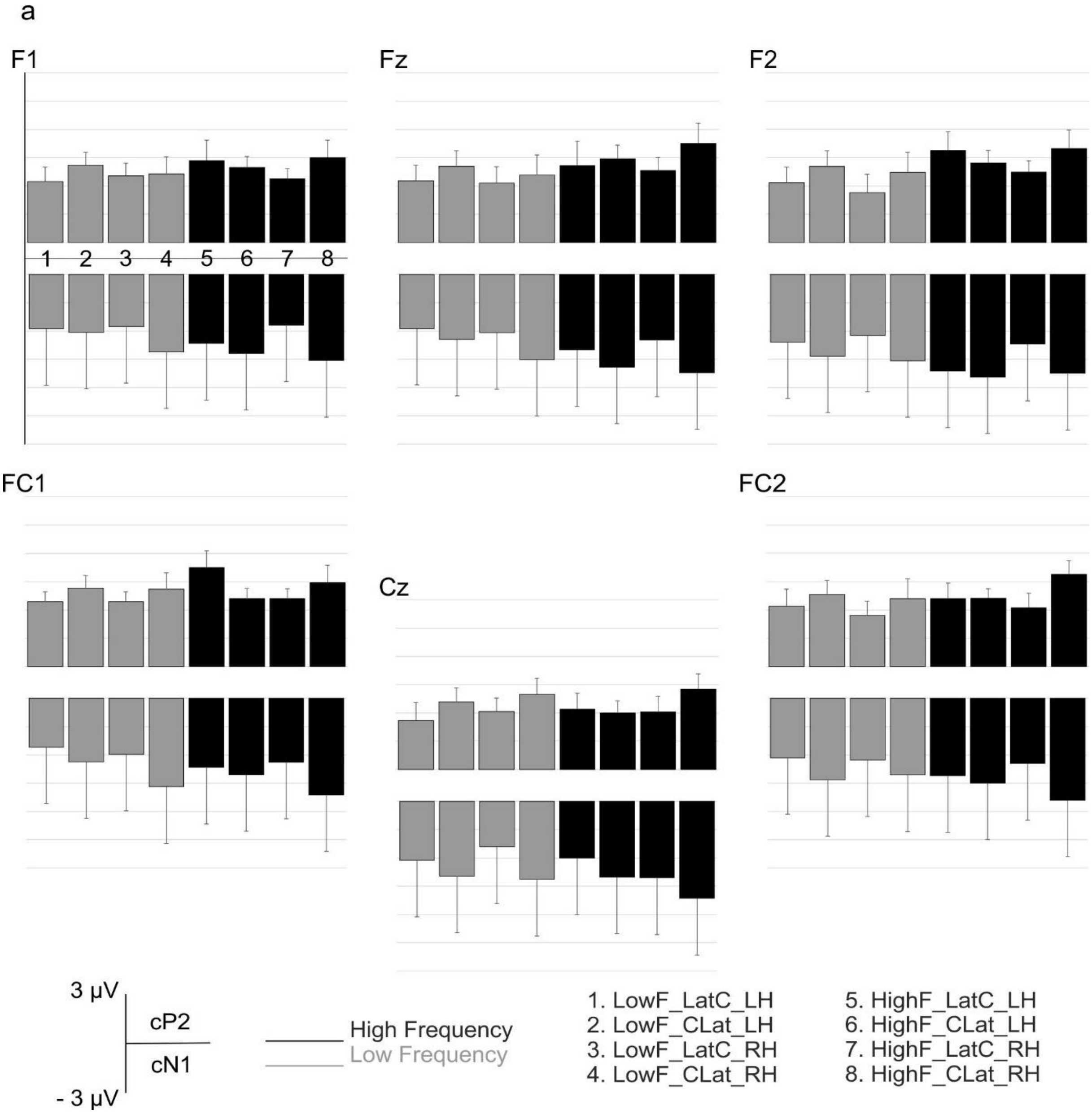


Fig. 4 *a-d:* Overlay of low- and high-frequency conditions of averaged difference waves at the Fz electrode, showing higher cN1 amplitude for the hf than lf. Grey lines represent responses from low frequency stimuli. Black lines represent responses from high frequency stimuli. *a*: Difference waves of low and high frequency stimuli presented at a lateral position in the left hemifield (LowF_LatC_LH and HighF_LatC_LH). *b*: Difference waves of low and high frequency stimuli presented at a central position in the left hemifield (LowF_CLat_LH and HighF_CLat_LH). *c*: Difference waves of low and high frequency stimuli presented at a central position in the right hemifield (LowF_CLat_RH and HighF_CLat_RH). *d*: Difference waves of low and high frequency stimuli presented at a lateral position in the right hemifield (LowF_LatC_RH and HighF_LatC_RH). Vertical solid lines show the onset of the stimuli. Vertical dotted lines show the onset of motion.



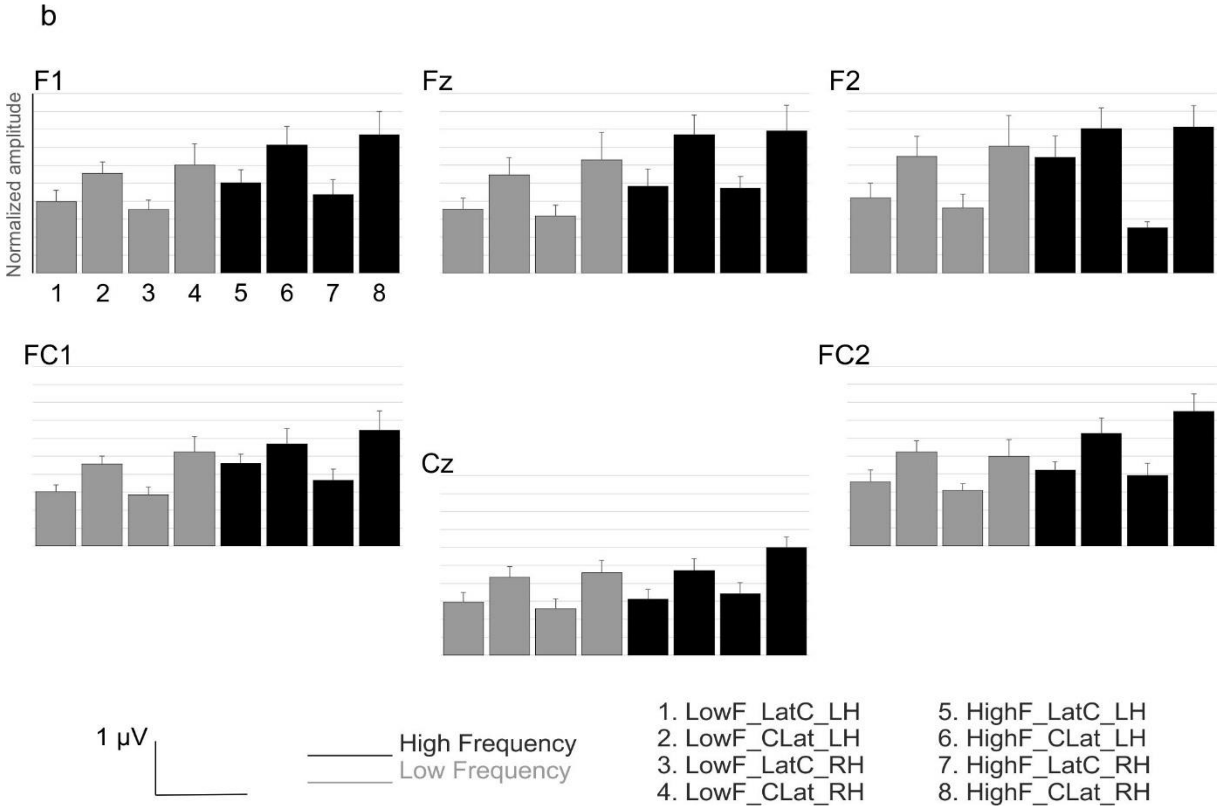


Fig. 5 a: Bar graph showing the results of the rm ANOVA on the amplitudes of the cN1 and cP2 from the difference waves for all the stimuli. Grey bars represent low frequency stimuli. Black bars represent high frequency stimuli. Error bars show the SEM. The electrode sets are arranged to represent the electrode positions on the cap. *b*: Bar graph showing the results of the post hoc comparisons between the normalized amplitudes, NcN1-cP2, of high and low frequency stimuli. Grey bars represent low frequency stimuli. Black bars represent high frequency stimuli. Error bars show the SEM.

4.4 Discussion

The current experiment aimed to analyse the different components of the MOR with respect to the frequency range of the stimuli and the direction of motion, i.e. to the front vs. to the side. In agreement with published data (Getzmann 2009, 2011; Getzmann and Lewald 2010, 2012; Grzeschik et al. 2013; Krumbholz et al. 2007), signal strength of MOR had its maximum at frontocentral electrode positions. The fact that MOR showed longer latencies than previously reported (Grzeschik et al. 2016; Kreitewolf et al. 2011; Krumbholz et al. 2007) might be due to the slower velocity of the motion stimuli (100°/sec) than the one employed in earlier studies. Also, due to the nature of MOR, which is elicited when motion begins after initial adaptation, the amplitude and latency of its components inevitably depends on how effective the adaptation is. This adaptation occurs during the stationary part of the stimulus. In the literature, the duration of the stationary part varies from 700-1000 ms (700 ms: Getzmann 2009, 2011; Getzmann and Lewald 2010, 2011, 2012; Kreitewolf et al. 2011; 1000 ms: Getzmann and Lewald 2014; Grzeschik et al. 2013, 2016; Krumbholz et al. 2007). In our experiment, however, we used a stationary part with a briefer duration (500 ms), which certainly provided less effective adaptation than a longer one. These differences might have caused smaller amplitudes and longer latencies of the different MOR components (Getzmann 2009). Still, despite these differences, MOR signals had the typical multi-peak morphology described earlier.

The current study showed that the amplitude of the cN1 component of MOR differs depending on the frequency range of Gaussian noise stimuli; high frequencies elicited higher amplitudes than low frequencies. This corresponds to psychophysical measurements; participants responded faster to sounds with high frequency. A congruence of significantly faster RT along with modulations in MOR has been reported previously: (i) in studies testing different stimulus velocities (Getzmann 2009), (ii) in studies comparing sound sources moving in the free field and moving with HRTF instead of ILD/ITD motion (Getzmann and Lewald 2010), and (iii) in studies comparing the effects of the position of the initial stationary part of the paradigm used for MOR elicitation (Geztmann and Lewald 2011). For the electrophysiological responses, high frequency stimuli might resemble a stimulus that allocates attention much faster than low frequency stimuli, causing the currently observed signal differences. Also, results from a recent study (Warren et al. 2016) showed that the propagation of the signal along the basilar membrane differs significantly for low frequency and high frequency sounds, because in the case of low frequency sounds, the signal follows an exponential decline of motion amplitude. This slow propagation of low frequency processing could also be reflected by frequency-specific temporal differences in neuronal activation throughout the ascending auditory pathway up to the cortex. The current results could also be explained by the contribution of high-frequency tuned neurons to the mechanisms eliciting MOR. Poirier and colleagues (2017) provided evidence for motion-specific areas in the primate cortex. It is possible that these areas receive more contribution from high-frequency tuned neurons and their signal is what contributes to the current results. Moreover, since the elicitation of MOR is based on a mechanism of adaptation by having the initial stationary part, the current data could arise from neurons that respond to stimuli with a low frequency range and are highly adapted from the initial stationary part, leading to lower

MOR amplitude. Another possible explanation is that MOR elicitation could be based on motion-sensitive cortical areas (Poirier et al. 2017) that respond strongly to stimuli with a high frequency range.

Additionally, the evidence from the literature for the stimulus frequency effect on the AEP varies. Early cortical responses of the auditory system have larger amplitudes and longer latencies as a response to low compared to high frequency stimuli (N1: Jacobson et al. 1992; MMN: Wunderlich and Cone-Wesson 2001; P150: Verkindt et al. 1994), but this is not true for components that are elicited earlier (brainstem and middle-latency components; Na, P1: Woods et al. 1995) or later (P3: Sugg and Polich 1995). To begin with, Jacobson and colleagues (1992) recorded EEG responses to low-, mid- and high-frequency tones and found the highest N1 amplitudes following low-frequency stimulation. They attributed their results to the position of the low frequency response areas that are located closer to the scalp than the high frequency ones, thus possibly contributing to stronger signal on the EEG. Also, basal areas of the basilar membrane may contribute to low frequency sounds and not high frequency sounds. Verkindt and colleagues (1994) analysed scalp current densities of the P2 response from EEG recordings from participants listening to tones of several frequencies and found that the P150 response that contributes to the P2 had decreasing amplitude with increasing stimulus frequency. Furthermore, Wunderlich and Cone-Wesson (2001) measured ERPs while participants were presented with consonant-vowel tokens and words on several speech frequency ranges in an oddball paradigm. They found decreasing amplitude and latency on the N1 component and the mismatched-negativity (MMN) response, and decreasing P2 amplitude with increasing stimulus frequency. However, the results look different for studies that investigated brainstem and middle-latency AEP and later cortical AEP (Sugg and Polich 1995; Woods et al. 1993; Woods et al. 1995). Woods and colleagues (1993) presented tones of either 250 Hz or 4kHz to participants while measuring their brainstem, middle, and long-latency AEP. They found that the high frequency tones evoked overall earlier responses than the low frequency ones. Later in another study (Woods et al. 1995), they replicated their results, and in addition found overall higher amplitudes as a response to high frequency tones, which reached significance for the Na and P1 components. Similarly, Sugg and Polich (1995) presented tones of several intensities and frequencies to participants, and their results showed that the P3 component had higher amplitude as a response to high frequency tones.

The behavioural results in this study could be explained by a tendency of overall quicker motor reactions to high frequency than low frequency sounds. For example, research in primates showed that scream vocalizations with a high frequency range, are perceived as ‘alarm calls’ and need to be well-localized by other members of the species (Tian and Rauschecker 2004). Also, high-frequency vocalizations in spider monkeys are produced by victims in attack situations and signalize danger (Ordonez-Gomez et al. 2015; for a recent review see Gruber and Grandjean 2017). It seems that the quick localization of high frequency sounds has an evolutionary purpose: if the data from the aforementioned studies hold for humans as well, then humans are more likely to react faster to high frequency sounds because they are not very common in our environment and could signal danger. Due to the timing of MOR elicitation, the respective signal occurs in the course of the stimulus presentation but the behavioural response is given after the end of the stimulus. This is because MOR is recorded over cortical areas that overlap with a motor response that is elicited during button pressing. Thus, the participants were instructed to reply at the end of the stimulus. The fact that this behavioural response is given at least 1000 ms after stimulus onset excludes the possibility that there is an immediate association between the behavioural response time and the stimulus frequency. The detection of frequency by the auditory system is made with the onset of sound, and thus there is a time delay until the button press. Still, the data show faster reaction times for high frequency stimuli.

Considering the presumed complexity of auditory motion processing, we also aimed to assess differences in the frequency effects with respect to the hemifield of sound presentation. It became evident that cN1 was elicited with shorter delays if the sounds were presented in the right compared to the left hemifield. These results are in line with the right ear advantage reported in otoacoustic emissions and linguistic studies; participants are more accurate to report to stimuli presented to their right ear in linguistic tests involving dichotic listening (Jerger and Martin 2004; McFadden 1993). Also, a study by Krumbholz and colleagues (2007) must be considered, in which it was hypothesized that the response to motion in each cortical hemisphere might be based on the contribution of different afferent connections to the two cortical hemispheres. A dipole analysis following the generation of motion percepts through earphone stimulation showed that the response in the right cortical hemisphere to rightward motion was faster than the response of the left hemisphere to leftward motion. The results were interpreted to be in line with findings from callosotomy studies, which showed that the left cortical hemisphere receives information through callosal projections to process auditory space

(Haumann et al. 2005). Also, the “neglect model” of hemispheric dominance in the auditory system suggests that stimuli presented in the right hemifield create equal right cortical responses as those presented in the left hemifield, but they also create stronger left cortical responses than those presented in the left hemifield (Teshiba et al. 2013). Thus, this extra boost in cortical response from the left hemisphere could account for the results in this study. The high temporal precision of EEG is specifically suited to disclose such a hemispheric advantage in signal processing.

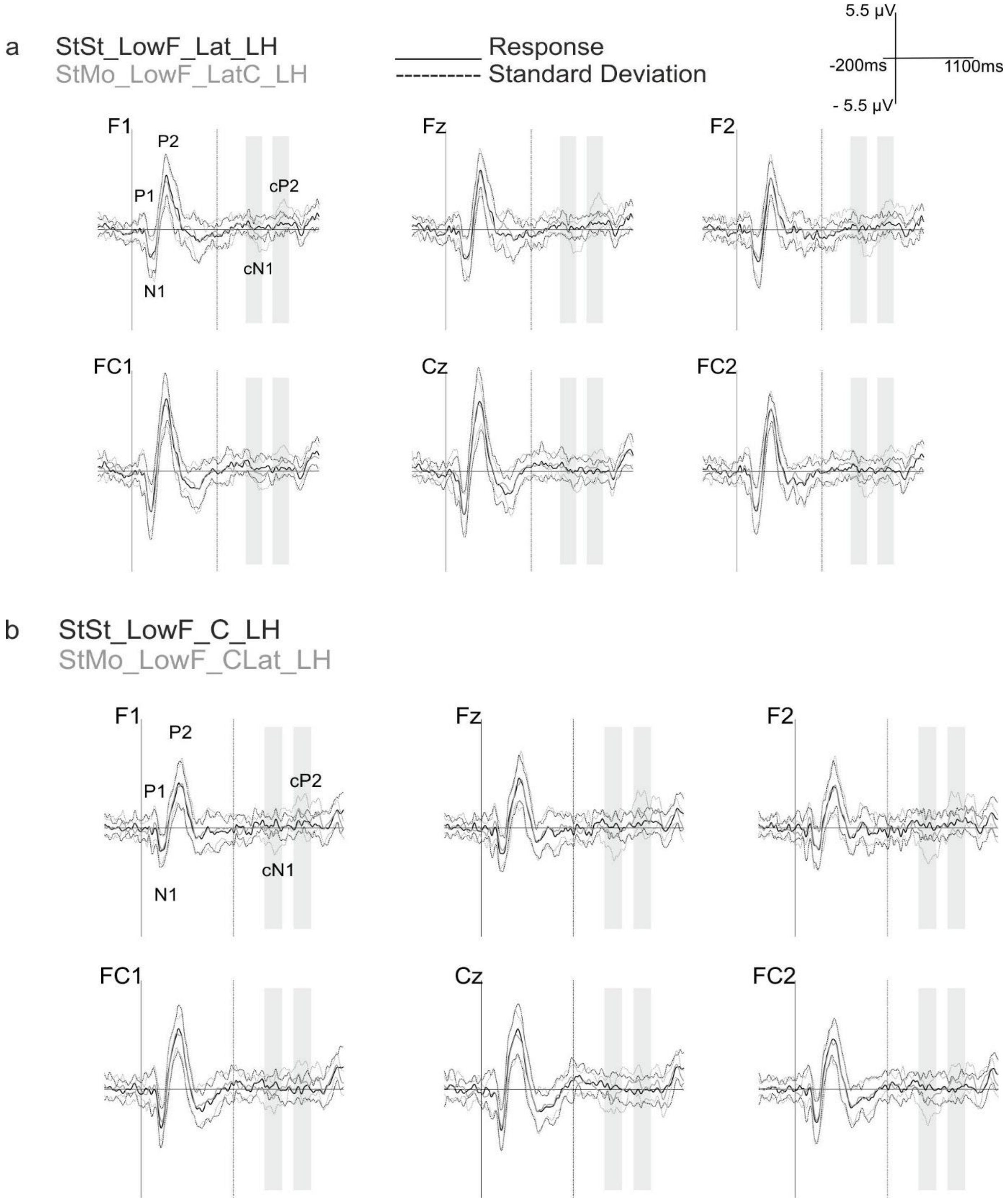
Presently, the effects of the inward and outward motion on MOR in both cortical hemispheres were examined by presenting the respective stimuli in each acoustical hemifield. Generally, the latency of cP2 was shorter and the amplitude of cN1 was higher for outward than inward motion. Previously, Magezi and Krumbholz (2010) showed that outward motion elicits stronger MOR amplitudes. The authors interpreted their results as evidence for the opponent-channel theory of cortical auditory space representation. The results of the present study are in line with this interpretation.

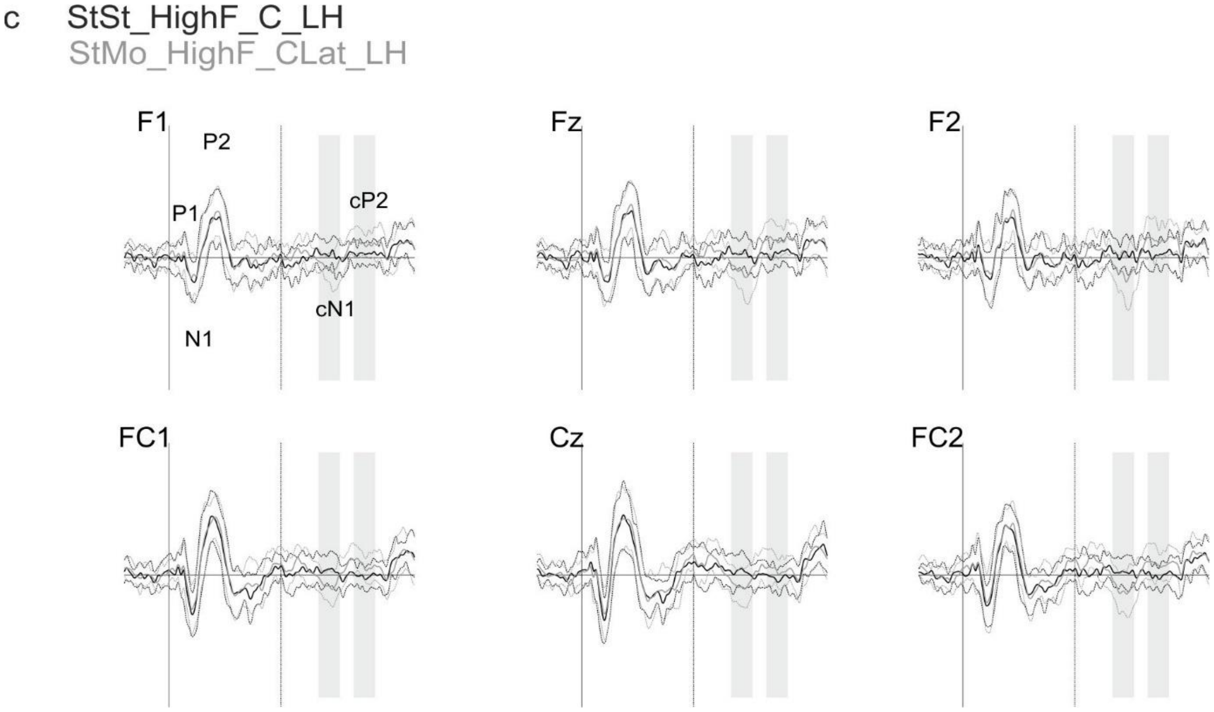
In previous studies, the cN1 was modulated based on the hemifield in which the sound motion occurred and the cP2 was modulated by the motion direction (Getzmann 2011) and attention (Kreitewolf et al. 2011). These studies (including Getzmann and Lewald 2010) suggest a modular organization of MOR, with cN1 and cP2 indicating different cortical processing stages of auditory motion. Also, higher motion velocity elicited stronger and faster cN1 and cP2 peaks (Getzmann 2009). Motion direction and velocity are inherent to motion stimuli, but the top-down attention to motion is not. In the current study, we showed that another feature of the signal, namely the frequency, has an effect on the amplitude of the cN1. This remains somewhat puzzling, since frequency is not inherent to motion. Still the frequency effect was specifically observed for MOR. Perhaps the propagation of the signal of frequency processing is reflected in MOR because of the mechanism of release-of-adaptation under which MOR is elicited (Krumbholz et al. 2007; Warren et al. 2016). When compared to the energy-onset response (EOR), MOR is lower in amplitude, and the dominance of energy onset in the AEP might be the reason why a respective frequency effect is not seen in the initial phase of the AEP.

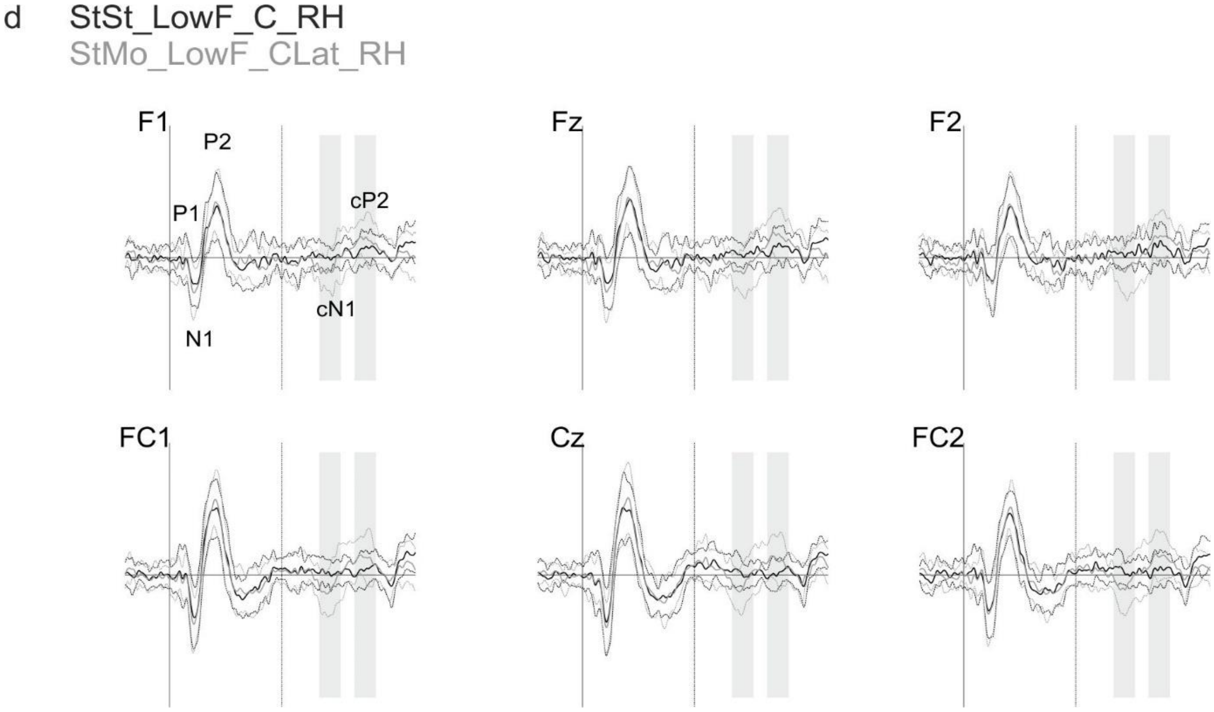
A possible confound in our results is that the noise stimuli were identical in ra within each frequency range, and this could have led to habituation effects. But as Butler showed, stimuli of different frequencies activate different neuronal populations (1968), and along with stimuli presented in different locations (1972), this decreases the effect of habituation. Also, previous studies used stimulus frequency change as a method of dishabituation or recovery (Polich 1989; Ritter et al. 1968). In the current experiment, we presented a variety of stimuli differing in type, frequency spectrum, spatial hemisphere, spatial location and direction of motion, without having two identical stimuli following each other. Thus, it is unlikely that the current data reflect any habituation effects.

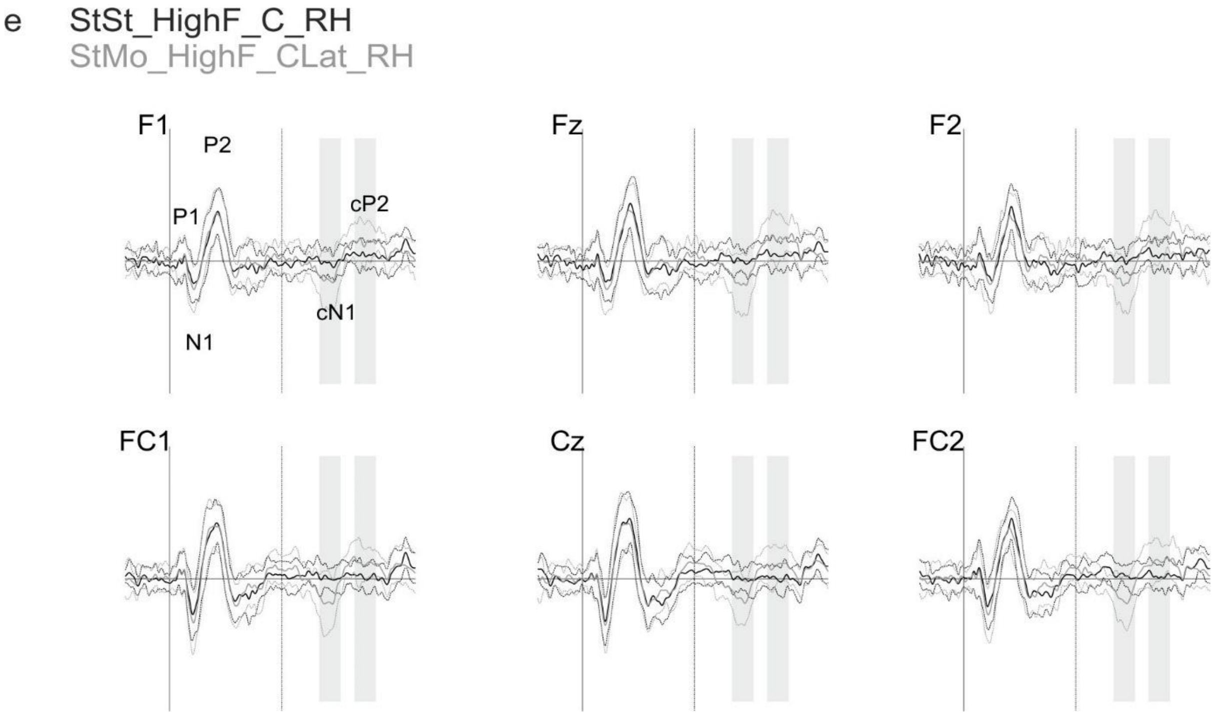
All in all, the current results show that late motion-related AEPs are modulated by the frequency of the stimulus. This is in agreement with an earlier study by Schröger (1996), who showed that the processing of ITDs and ILDs (which are indicative of localizing different frequency domains [Rayleigh 1907; Stevens and Newman 1936]) follow segregated paths. This reveals the importance of careful stimulus design in studied that aim to measure fine details of cortical auditory processing.

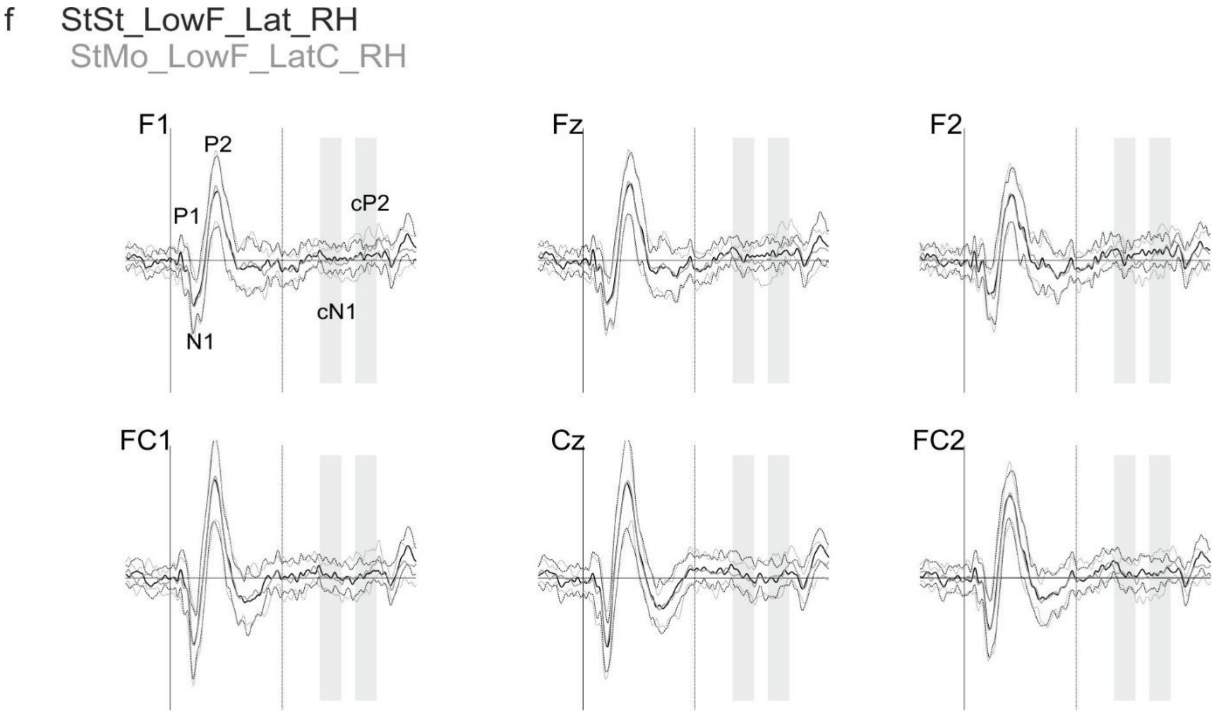
4.5 Supplementary material




c
StSt_HighF_C_LH
StMo_HighF_CLat_LH
F1
P2
P1
cP2
cN1
N1
Fz
F2
FC1
Cz
FC2


d
StSt_LowF_C_RH
StMo_LowF_CLat_RH
F1
P2
P1
cP2
cN1
N1
Fz
F2
FC1
Cz
FC2


e
StSt_HighF_C_RH
StMo_HighF_CLat_RH
F1
P2
P1
cP2
cN1
N1
Fz
F2
FC1
Cz
FC2


f
StSt_LowF_Lat_RH
StMo_LowF_LatC_RH
F1
P2
P1
cP2
cN1
N1
Fz
F2
FC1
Cz
FC2

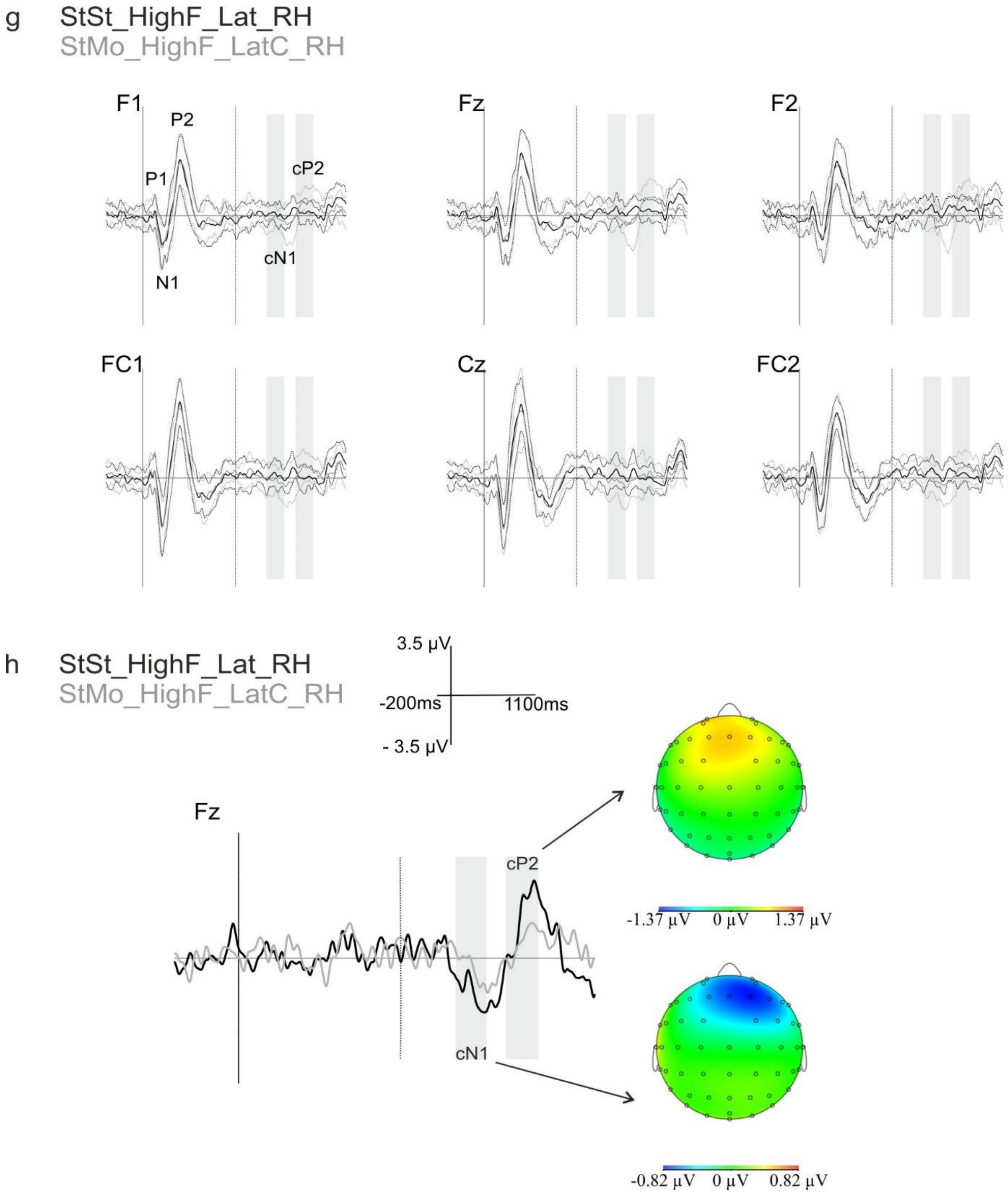


Fig. 6 Grand averages of raw waveforms. *a*: Raw waveforms of StSt of low frequency presented at the lateral position in the left hemifield (black line) overlaid with StMo of low frequency presented at a lateral position in the left hemifield (grey line). *b*: Raw waveforms of StSt of low frequency presented at the central position in the left hemifield (black line) overlaid with StMo of low frequency presented at a central position in the left hemifield (grey line). *c*: Raw waveforms of StSt of high frequency presented at the central position in the left hemifield (black line) overlaid with StMo of high frequency presented at a central position in the left hemifield (grey line). *d*: Raw waveforms of StSt of low frequency presented at the central position in the right hemifield (black line) overlaid with StMo of low frequency

presented at a central position in the right hemifield (grey line). *e*: Raw waveforms of StSt of high frequency presented at the central position in the right hemifield (black line) overlaid with StMo of high frequency presented at a central position in the right hemifield (grey line). *f*: Raw waveforms of StSt of low frequency presented at the lateral position in the right hemifield (black line) overlaid with StMo of low frequency presented at a lateral position in the right hemifield (grey line). *g*: Raw waveforms of StSt of high frequency presented at the lateral position in the right hemifield (black line) overlaid with StMo of high frequency presented at a lateral position in the right hemifield (grey line). Solid lines show the waveforms and the dashed lines show the standard deviation. The early peaks (grey arrows) represent the initial reaction of the system to the stimuli, with the complex P1-N1-P2 and are similar in both conditions. With the onset of motion (vertical dotted lines), there is the usual motion-onset response evident by the prominent peaks cN1 and cP2 (highlighted with grey boxes). The subtraction StMo minus StSt of this particular pair, led to the difference waves shown in Fig.4. *h*: In black, difference wave of high frequency stimuli presented at a lateral position in the left hemifield (HighF_LatC_LH) and in grey, difference wave of low frequency stimuli presented at a lateral position in the left hemifield (LowF_LatC_LH). Grey boxes show the cN1 and cP2 of the motion-onset response. The scalp maps show the signal distribution for the time frame of cN1 and cP2. In *h*, during the initial stimulus presentation, when the characteristics of the acoustic stimuli are identical, there is no response evident in the difference wave, whereas with the onset of motion, two distinct peaks (one negative and one positive) are present. Vertical solid line shows the onset of the stimuli and the vertical dotted line shows the onset of motion.

Table 1 Table with descriptive information for the latency of cN1 of all conditions and analyzed electrodes. N = number of participants, Std. Error = standard error, Std. Deviation = standard deviation.

Electrode_Condition	**N**	**Mean (ms)**	**Std. Error**	**Std. Deviation**
Fz_LowF_LatC_LH	14	727.29	7.106	26.586
Fz_HighF_LatC_LH	14	719.71	7.643	28.596
Fz_LowF_CLat_LH	14	715.71	6.62	24.768
Fz_HighF_CLat_LH	14	710.14	5.947	22.25
Fz_LowF_CLat_RH	14	701.14	5.393	20.179
Fz_HighF_CLat_RH	14	715.29	6.701	25.074
Fz_LowF_LatC_RH	14	707.43	6.263	23.435
Fz_HighF_LatC_RH	14	716.43	6.259	23.418
FC1_LowF_LatC_LH	14	711.57	7.893	29.532
FC1_HighF_LatC_LH	14	714	7.116	26.626
FC1_LowF_Clat_LH	14	713.71	5.152	19.277
FC1_HighF_Clat_LH	14	722.43	4.864	18.199
FC1_LowF_Clat_RH	14	705.71	6.108	22.855
FC1_HighF_Clat_RH	14	707.14	5.514	20.632
FC1_LowF_LatC_RH	14	714.71	6.437	24.085
FC1_HighF_LatC_RH	14	722.14	5.998	22.443
FC2_LowF_LatC_LH	14	722.43	7.481	27.991

FC2_HighF_LatC_LH	14	726	7.672	28.705
FC2_LowF_Clat_LH	14	716	4.858	18.179
FC2_HighF_Clat_LH	14	714	5.84	21.853
FC2_LowF_Clat_RH	14	702.14	4.972	18.605
FC2_HighF_Clat_RH	14	705.43	5.678	21.245
FC2_LowF_LatC_RH	14	703.29	7.388	27.642
FC2_HighF_LatC_RH	14	713	6.382	23.881
Cz_LowF_LatC_LH	14	712.57	7.163	26.803
Cz_HighF_LatC_LH	14	715.29	6.392	23.918
Cz_LowF_Clat_LH	14	711.14	6.655	24.902
Cz_HighF_Clat_LH	14	721.14	5.039	18.855
Cz_LowF_Clat_RH	14	700.43	5.837	21.841
Cz_HighF_Clat_RH	14	702.86	5.303	19.841
Cz_LowF_LatC_RH	14	715.14	6.793	25.416
Cz_HighF_LatC_RH	14	720.57	6.834	25.57
F1_LowF_LatC_LH	14	726.14	7.102	26.573
F1_HighF_LatC_LH	14	722	8.36	31.28
F1_LowF_Clat_LH	14	714	4.786	17.906
F1_HighF_Clat_LH	14	717.43	5.898	22.069
F1_LowF_Clat_RH	14	697.29	5.573	20.853
F1_HighF_Clat_RH	14	711.86	7.206	26.964
F1_LowF_LatC_RH	14	713.71	6.166	23.07
F1_HighF_LatC_RH	14	710.86	6.854	25.645
F2_LowF_LatC_LH	14	722.43	7.169	26.823
F2_HighF_LatC_LH	14	739.43	4.851	18.152
F2_LowF_Clat_LH	14	714.43	6.732	25.191
F2_HighF_Clat_LH	14	716.71	7.337	27.452
F2_LowF_Clat_RH	14	711.86	6.545	24.488
F2_HighF_Clat_RH	14	721.14	6.31	23.609
F2_LowF_LatC_RH	14	711.86	7.921	29.638
F2_HighF_LatC_RH	14	715.14	5.91	22.115

Table 2 Table with descriptive information for the amplitude of cN1 of all conditions and analyzed electrodes. N = number of participants, Std. Error = standard error, Std. Deviation = standard deviation.

Electrode_Condition	**N**	**Mean (µV)**	**Std. Error**	**Std. Deviation**
Fz_LowF_LatC_LH	14	-0.9533981	0.17837917	0.667433727
Fz_HighF_LatC_LH	14	-1.3329714	0.15240063	0.570230938
Fz_LowF_CLat_LH	14	-1.1475319	0.23571383	0.881960405
Fz_HighF_CLat_LH	14	-1.6386224	0.25088282	0.938717564
Fz_LowF_CLat_RH	14	-1.5056827	0.23011622	0.861016036
Fz_HighF_CLat_RH	14	-1.7390669	0.25624772	0.958791167
Fz_LowF_LatC_RH	14	-1.0290754	0.19515371	0.730198325
Fz_HighF_LatC_RH	14	-1.1593929	0.15728684	0.588513479
FC1_LowF_LatC_LH	14	-0.8609036	0.17215177	0.644132954
FC1_HighF_LatC_LH	14	-1.2195417	0.15404528	0.576384662
FC1_LowF_Clat_LH	14	-1.1208786	0.16506843	0.617629509

FC1_HighF_Clat_LH	14	-1.3504879	0.13391489	0.501063638
FC1_LowF_Clat_RH	14	-1.5605914	0.25869361	0.967942848
FC1_HighF_Clat_RH	14	-1.707226	0.25938718	0.970537954
FC1_LowF_LatC_RH	14	-0.9882686	0.1907373	0.713673638
FC1_HighF_LatC_RH	14	-1.1298037	0.22357153	0.836528059
FC2_LowF_LatC_LH	14	-1.0474349	0.13160652	0.492426491
FC2_HighF_LatC_LH	14	-1.3682376	0.17256818	0.645691004
FC2_LowF_Clat_LH	14	-1.4381152	0.20233753	0.757077702
FC2_HighF_Clat_LH	14	-1.4990549	0.26538232	0.992969712
FC2_LowF_Clat_RH	14	-1.3550311	0.24152341	0.903697833
FC2_HighF_Clat_RH	14	-1.8002846	0.1860906	0.696287248
FC2_LowF_LatC_RH	14	-1.0886851	0.16127514	0.603436305
FC2_HighF_LatC_RH	14	-1.1515439	0.2289722	0.856735518
Cz_LowF_LatC_LH	14	-1.0404691	0.22269007	0.833229948
Cz_HighF_LatC_LH	14	-1.0027704	0.2304393	0.862224921
Cz_LowF_Clat_LH	14	-1.324777	0.16986944	0.635593226
Cz_HighF_Clat_LH	14	-1.337369	0.18768265	0.702244155
Cz_LowF_Clat_RH	14	-1.3809348	0.26846382	1.004499622
Cz_HighF_Clat_RH	14	-1.7156547	0.17070451	0.638717805
Cz_LowF_LatC_RH	14	-0.8031333	0.1804858	0.675316038
Cz_HighF_LatC_RH	14	-1.3536626	0.26980203	1.009506743
F1_LowF_LatC_LH	14	-0.9573564	0.16775388	0.627677558
F1_HighF_LatC_LH	14	-1.2197777	0.1587116	0.593844443
F1_LowF_Clat_LH	14	-1.021361	0.21891104	0.819090113
F1_HighF_Clat_LH	14	-1.3952368	0.2068812	0.774078577
F1_LowF_Clat_RH	14	-1.3700921	0.23641932	0.884600109
F1_HighF_Clat_RH	14	-1.5219299	0.22913408	0.857341222
F1_LowF_LatC_RH	14	-0.9193201	0.17886492	0.669251229
F1_HighF_LatC_RH	14	-0.8951856	0.15472327	0.578921477
F2_LowF_LatC_LH	14	-1.1944026	0.16748036	0.626654112
F2_HighF_LatC_LH	14	-1.7082179	0.24043209	0.899614507
F2_LowF_Clat_LH	14	-1.4429811	0.24938905	0.933128373
F2_HighF_Clat_LH	14	-1.8125069	0.25883272	0.968463358
F2_LowF_Clat_RH	14	-1.5230751	0.26321058	0.984843824
F2_HighF_Clat_RH	14	-1.7489651	0.2059706	0.770671418
F2_LowF_LatC_RH	14	-1.0740502	0.20025368	0.749280651
F2_HighF_LatC_RH	14	-1.233852	0.17541183	0.656330965

Table 3 Table with descriptive information for the latency of cP2 of all conditions and analyzed electrodes. N = number of participants, Std. Error = standard error, Std. Deviation = standard deviation.

Electrode_Condition	**N**	**Mean (ms)**	**Std. Error**	**Std. Deviation**
Fz_LowF_LatC_LH	14	878.14	5.489	20.538
Fz_HighF_LatC_LH	14	885.14	6.158	23.041
Fz_LowF_CLat_LH	14	858.86	5.802	21.707
Fz_HighF_CLat_LH	14	860	6.401	23.949
Fz_LowF_CLat_RH	14	875	6.348	23.752
Fz_HighF_CLat_RH	14	879.43	6.541	24.475

Fz_LowF_LatC_RH	14	878.14	6.089	22.783
Fz_HighF_LatC_RH	14	875.71	5.764	21.567
FC1_LowF_LatC_LH	14	876.29	7.186	26.888
FC1_HighF_LatC_LH	14	893.57	5.296	19.817
FC1_LowF_Clat_LH	14	863	5.675	21.235
FC1_HighF_Clat_LH	14	873	6.471	24.214
FC1_LowF_Clat_RH	14	857.86	6.704	25.084
FC1_HighF_Clat_RH	14	868.86	6.197	23.188
FC1_LowF_LatC_RH	14	870	8.297	31.043
FC1_HighF_LatC_RH	14	877.86	6.114	22.877
FC2_LowF_LatC_LH	14	878.86	7.241	27.092
FC2_HighF_LatC_LH	14	887.29	7.526	28.16
FC2_LowF_Clat_LH	14	870.29	7.091	26.531
FC2_HighF_Clat_LH	14	868.14	7.024	26.282
FC2_LowF_Clat_RH	14	866.14	6.443	24.108
FC2_HighF_Clat_RH	14	880.71	5.605	20.97
FC2_LowF_LatC_RH	14	867.71	6.767	25.321
FC2_HighF_LatC_RH	14	874.43	6.768	25.325
Cz_LowF_LatC_LH	14	861	7.424	27.777
Cz_HighF_LatC_LH	14	882.14	5.214	19.509
Cz_LowF_Clat_LH	14	872.43	6.992	26.161
Cz_HighF_Clat_LH	14	868.43	6.019	22.521
Cz_LowF_Clat_RH	14	862	5.398	20.199
Cz_HighF_Clat_RH	14	872.29	6.427	24.049
Cz_LowF_LatC_RH	14	871.14	6.351	23.764
Cz_HighF_LatC_RH	14	872.86	5.947	22.253
F1_LowF_LatC_LH	14	883.29	6.464	24.187
F1_HighF_LatC_LH	14	882.86	6.981	26.121
F1_LowF_Clat_LH	14	858.14	5.206	19.477
F1_HighF_Clat_LH	14	865.14	7.382	27.621
F1_LowF_Clat_RH	14	865.14	6.403	23.958
F1_HighF_Clat_RH	14	872.57	7.08	26.492
F1_LowF_LatC_RH	14	875.86	7.234	27.066
F1_HighF_LatC_RH	14	876	5.742	21.483
F2_LowF_LatC_LH	14	878.71	6.309	23.607
F2_HighF_LatC_LH	14	892.57	4.489	16.796
F2_LowF_Clat_LH	14	866.71	7.482	27.996
F2_HighF_Clat_LH	14	870	5.058	18.925
F2_LowF_Clat_RH	14	880.57	8.078	30.224
F2_HighF_Clat_RH	14	876.71	5.577	20.867
F2_LowF_LatC_RH	14	871.86	5.635	21.085
F2_HighF_LatC_RH	14	887.71	6.307	23.597

Table 4 Table with descriptive information for the amplitude of cP2 of all conditions and analyzed electrodes. N = number of participants, Std. Error = standard error, Std. Deviation = standard deviation.

Electrode_Condition	N	Mean (µV)	Std. Error	Std. Deviation

Fz_LowF_LatC_LH	14	1.09427957	0.27460888	1.02749235
Fz_HighF_LatC_LH	14	1.36023493	0.43284979	1.61957561
Fz_LowF_CLat_LH	14	1.34852036	0.27573293	1.03169815
Fz_HighF_CLat_LH	14	1.48474136	0.23686102	0.88625279
Fz_LowF_CLat_RH	14	1.19458814	0.35133952	1.3145921
Fz_HighF_CLat_RH	14	1.74807207	0.36655703	1.37153081
Fz_LowF_LatC_RH	14	1.05140393	0.28945604	1.08304532
Fz_HighF_LatC_RH	14	1.27126157	0.22814304	0.8536331
FC1_LowF_LatC_LH	14	1.14338521	0.17829072	0.66710277
FC1_HighF_LatC_LH	14	1.74815536	0.29645231	1.10922297
FC1_LowF_Clat_LH	14	1.3899515	0.21798353	0.8156197
FC1_HighF_Clat_LH	14	1.19703357	0.18130753	0.67839064
FC1_LowF_Clat_RH	14	1.36427657	0.29358999	1.09851316
FC1_HighF_Clat_RH	14	1.48389557	0.30456615	1.1395822
FC1_LowF_LatC_RH	14	1.1480375	0.17617155	0.65917357
FC1_HighF_LatC_RH	14	1.198957	0.17252864	0.64554307
FC2_LowF_LatC_LH	14	1.06550371	0.30319598	1.13445549
FC2_HighF_LatC_LH	14	1.19889407	0.27471957	1.0279065
FC2_LowF_Clat_LH	14	1.27588971	0.2489238	0.93138757
FC2_HighF_Clat_LH	14	1.20849943	0.16904039	0.63249123
FC2_LowF_Clat_RH	14	1.20070871	0.34723298	1.29922685
FC2_HighF_Clat_RH	14	1.62789471	0.24307754	0.90951286
FC2_LowF_LatC_RH	14	0.90606836	0.24902879	0.93178043
FC2_HighF_LatC_RH	14	1.04071021	0.25123835	0.94004782
Cz_LowF_LatC_LH	14	0.86497279	0.32299962	1.20855393
Cz_HighF_LatC_LH	14	1.06882429	0.27617259	1.0333432
Cz_LowF_Clat_LH	14	1.190092	0.2482357	0.92881296
Cz_HighF_Clat_LH	14	0.99878993	0.21485063	0.80389745
Cz_LowF_Clat_RH	14	1.32513936	0.29215565	1.09314635
Cz_HighF_Clat_RH	14	1.42316764	0.26944231	1.00816082
Cz_LowF_LatC_RH	14	1.02748779	0.23377174	0.87469375
Cz_HighF_LatC_RH	14	1.02123707	0.27415363	1.02578896
F1_LowF_LatC_LH	14	1.08118814	0.25654927	0.95991949
F1_HighF_LatC_LH	14	1.44946307	0.36191732	1.3541706
F1_LowF_Clat_LH	14	1.366416	0.22643021	0.84722428
F1_HighF_Clat_LH	14	1.32598021	0.19651907	0.73530704
F1_LowF_Clat_RH	14	1.21039957	0.30248091	1.13177994
F1_HighF_Clat_RH	14	1.501459	0.30946834	1.15792451
F1_LowF_LatC_RH	14	1.17701314	0.22421787	0.83894646
F1_HighF_LatC_RH	14	1.12915807	0.17669669	0.66113849
F2_LowF_LatC_LH	14	1.06148393	0.27315082	1.02203678
F2_HighF_LatC_LH	14	1.62596129	0.32870887	1.22991597
F2_LowF_Clat_LH	14	1.34719971	0.27603671	1.03283479
F2_HighF_Clat_LH	14	1.40831936	0.22319152	0.83510621
F2_LowF_Clat_RH	14	1.23789029	0.35649957	1.33389925
F2_HighF_Clat_RH	14	1.66335943	0.32698151	1.22345277
F2_LowF_LatC_RH	14	0.88205879	0.32818129	1.22794195
F2_HighF_LatC_RH	14	1.24849743	0.1958528	0.73281408

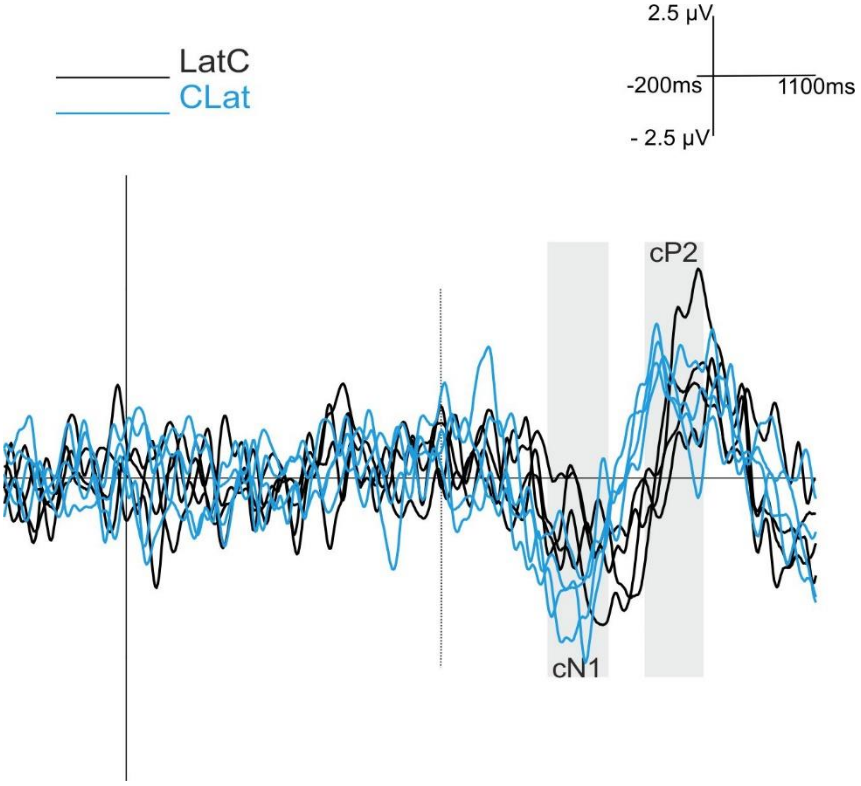


Fig. 7 Overlay of all conditions of averaged difference waves at the Fz electrode, showing higher cN1 amplitude for the CLat than the LatC. Black lines represent responses from LatC stimuli. Blue lines represent responses from CLat stimuli. Vertical solid line shows the onset of the stimuli. Vertical dotted line shows the onset of motion.

Chapter 5. *Study 3*: The motion-onset response is depended on the initial adaptation from the delayed motion-onset stimuli.

Abstract

The current study analyses data from the whole experimental session described in Chapter 4. The focus is on the motion-onset response elicited using a stimulus paradigm that has an initial moving part instead of a stationary part and how it compares to the response elicited using the usual stimulus paradigm. These responses were also analysed based on the effects of stimulus frequency. Difference waves were calculate to extract the motion responses and analyses were employed on the peak amplitude and latencies of the cN1 and cP2. The results showed that the stimulus paradigm with the initial moving part elicited a response that resembles the MOR but has lower amplitude. In addition, the effects of stimulus frequency evident from the previous analysis (Chapter 4) apply for data collected from both conditions, with high frequency stimuli eliciting higher MOR amplitude than low frequency stimuli.

5.1 Introduction

Localization of stationary sound sources on the azimuthal level is based on the interaural time and level differences (ITD and ILD, respectively) (Schnupp et al. 2011) and localization of dynamic sound sources is based on the dynamic aspects of these differences (Middlebrooks and Green 1991). The electrophysiological signature to auditory motion is the cortical motion-onset response (MOR) (Krumbholz et al. 2007). This response is based on a release-from-adaptation mechanism created from a delayed motion-onset stimulus and it is recorded from the electroencephalogram (EEG). This stimulus introduces an initial stationary sound, during which the usual deflections emerging from subcortical and cortical structures appear in the EEG signal: P1, N1, P2 (energy-onset response or EOR). After some time the signal returns to baseline. The subsequent onset of motion gives rise to the MOR, which resembles the EOR, with the peaks cP1, cN1 and cP2 (Getzmann 2009; Getzmann & Lewald 2010; Getzmann 2011; Getzmann & Lewald 2012; Grzeschik et al. 2013; Kreitewolf et al. 2011; Krumbholz et al. 2007). Each of these peaks has been associated with different aspects of auditory motion characteristics. Getzmann (2011) presented to participants rightward and leftward delayed motion-onset stimuli within each hemifield in a freefield study while measuring EEG data. His results showed that whilst both prominent peaks of the MOR, cN1 and cP2, exhibited a pattern of contralaterality, each was associated with different sound features. The cN1 was contralateral to the hemisphere of motion presentation, having for example higher amplitude at the right hemisphere when the sound was presented at the left hemifield and vice versa. The cP2 was contralateral to the direction of motion, having higher amplitude at the right hemisphere with the presentation of centrifugal motion direction in the left hemifield and vice versa. The author concluded that the cN1 is location-specific and the cP2 is direction-specific and these are evidence of a modular organization for the processing of auditory motion. In a later study (Getzmann & Lewald 2012), the MOR was compared to responses to abrupt and scatter spatial sound shifts after the initial stationary sound presentation. The EEG results showed earlier and stronger cP1, cN1 and cP2 peaks with the abrupt and scatter than those elicited from the motion sounds. The researchers suggested that the MOR is a processing mechanism for any spatial change that it is not limited to motion. Therefore, they concluded that the mechanism underlying the MOR could be related to the smooth release-from-adaptation of the spatially encoding neurons adapted with the initial stationary part and thus, the sudden change of location from the abrupt and scatter stimuli create sudden excitation to neurons responding to more lateral locations that leads to stronger EEG responses, giving it the name “spatial change response” (SCR).

The frequency-specificity of the processing of auditory motion comes from a behavioural study on the auditory motion aftereffect (aMAE) (Dong et al. 2000). The researchers presented pairs of adaptor and probe auditory moving stimuli and the participants were requested to indicate the direction of the probe stimuli. In one of their experiments, they presented frequency differences between the adaptor and the probe and measured the magnitude of the adaptation effect. Their results showed that the adaptation was stronger when the pairs had the same frequency spectrum and it was even stronger for low than middle or high frequencies. If, however, the spectrum of the pairs differed, then stronger adaptation was observed when the adaptor had lower frequency spectrum than the probe, than the other way around. The researchers interpreted it as evidence for more high than low frequency tuning of the motion-sensitive neurons. This conclusion is supported from behavioural and electrophysiological data on the MOR from our lab (Chapter 4). The presentation of delayed motion-onset stimuli elicited higher amplitudes at the cN1 peak of the MOR when these stimuli had high than low frequency. These results were also accompanied with faster reaction times after the high frequency stimuli. Since the underlying mechanism is a release-from-adaptation, we interpreted the results arising from a mechanism that adapts stationary sounds stronger when they have low than high frequency or a mechanism whose neurons responding to motion-onset are more high- than low-frequency tuned. Either mechanism could lead to stronger amplitudes on the MOR peaks, thus providing an explanation of the previous results, including those from Dong et al. (2000).

The current study made use of the data presented in our previous study (Chapter 4) and also the rest of the data collected from the same experiment. Whereas the previous study focused on data eliciting the MOR and analysing these data for the variables frequency spectrum, hemifield presentation and motion direction, the current study analyses the data only with regards to frequency and compares them with data having a delayed motion-onset after initial motion presentation. If the MOR mechanism is only a SCR mechanism, then the onset of motion after an initial moving phase (MoMo condition), should elicit a response that resembles the one elicited from the onset of motion after an initial stationary phase (StMo condition). Any differences between the two should be based on the rate of adaptation a stationary sound gives in comparison to that from a moving sound. In addition, if the neurons responding to auditory motion are indeed frequency-specific, then this effect should be the same in both StMo and MoMo conditions.

5.2 Method

Participants

The same participants of Chapter 4 are included in this study as well. A total of 14 healthy volunteers participated in the study (10 females) with a mean age of 24.9 years (13 right-handed, 1 ambidextrous) in exchange of small monetary compensation. All participants were informed about the scope of the experiment and had signed a written consent form for their participation. The experimental procedures were in agreement with the Declaration of Helsinki (World Medical Association 2000).

Setup

The experiment was conducted in complete darkness in a semi-anechoic chamber (40 m^2; Industrial Acoustics Company, Niederkrüchten, Germany) in which 47 loudspeakers (Visaton, FRS8 4 Ohm, Haan, Germany) with radius 2.35 m are arranged in an azimuthal, semi-circular plane spanning from 98° left to 98° right in the frontal hemifield. The distance between the loudspeakers was 4.3°. Each loudspeaker was equilibrated individually. For this, the transmission spectrum was measured using the a Bruel & Kjaer measuring amplifier (B&K 2610), a microphone (B&K 2669, pre-amplifier B&K 4190) and a real-time signal processor (RP 2.1, System3, Tucker Davis Technologies, TDT, Alachua, Florida, USA). For each loudspeaker, a calibration file was generated in Matlab 6.1 (The MathWorks Inc., Natick, Massachusetts, USA) and was later used to ensure flat spectra of the acoustic stimuli across the frequency range tested. The loudspeakers were covered with a sound transparent black curtain to prevent participants having visual location information of the loudspeakers' positions. The generation of the stimuli was digital using real-time TDT processors (RX8, System3) and controlled with custom-made scripts in MATLAB 7.5 (R2007b, The MathWorks Inc., Natick, Massachusetts, USA). An infrared camera was installed in this chamber to oversee participants during testing.

Stimuli

The stimuli used were the same as described in the *Stimuli* section of Chapter 4. The stimuli were 2 Gaussian noise signals, one at a low-to-moderate frequency range (300-1200 Hz) and one at high frequency range (2000-8000 Hz). The intensity was set on 50 dB sensation level (SL; above individual hearing threshold) with a rowing level +/- 3dB. For each participant, the individual hearing threshold for the low- and high-frequency noise bands was measured at the beginning of each session using a yes/no (heard/not heard) paradigm for stationary and moving sound stimuli.

Acoustic stimuli had a duration of 1000ms separated in two 500ms-components (without any temporal gap between them), either stationary or moving. Four different signal combinations were used: (i) stationary-stationary (StSt), (ii) stationary-moving (StMo), (iii) moving-stationary (MoSt) (Fig. 1 A-D), (iv) moving-moving (MoMo) (Fig. 1 E-H). For stationary stimulus components a single speaker was activated at a particular azimuthal position. Moving sounds were created by sequential activation of adjacent speakers with linear cross-fading adjusting the intensity of the output signal voltage of neighbouring speakers. Stationary sound stimuli were presented in one of the following locations (negative numbers show locations left from midline): -60°, -10°, +10°, +60°. The velocity of the moving sounds was invariantly set at 100°/sec. Sound presentation started at 10° or 60° in either hemifield. For StSt the sound source position remained unchanged. For moving sounds, either being the first and/or the second component of the 1000ms stimulus, the movement starting at +/-10° was laterally directed to +/-60° and the movement starting at +/-60° was frontally directed to +/-10°. This way sound stimuli being stationary or moving (and the latter moving either inward or outward) remained within the same hemifield.

Altogether 32 stimulus conditions were explored: 4 signal combinations (StSt, Stmo, MoSt, MoMo), 2 frequency ranges (lf, hf), both hemifields (left, right). Each stimulus combination was presented 60 times in a pseudo-random manner resulting in 1920 stimulus presentations. The interstimulus interval (ISI) varied between 900-1000 ms starting from the time point a participant gave her response (see below). Because of the long duration needed to collect a complete dataset from a single participant, data acquisition was divided into 2 sessions. Within each session 960 stimuli were presented. In each, short breaks were inserted every about 12 minutes.

Procedure

The procedure followed is the same as the one described in *Procedure* in Chapter 4. Participants were asked to sit in a comfortable chair in the middle of a semi-circular loudspeaker array in the semi-anechoic chamber. In order to keep the head in a steady position, they were instructed to keep their eyes on a fixation cross during the audiogram and the experiment. Both of these sessions were contacted in complete darkness.

The heard/not-heard audiogram was combined with a simple staircase paradigm used to determine participants' individual hearing threshold for stationary and moving sounds of low and high frequency. The stationary sounds were presented at 0° in front of the participants and the moving sounds spanned a trajectory from -30° (left from midline) to +30°, each lasting 1000ms, with an initial intensity of 62 dB SPL. Participants were instructed to press a left button on a response box to indicate that they have detected a sound (*heard* response) and the right button when they did not detect the sound (*not-heard* response). The stimulus' intensity was decreased by 5 dB for each *heard* response and increased by 5 dB for each *not-heard* response. The average of the stationary and moving noise stimuli of low frequency range defined the SL of the low frequency noise stimuli; the average of the stationary and moving noise stimuli of high frequency range defined the SL of the high frequency noise stimuli.

During the experiment, the participants were instructed to indicate by pressing the right button on a response box when the presented sound stimulus included a change in modality (stationary-to-moving or moving-to-stationary) and the left button if no such change occurred (stationary-stationary or moving-moving). The participants were asked to reply as quick and as correct as possible. During the analysis the reaction times (RT) were quantified, defined as the time from the end of the stimulus to the button press, only for the correct responses. If the answer was not given within a 1000ms time window, then the trial was noted as missed. Before starting the data acquisition, the participants were allowed to familiarize themselves with the task in a demonstration block.

In a previous chapter (Chapter 4), the data corresponding to the StMo and StSt were analysed to study the frequency-specificity of the motion-onset response (MOR) with a stationary sound as history, compared for frequency spectrum, motion direction and hemifield presentation. In the current chapter, the 960 stimuli corresponding to StSt and StMo are averaged based on their frequency range and the 960 stimuli corresponding to MoMo and MoSt are also averaged based on their frequency range to study how the frequency-specificity of the MOR changes based on history (stationary versus moving).

The mechanism of the motion-onset is the release-of –adaptation of the initial stationary phase (Krumbholz et al. 2007). If the initial adaptation is done with a moving stimulus, then we expect lower amplitude that usual on the electroencephalographic response.

Data Recording and Analysis

The EEG data were recorded using actiCAP Standard-2 (Brain Products GmbH, Gilching, Germany) at a sampling rate of 500 Hz. The recording consisted of 64 Ag/AgCl active electrodes placed on the scalp according to the international 10-20 system (Towle et al. 1993). The horizontal EOG (hEOG) was recorded by placing 2 electrodes next to the eyes and the vertical EOG (vEOG) was recorded by placing 2 electrodes above and below the left eye; both placements were at the level of the retina. The ground electrode was placed in the middle of the forehead just above the nation and the online reference electrode was placed at the tip of the nose. Two additional electrodes were placed on the mastoids. Impedances were kept below 5 kΩ.

The EEG data were pre-processed with Brain Vision Analyzer 2.0 (Brain Products GmbH, Gilching, Germany). The data were bandpass filtered (IIR, cut-off frequencies 0.1 40Hz; slope 48 dB/octave) and re-referenced to the average of 60 electrodes (58 scalp channels and 2 mastoids; implicit reference included in the re-reference). Then they underwent ocular correction using the Gratton and Coles procedure (Gratton et al. 1983).

The signal was averaged among all StSt, StMo, MoSt (Fig. 1 A-D) and MoMo (Fig. 1 E-H) of each frequency spectrum separately. The epochs were set at -200ms to 1100ms according to stimulus onset (200ms baseline period). Those which exceeded an amplitude of ±100 μV were excluded from further analysis. They were then baseline corrected for the period of 200ms before motion onset and averaged between the trials.

Statistical analysis

The statistical analysis was done using IBM SPSS Statistics for Windows, version 22.0 (IBM Corporation, Armonk, New York, USA). For the EEG results, the presence of the MOR was calculated by creating four difference waves: two between the StSt and StMo conditions, StMo minus StSt, for low and high frequency conditions, and another two between the MoSt and MoMo conitions, MoMo minus MoSt, for low and high frequency conditions. The cN1 peaks were defined as the maximal local negativity between 170-265 ms after motion onset and the cP2 peaks as the maximal local positivity between 326-421 ms after motion onset on the difference waves for the electrodes F1, Fz, F2, FC1, FC2 and Cz. The time frames and the electrodes were chosen to match the ones used in the previous analysis (Chapter 4). The latency of these peaks was also extracted and analysed. Both extractions were done automatically.

To investigate the frequency and history effects on the cN1, repeated-measure (rm) ANOVAs were carried out at the extracted values (latency and amplitude) with the factors electrode, frequency (low and high, LF and HF, respectively) and history (stationary and moving, S and M, respectively), with p values set at .05. Then, to investigate the frequency and history effects on the cP2, rm ANOVAs were carried out at the extracted values with the same factors and p value as the ones used for cN1. If the assumption of sphericity was violated, then Greenhouse-Geisser correction was applied. Multiple comparisons were reported with Bonferroni correction.

The recorded behavioural measures were reaction times (RT) and hit rates (HR). RT of correct responses to all StMo and MoMo were averaged for each participant for each frequency range (LF, HF) and entered in a repeated-measures ANOVA for statistical analysis, with the second factor history (S, M). Wrong and missed responses were excluded from RT analysis. HR were analysed separately for correct, wrong and missed responses for each stimulus condition: repeated-measures ANOVA was performed with the factors frequency (LF, HF) high), history (S, M) and type of response (correct, wrong, missed) for statistical analysis.

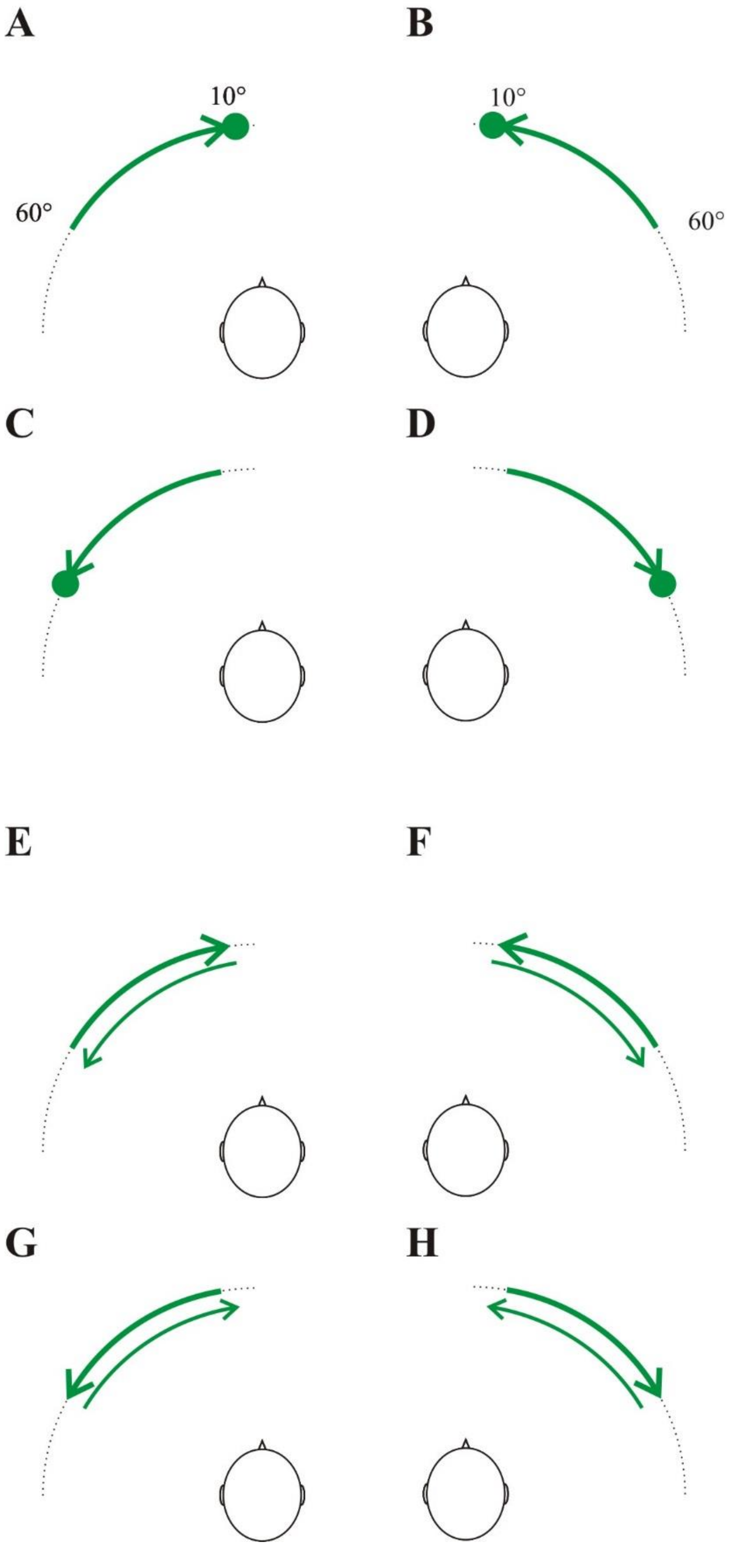


Fig. 1 Graphic representation of stimuli conditions. In each hemifield, there were 2 types of MoSt (A-D) stimuli: one started from a near-central position (C and D) and the other one at a lateral position (A and B) and moved to the other location within the same hemifield in 500ms (green arrows at A-D). Then, the stimuli remained at

the last location for 500ms (green circles). Also, in each hemifield, there were also 2 types of MoMo (E-H) stimuli. In this case, the sound began moving from a lateral position in each hemifield (thick green arrows at E and F) to a near-central within 500ms, and then back to the original position (thin green arrows at E and F) within 500ms. This was procedure was also done with the initial-/final- and mid-locations switched (G and H). All of the stimuli illustrated here were presented in lf (300-1200 Hz) and hf (2000-8000 Hz). Left hemifield: A, C, E, G. Right hemifield: B, D, F, H.

5.3 Results

Behavioural results

Behavioural rm ANOVA analysis for RT focused on the time the participants needed to identify the change in the type of sound within a stimulus. In particular, it focused on conditions that a motion response was expected: the StMo and MoMo. Participants were faster replying to sounds with initial stationary phase (M=.341), p<.05, F(1,13)=13.21, than initial moving phase (M=.451) (Fig. 2). No frequency effect was observed, however, there was an interaction between the initial phase of the stimuli and their frequency, p<.05, F(1,13)=15.61. Participants reply faster to sounds with initial stationary phase when they had high (M=.323) than low frequency spectrum (M=.358) and they replied faster to sounds with initial moving phase when they had low (M=.422) than high frequency spectrum (M=.48).

For the behavioural accuracy analysis, the results showed overall correct (M=199.125) than wrong (M=23.232) or missed (M=17.429) responses for all conditions. In addition, there was an interaction between the initial phase of the stimuli and the accuracy, with overall more correct responses when the stimuli had an initial stationary (M=209.071) than moving (M=189.179) phase.

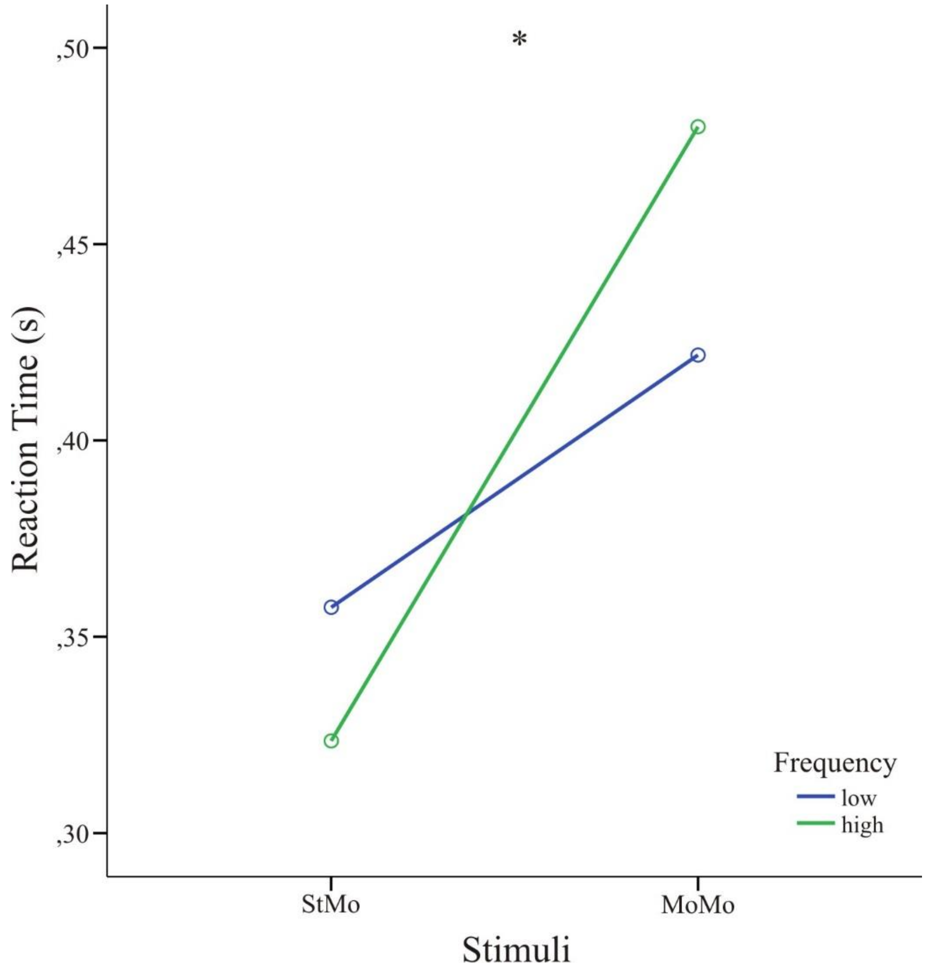


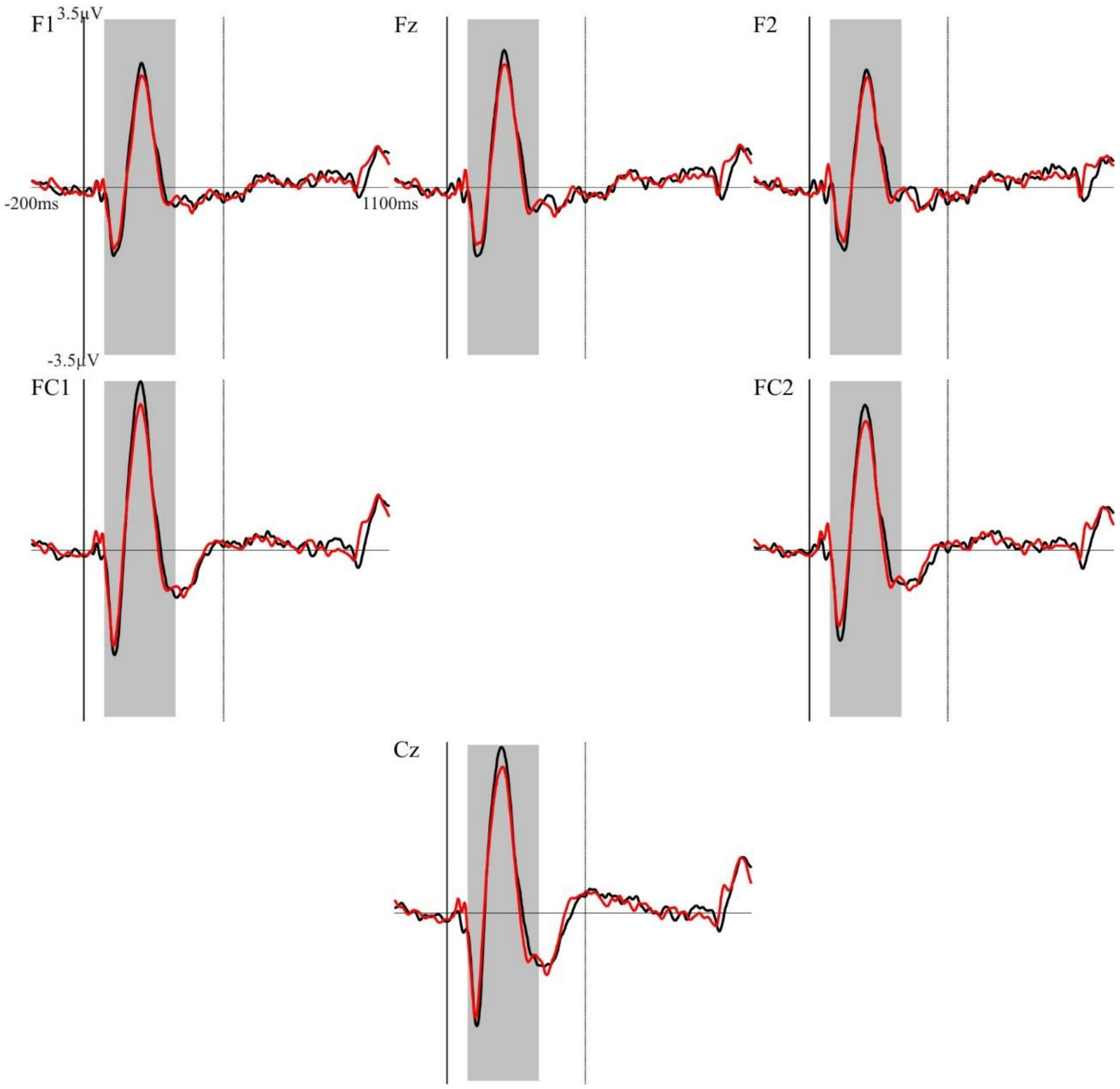
Fig. 2 Effect of the initial type of stimulus on reaction times, p<.05. Blue line shows the reaction times of the conditions with low frequency spectrum. Green line shows the reaction times of the conditions with high frequency spectrum.

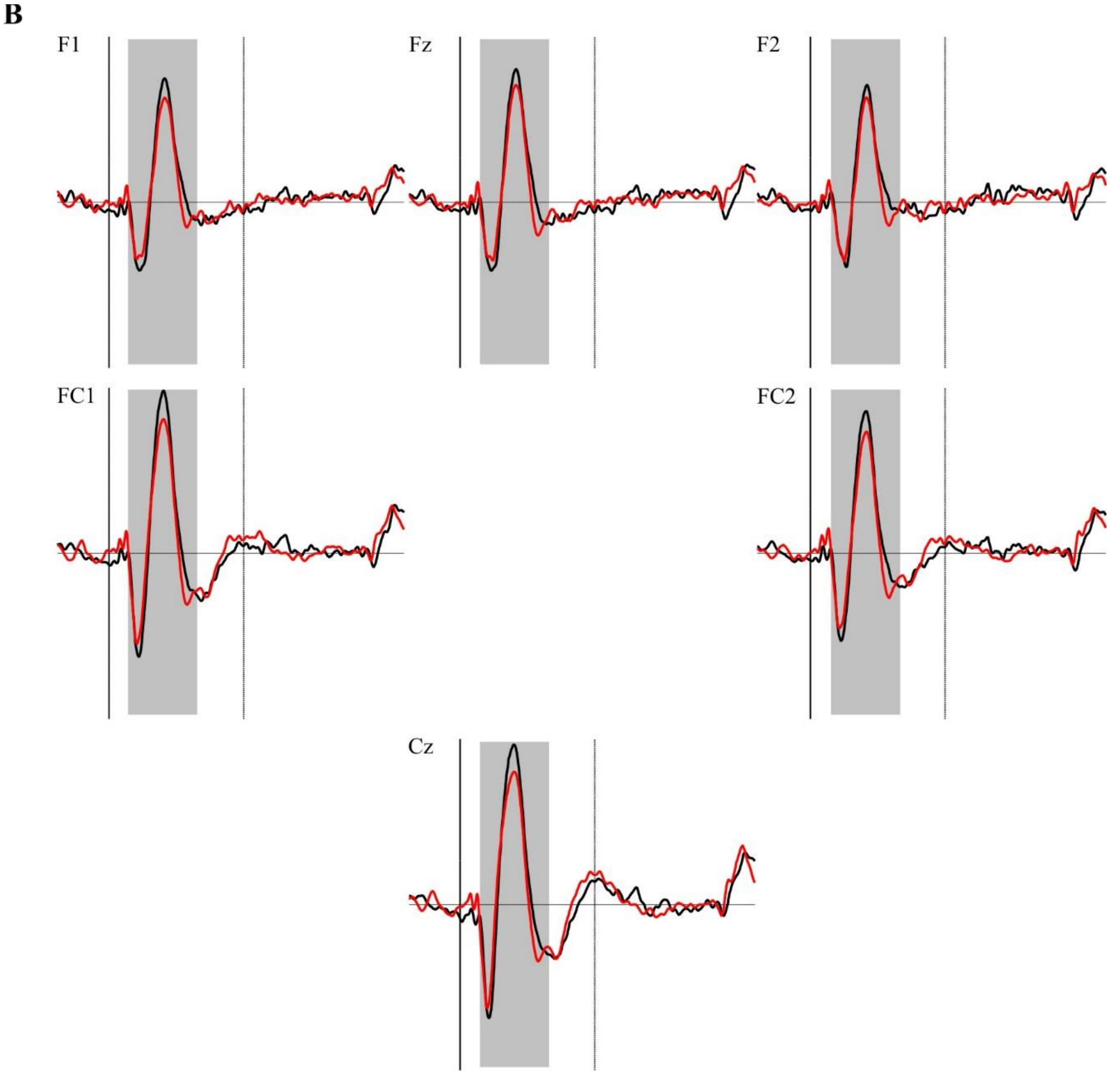
Electrophysiological results

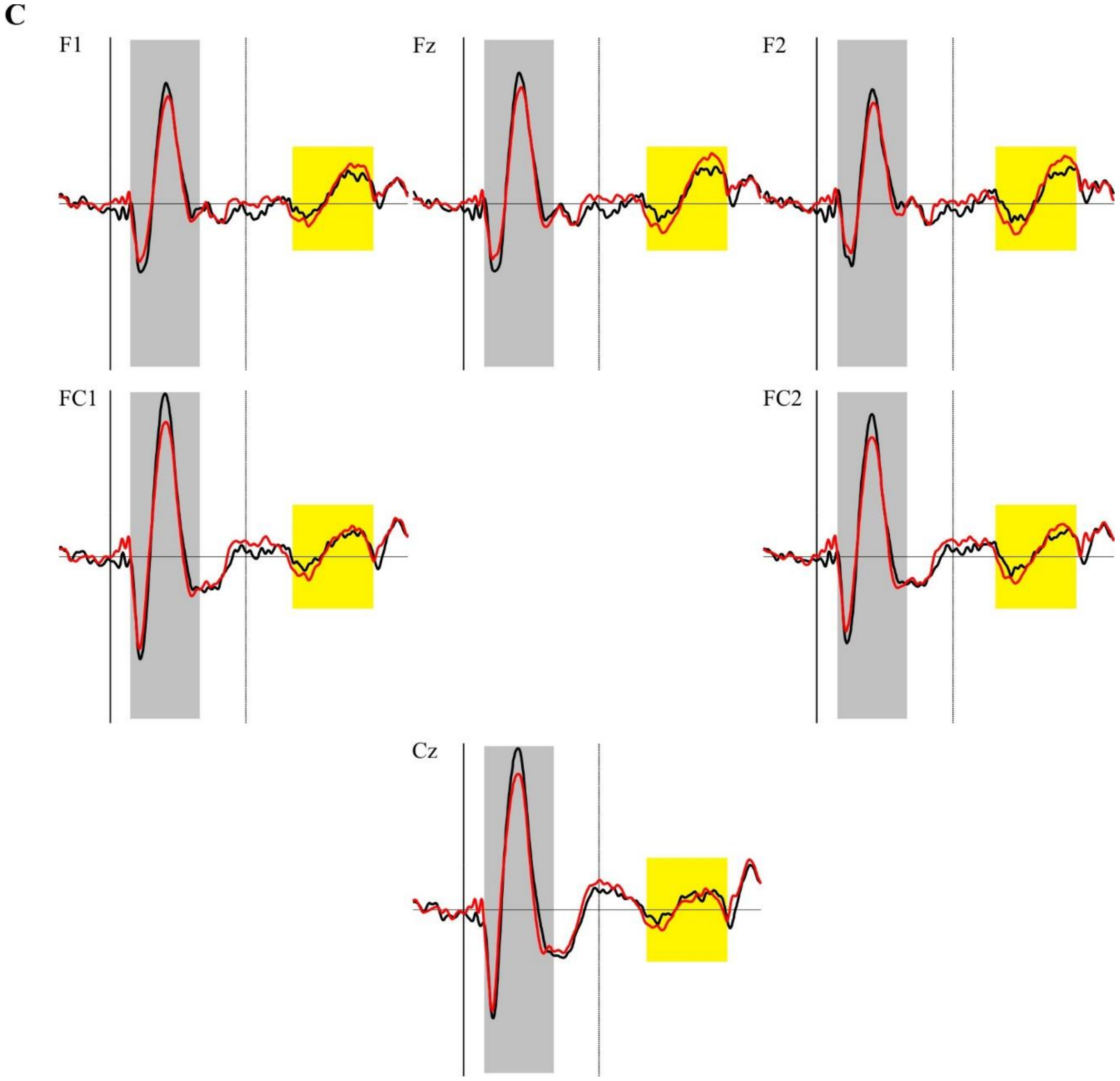
In all conditions, the onset of the stimuli elicited the energy-onset response (EOR) with transient P1, N1, P2 peaks (Fig. 3) between 50-200ms, which is in line with the literature (Chait et al. 2004). These peaks had lower amplitude for the high frequency than the low frequency sounds for all conditions. In addition, the onset of motion for the StMo condition elicited the motion-onset response (MOR), with the peaks cN1 and cP2 (Krumbholz et al. 2007). In this case, the peaks have higher amplitude for the high frequency than the low frequency condition (Fig. 3C). For the MoMo condition, the onset of the second motion, elicited a response that deviated from baseline but it resembled the MOR within the MOR time frame (Fig. 3D). The high frequency condition showed more prominent waveform than the low frequency one. For the MoSt condition, the change of mode within the stimulus from moving to stationary did not elicit any transient response and the signal remained around the baseline (Fig. 3B). For the StSt condition, the stimuli remained at the same location throughout the whole stimulus duration and so no transient responses were observed (Fig. 3A).

Four difference waves were calculated, two from the subtraction of StSt from the StMo condition for high and low frequency conditions (Fig. 4A) and another two from the subtraction of MoSt from the MoMo condition for high and low frequency conditions (Fig. 4B). Then, we extracted the peak amplitudes and the peak latencies as the maximal local negativity and positivity for cN1 and cP2. To compare frequency and history effects on the cN1, we performed rm ANOVA on the peak latencies and peak amplitudes of the difference waves StMo minus StSt and MoMo minus MoSt, with the factors history (S, M) and frequency (L, H). The analysis on the latencies showed a main effect of electrode, $p<.05$, $F(5,65) = 2.436$ (Table 1). The same analysis on the amplitudes of the cN1 showed a main effect of history, $p<.05$, $F(1,13) = 12.306$, with the StMo conditions eliciting stronger amplitudes ($M=-.852$) than the MoMo conditions ($M=-.49$). Also, there was a main effect of frequency, $p<.05$, with the high frequency conditions eliciting higher amplitudes than the low frequency conditions ($M=-.746$, and $M=-.596$, respectively). The same analysis was carried out at the extracted values of the cP2. The analysis of the cP2 latencies did not yield any significance. The analysis on the cP2 amplitudes showed a main effect of history, $p<.05$, $F(1,13)=6.655$, with stronger amplitudes elicited from the StMo ($M=.86$) than the MoMo ($M=.599$) conditions. There was also an interaction between the electrodes and the frequency of the conditions but because there was no main effect of either of the 2 variables, the results of this interaction are not presented here.

A




B
F1
Fz
F2
FC1
FC2
Cz


C
F1
Fz
F2
FC1
FC2
Cz

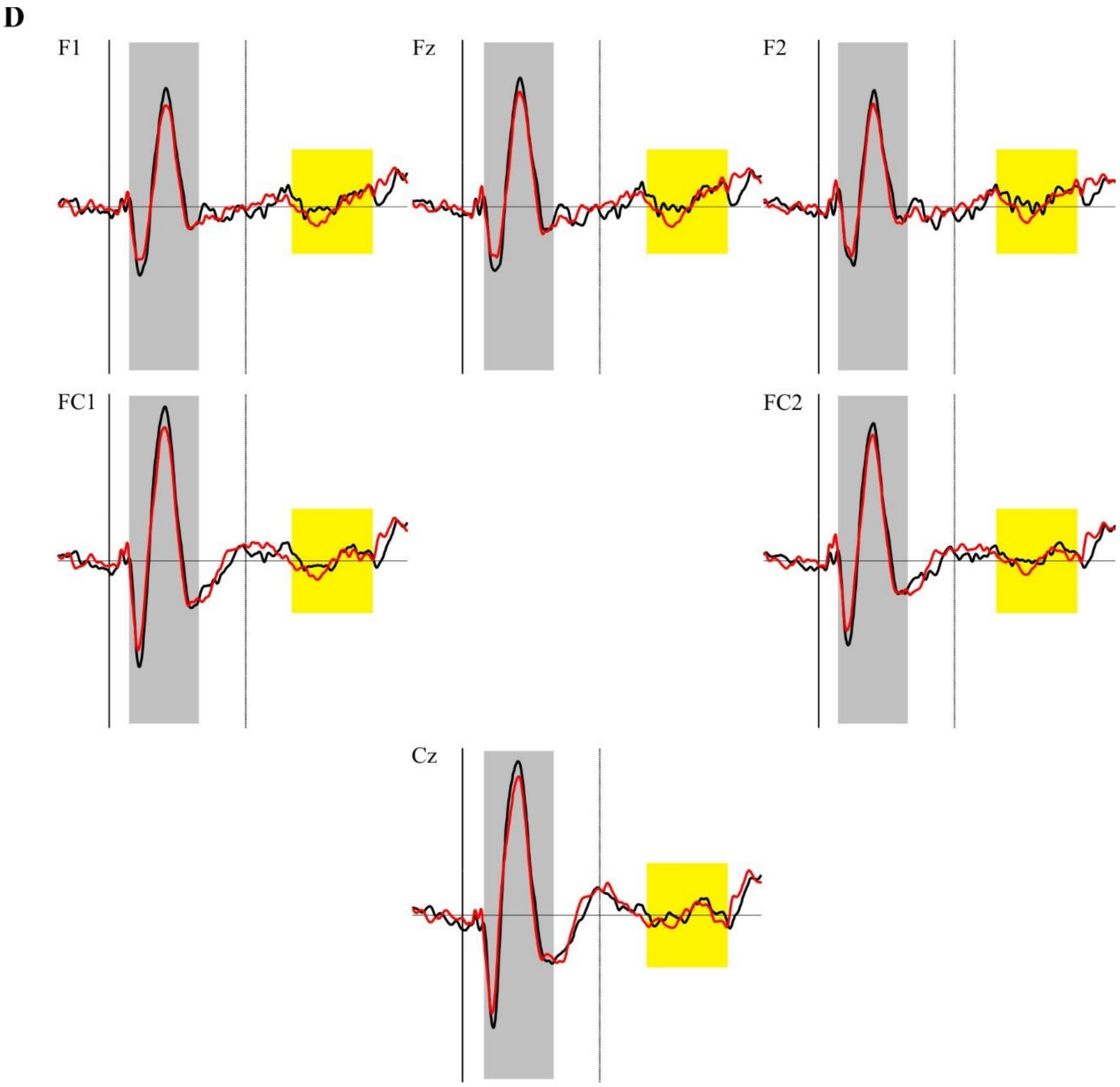
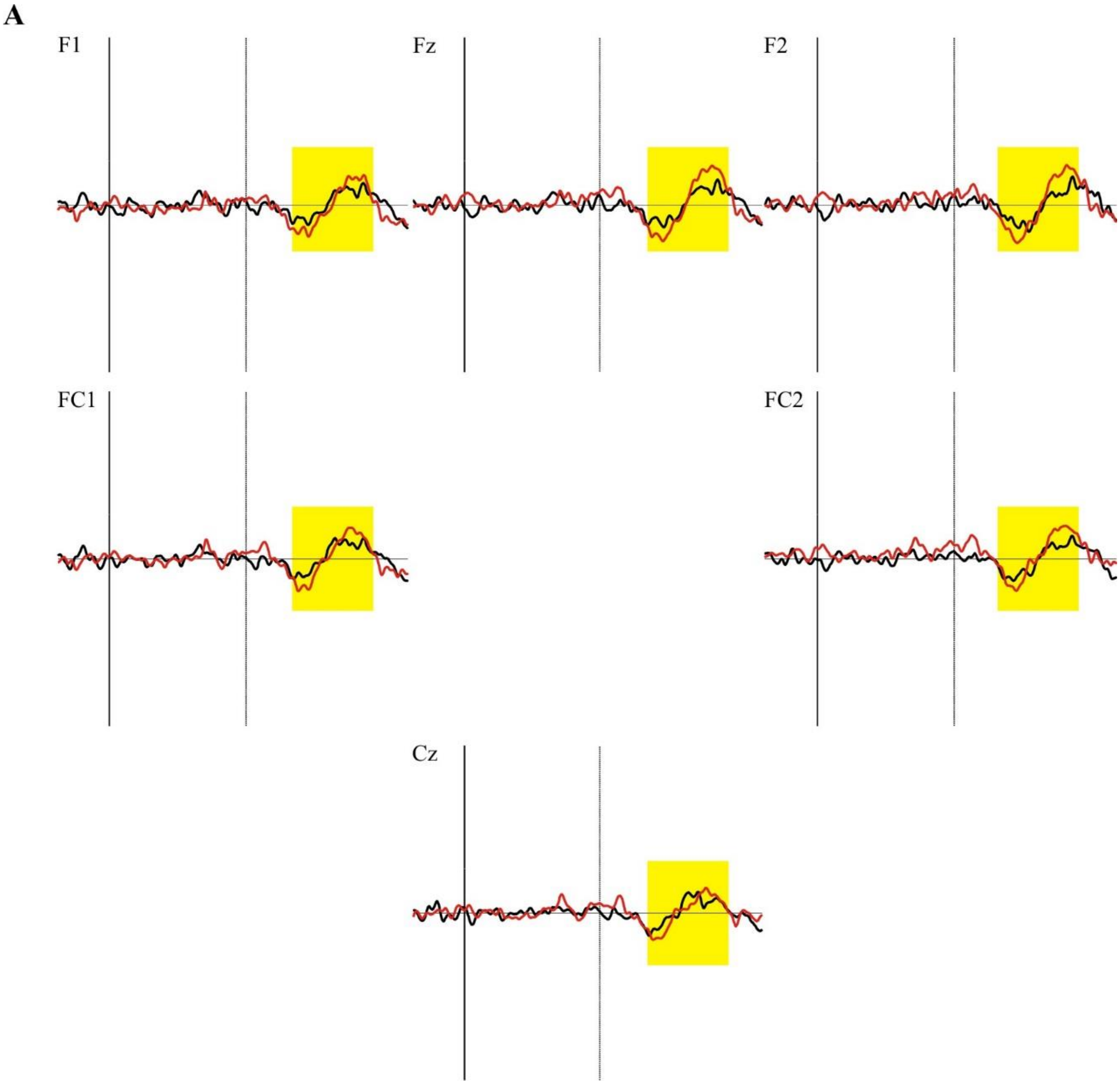


Fig. 3 Grand averages of raw waveforms at the electrodes used for analysis. In A-D, black waveforms represent low frequency conditions and red waveforms represent high frequency conditions. The grey boxes show the EOR with the N1 and P2 peaks. The yellow boxes show the MOR with the cN1 and cP2 peaks. Vertical solid lines show the onset of the stimuli and the vertical dashed lines show the onset of change within the stimuli for MoSt (B), StMo (C) and MoMo (A) at 500ms. **A**: StSt **B**: MoSt **C**: StMo **D:** MoMo

Table 1 Mean latencies and standard error in parenthesis for each electrode resulting from the rm analysis of the cN1.

Electrode	Mean in ms (Standard Error)
Fz	722.964 (3.527)
FC1	717.250 (3.333)
FC2	717.5 (4.19)
Cz	712.214 (4.567)
F1	719.357 (2.569)
F2	721.393 (4.34)



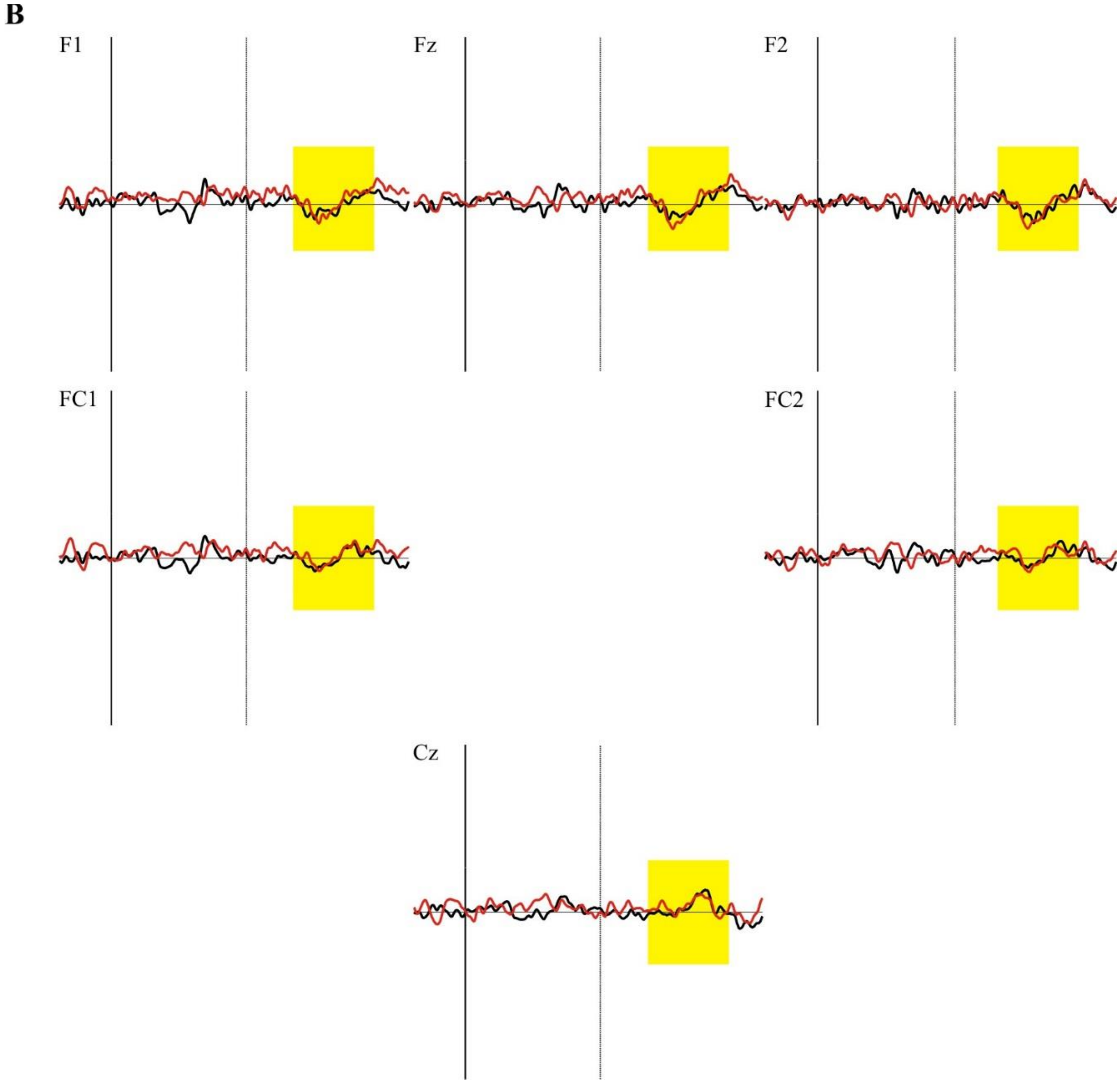


Fig. 4 Difference waves of grand average waveforms at the electrodes used for analysis. In A and B, black waveforms represent low frequency conditions and red waveforms represent high frequency conditions. The yellow box in A shows the MOR with the cN1 and cP2 peaks. The yellow box in B shows the corresponding response at the MoMo condition, in which peaks similar to the cN1 and the cP2 appear. Vertical solid lines show the onset of the stimuli and the vertical dashed lines show the onset of motion at 500ms. **A**: Difference waves resulting from StMo minus StSt for the low frequency (black) and high frequency (red) conditions. **B**: Difference waves resulting from MoMo minus MoSt for the low frequency (black) and high frequency (red) conditions.

5.4 Discussion

The purpose of the current study was to investigate whether the MOR receives contributions from neuronal populations that encode overall spatial change or motion per se by using a delayed motion-onset stimulus paradigm with an initial moving phase (MoMo). It was compared with the MOR elicited from the usual delayed motion-onset paradigm with an initial stationary phase (StMo). Sounds for both MoMo and StMo conditions were presented in low and high frequency spectrum and the isolation of the components was done using difference waves. All conditions elicited the EOR at the frontocentral electrode sites with low frequency having higher amplitudes than high frequency, a finding that is in line with previous literature (Jacobson et al. 1992; Verkindt et al. 1994). Jacobson and colleagues (1992) recorded EEG responses to low-, mid- and high-frequency tones. Their results showed that the N1 had higher amplitudes as a response to low- than to mid- or high-frequency tones. They attributed the latency effects as a combination of the signal delay travelling along the basilar membrane from the base to the apex of the cochlea and the difference in transmission speed within the primary auditory cortex between the areas that respond mainly to low than those responding mainly to high frequencies. In addition, they associated their amplitude effects to the position of the low frequency response areas as they are located closer to the scalp as the high frequency ones and also to the contribution of basal areas of the basilar membrane to low frequency sounds. Moreover, Verkindt and colleagues (1994) recorded EEG from participants that were also listening to tones of several frequencies. They analysed scalp current densities of the P2 response and they found that the P150, which contributes to the P2, had decreasing frontocentral positivity with increasing stimulus frequency. The researchers interpreted these results as evidence of tonotopical organization of the underlying generators different from those generating the N1.

The onset of the second motion in the MoMo conditions elicited a response that resembled the MOR and it was within the same timeframe but had lower amplitude. The peak elicited in the MoMo conditions corresponding to cN1 was especially prominent in the high frequency conditions. Analysis of the difference waves showed higher amplitudes for the cN1 and cP2 at the StMo minus StSt wave than the MoMo minus MoSt wave. This was also accompanied by significantly faster and more correct responses at the StMo than MoSt conditions. These results were expected since the change of mode from stationary to a moving sound is more prominent than presenting motion after a motion phase. It is worth mentioning, however, that although the MoMo condition had constant motion, it introduced a change of direction within the stimulus. Therefore, the lower amplitudes from the MoMo minus MoSt wave could be because the locations travelled at the second part of the MoMo stimuli were adapted by the presentation of motion in the first phase of the stimuli. Therefore, the adaptation expected from the first part might be location- instead of motion-specific. The presentation of an only-moving stimulus having the same direction throughout its duration would not be ideal since it would not introduce any specific change within it to elicit a change-response. If we assume, however, that the activity we observed is the result of the mechanism of motion-specific or spatial change detectors as described by Getzmann and Lewald (2012), then we might attribute the lower amplitudes of the MOR-like response of the MoMo minus MoSt wave in comparison to the amplitudes of the MOR resulting from the StMo minus StSt wave to the activity of the spatial change detectors contributing to the EOR, the early signal elicited by the first difference wave and therefore, they have lower amplitude as a response to the second motion-onset.

In addition, higher amplitudes were elicited at cN1 from high than low frequency conditions, for both difference waves. This result is in line with a previous more detailed MOR analysis on the StMo minus StSt wave (Chapter 4) and here it is extended to the results of the MoMo conditions. In the previous analysis, we proposed the stronger contribution of high-frequency tuned neuronal ensembles to the MOR signal perhaps elicited from motion-specific cortical areas (Poirier et al. 2017) since research supports the motion-specific mechanisms in the auditory cortex (Magezi et al. 2013). The same mechanism of higher contribution from high-frequency tuned neurons could contribute to the current data, even if the MOR is based on spatial change-specific mechanisms. Interestingly, the behavioural results did not accompany the electrophysiological results as in the previous analysis. Reaction times were faster for the StMo stimuli with high frequency and for the MoMo stimuli with low frequency. Since the responses were given after the end of the stimuli and 500ms after the change within the stimuli had occurred, we cannot interpret the behavioural results as pure reflection of variable manipulation and not as artifact.

There was also a main effect of electrode for the latency of the cN1. Although, one particular electrode site did not pop out, a pattern was observed. The cN1 was elicited earlier at the central vertex (Cz) electrode, moved to the two frontocentral sites (FC1, FC2) and then moved to the two frontal sites (F1, F2). It is safe to assume that the underlying neuronal generators of all signals in the electrode sites are the same (Luck 2014). Nevertheless, there is a distinct propagation of the signal, with the generator’s signal moving from the central hemispheric position to the right and left cortical fields.

Chapter 6. *Study 4*: Fast velocity and strong adaptation lead to stronger motion-onset response and split its latter phase into its cumulative parts.

Abstract

In the current study, the effects of the adaptation time and the velocity of motion were investigated for the motion-onset response (MOR). The modulation of the velocity was done by changing the duration of motion. The results showed that faster velocity elicited higher amplitudes on the peak-to-peak comparison. Separate analysis on the MOR components, showed that this effect was based on higher cN1 amplitude. A separate analysis between the electrodes over the left and right hemisphere, showed that the peak-to-peak amplitude was stronger on the electrodes over the right hemisphere. Separate analyses on the MOR components showed that this effect was based on higher amplitude on the cP2 for the right hemisphere. Lastly, the strong adaptation created by the long duration of the initial stationary part provided abundant evidence of auditory motion, which led to the separation of the cP2 into its constituent parts.

6.1 Introduction

Localization of moving sounds is an everyday task that the brain handles with easiness. Recent studies highlight that there are motion-specialized areas and mechanisms in the brain (Magezi et al. 2013; Poirier et al. 2017; Xiang et al. 2002). Functional magnetic resonance imaging (fMRI) data from primates showed that contrasting responses between moving and stationary auditory stimuli revealed specific parts of cortical areas that respond only to motion and the authors considered that this signal arises from neurons that are activated as a response to auditory motion (Poirier et al. 2017). Magnetoencephalographic (MEG) data from human participants to stationary and moving sounds identified responses elicited only from moving sounds and whose source localization showed that they arose from cortical areas associated with auditory motion processing (Xiang et al. 2002). Magezi and colleagues (2013) measured the electroencephalogram (EEG) from participants while being presented with stationary and different kinds of moving sounds. Their results showed that the global field potential from responses to stationary sounds differed significantly from the responses to moving sounds giving evidence that the cortical responses to auditory stimuli show distinction based on whether they are moving or are stationary. The timeframe of the difference in their results coincides with the timeframe of the motion-onset response (MOR). This response is elicited from the onset of motion after an adaptor stationary sound or in other words, from a delayed motion-onset (Krumbholz et al. 2007). It has distinct peaks, one negative (cN1) and one positive (cP2) that resemble the ones elicited initially from the auditory system, the energy-onset response (EOR) (N1, P2).

Since the initial mentioning of the MOR in the literature (Krumbholz et al. 2007), a series of studies investigated several aspects of how this response is effected and what aspects it represents. EEG data showed the different peaks of the MOR represent different stages of the auditory motion perception, with the early cN1 representing the hemifield location of the sound and the latter cP2 the direction of motion (Getzmann 2011). These results were later confirmed with electrophysiological responses to moving stimuli that showed that the N1 and P2 exhibit similar patterns of activity (Shestopalova et al. 2016). In addition, the velocity of the auditory stimuli is a characteristic that affects the morphology of the motion responses (Getzmann 2009; Shestopalova et al. 2016). Getzmann (2009) presented to participants with earphones sounds with a delayed motion-onset paradigm and recorded their cortical response with EEG. The stimuli were spanning the same trajectory but varied in their duration, modifying in this way their velocities. His results showed that the responses to sounds with high velocity led to earlier and higher in amplitude cortical responses and it also led to increased contralaterality. The latter result was not replicated in the data from Shestopalova et al. (2016). In the latter study, the researchers presented moving sounds of both slow or fast velocity and stimuli with abrupt spatial displacement using earphones while conducting EEG measurement. Their study followed an oddball paradigm with the step and fast stimuli serving as the deviants. Their results showed higher P2 amplitudes and larger mismatch negativity (MMN) as a response to the step stimuli but no contralaterality pattern neither for the ERPs nor for the MMN. The authors explained these results based on the more prominent angular shifts resulting from the step stimuli in comparison to those resulting from continuous motion. Moreover, another study compared the MOR elicited from continuous moving sounds, step moving sounds and scatter sounds and showed that the responses to scattered and step stimuli were elicited earlier and with higher amplitudes (Getzmann & Lewald, 2012). The authors of the afforementioned study suggested that there is a system that detects spatial change within the auditory system that is based on the angular velocity of the acoustic event. Therefore, the stronger ERPs are elicited from those stimuli providing the more noticeable spatial change that in the case of the last two aforementioned studies was the step stimulus. Another study used the ‘discrimination contours’ technique to investigate different combinations of distance, duration and speed cues in the discrimination of motion direction (Freeman et al. 2014). In this behavioural study, the researchers employed an oddity task during which they presented participants with three stimuli out of which one of them differed from the other two by modifying its velocity either in duration of motion or distance of motion. The participants were requested to detect the odd stimulus. Their results showed that motion discrimination is duration-sensitive rather than distance-sensitive. Taken together, the results from the studies on velocity show that the part of the auditory system that is motion-specific is sensitive to this cue. However, there was no study so far showing the effect of velocity by modifying distance of motion on the MOR. This particular issue is addressed in the current study.

In addition, as mentioned earlier, the mechanism underlying the elicitation of the MOR is based on a delayed motion-onset paradigm after the initial presentation of a stationary sound. This initial stationary sound results in the activity from neurons corresponding to its characteristics.

The prolonged presentation of this stimulus leads to adaptation of these neurons. Therefore, the subsequent introduction of motion recruits neurons that have not been adapted during the early stimulus phase, such as those responding to motion or spatial-change cues and thus, the MOR is consider a true response of auditory motion (Krumbholz et al. 2007). What yet remains unanswered is how long the initial adaptor should be in order to create a sustained adaptation effect for the MOR to occur. In the literature, the duration of the stationary part varies from 700-1000ms (700ms: Getzmann 2009; 2011; Getzmann & Lewald 2010; 2011; 2012; Kreitewolf et al. 2011, 1000ms: Getzmann & Lewald 2014; Grzeschik et al. 2013; 2016; Krumbholz et al. 2007). Previous research from our lab (Chapter 4 and Chapter 5) used a stationary phase of 500ms and the elicitation of the MOR was successful. However, the MOR peaks were elicited a few milliseconds later than previously reported in the literature. In addition, earlier data from our lab (Chapter 3) showed that a 250ms stationary sound did not elicit any MOR even if the velocity of motion was relatively medium (120 °/sec).

Therefore, the current study investigated the modulation of the MOR as a function of velocity based on distance and the duration of initial stationary part in the freefield. We used four types of stimuli, out of which two had an initial stationary period of 500ms and the other two had an initial stationary part of 700ms. One stimulus from each of the previous stationary durations had a velocity of 100°/sec or a velocity of 160°/sec by keeping the duration of the moving sound object constant but changing its final position to increase its velocity.

6.2 Method

Participants

A total of 9 healthy volunteers participated in the study (7 females) with a mean age of 26.6 years (all right-handed). All participants were informed about the scope of the experiment and had signed a written consent form for their participation. The experimental procedures were in agreement with the Declaration of Helsinki (World Medical Association 2000).

Setup

The experiment was conducted in complete darkness in a semi-anechoic chamber (40 m^2; Industrial Acoustics Company, Niederkrüchten, Germany) in which 47 loudspeakers (Visaton, FRS8 4 Ohm, Haan, Germany) with radius 2.35 m are arranged in an azimuthal, semi-circular plane spanning from 98° left to 98° right in the frontal hemifield. The distance between the loudspeakers was 4.3°. Each loudspeaker was equilibrated individually. For this, the transmission spectrum was measured using the a Bruel & Kjaer measuring amplifier (B&K 2610), a microphone (B&K 2669, pre-amplifier B&K 4190) and a real-time signal processor (RP 2.1, System3, Tucker Davis Technologies, TDT, Alachua, Florida, USA). For each loudspeaker, a calibration file was generated in Matlab 6.1 (The MathWorks Inc., Natick, Massachusetts, USA) and was later used to ensure flat spectra of the acoustic stimuli across the frequency range tested. The loudspeakers were covered with a sound transparent black curtain to prevent participants having visual location information of the loudspeakers' positions. The generation of the stimuli was digital using real-time TDT processors (RX8, System3) and controlled with custom-made scripts in MATLAB 7.5 (R2007b, The MathWorks Inc., Natick, Massachusetts, USA). An infrared camera was installed in this chamber to oversee participants during testing.

Stimuli

The stimuli were 2 Gaussian noise signals at a low-to-moderate frequency range (300-1200 Hz). The intensity was set on 50 dB sensation level (SL; above individual hearing threshold) with a rowing level +/- 3dB. For each participant, the individual hearing threshold for the low-frequency noise band was measured at the beginning of each session using a yes/no (heard/not heard) paradigm for stationary and moving sound stimuli.

For the experiment, acoustic stimuli had a duration of either 1000ms separated in two 500ms-components (without any temporal gap between them), or 1200ms separated in one part of 700ms duration and a second part of 500ms duration (also, without any temporal gap between them) (Fig. 1). The first part was always stationary and the second was either stationary (StSt) or moving (StMo). For the stationary part of the stimuli a single speaker was activated at a particular azimuthal position. The moving part of the stimuli were created by sequential activation of adjacent speakers with linear cross-fading adjusting the intensity of the output signal voltage of neighbouring speakers. Two signal combinations were formed and they were presented in both the left and the right hemifield: (i) stationary-stationary (StSt) and (ii) stationary-moving (StMo). Stationary sound stimuli were presented in one of the following locations (negative numbers show locations left from midline): -10°, +10°. Moving stimuli moved from either from -10° to -60° and from +10° to +60° in 2 blocks or, from -10° to -90° and from +10° to +90° in the other 2 blocks. The velocity of the moving sounds was set at 100°/sec for the distance between 10° and 60° and at 160°/sec for the distance between 10° and 90°. The sound stimuli remained within the same hemifield.

Altogether 8 stimulus conditions were explored: 2 signal combinations (StSt, StMo), 2 durations for the stationary part (500ms, 700ms) and 2 velocities (100°/sec, 160°/sec). The stimuli were averaged between left and right hemifield presentation. Each stimulus combination was presented 160 times in a pseudo-random manner resulting in 1280 stimulus presentations. The interstimulus interval (ISI) varied between 900-1000 ms starting from the time point a participant gave her response (see below). Short breaks were inserted within the experiment to avoid fatigue effects.

Procedure

Participants were asked to sit in a comfortable chair in the middle of a semi-circular loudspeaker array in the semi-anechoic chamber. In order to keep the head in a steady position, they were instructed to keep their eyes on a fixation cross during the audiogram and the experiment. Both of these sessions were contacted in complete darkness.

The heard/not-heard audiogram was combined with a simple staircase paradigm used to determine participants' individual hearing threshold for stationary and moving sounds of low frequency. The stationary sounds were presented at 0° in front of the participants and the moving sounds spanned a trajectory from -30° (left from midline) to +30°, each lasting 1000ms, with an initial intensity of 62 dB SPL. Participants were instructed to press a left button on a response box to indicate that they have detected a sound (*heard* response) and the right button when they did not detect the sound (*not-heard* response). The stimulus' intensity was decreased by 5 dB for each *heard* response and increased by 5 dB for each *not-heard* response. The average of the stationary and moving noise stimuli defined the SL of the noise stimuli.

During the experiment, the participants were instructed to indicate by pressing the right button on a response box when the presented sound stimulus included a change in modality (stationary-to-moving) and the left button if no such change occurred (stationary-stationary). The participants were asked to reply as quick and as correct as possible. During the analysis the reaction times (RT) were quantified, defined as the time from the end of the stimulus to the button press, only for the correct responses. If the answer was not given within a 1000ms time window, then the trial was noted as missed. Before starting the data acquisition, the participants were allowed to familiarize themselves with the task in a demonstration block.

In a previous chapter (Chapter 4), the data corresponding to the StMo and StSt were analysed to study the frequency-specificity of the motion-onset response (MOR) with a stationary sound as history, compared for frequency spectrum, motion direction and hemifield presentation. In another chapter (Chapter 5), the stimuli corresponding to StSt and StMo were averaged based on their frequency range and the stimuli corresponding to MoMo and MoSt were also averaged based on their frequency range to study how the frequency-specificity of the MOR changes based on history (stationary versus moving). In the current chapter, we wanted to study the effects of velocity and the prolonged adaptation on the MOR.

The mechanism of the motion-onset is the release-of –adaptation of the initial stationary phase (Krumbholz et al. 2007). If the initial adaptation is more prolonged, then we expect that the adaptation will be stronger, leading to higher amplitudes on the MOR with the onset of motion. We also expect stronger MOR signal for higher velocities (Getzmann 2009). We do not know however, how the 2 variables interact with each other.

Data Recording and Analysis

The EEG data were recorded using actiCAP Standard-2 (Brain Products GmbH, Gilching, Germany) at a sampling rate of 500 Hz. The recording consisted of 64 Ag/AgCl active electrodes placed on the scalp according to the international 10-20 system (Towle et al. 1993). The horizontal EOG (hEOG) was recorded by placing 2 electrodes next to the eyes and the vertical EOG (vEOG) was recorded by placing 2 electrodes above and below the left eye; both placements were at the level of the retina. The ground electrode was placed in the middle of the forehead just above the nation and the online reference electrode was placed at the tip of the nose. Two additional electrodes were placed on the mastoids. Impedances were kept below 5 kΩ.

The EEG data were pre-processed with Brain Vision Analyzer 2.0 (Brain Products GmbH, Gilching, Germany). The data were bandpass filtered (IIR, cut-off frequencies 0.1 40Hz) and re-referenced to the average of 60 electrodes (58 scalp channels and 2 mastoids; implicit reference included in the re-reference). Then they underwent ocular correction using the Gratton and Coles procedure (Gratton et al. 1983). The epochs were set at -200ms to 1200ms for the stimuli that had 1000ms duration and at -200 to 1400ms for the stimuli that had 1200ms duration, based on stimulus onset (200ms baseline period). Those which exceeded an amplitude of ±100 µV were excluded from further analysis. They were then baseline corrected for the period of 200ms before motion onset and averaged between the trials.

Statistical analysis

The statistical analysis was done using IBM SPSS Statistics for Windows, version 22.0 (IBM Corporation, Armonk, New York, USA). The data were analysed in a 2X2 design: time (brief, long), velocity (slow, fast). Therefore, the epochs from the sounds presented at the left hemifield and those presented at the right were averaged for each condition. The presence of the MOR was calculated by creating 4 difference waves between the StSt and StMo conditions, StMo minus StSt, 2 for the 500ms stationary conditions and 2 for the 700ms stationary conditions, out of which 1 corresponded to the 60° and 1 to the 90° conditions. The cN1 peaks were defined as the maximal local negativity between 150-250 ms after motion onset and the cP2 peaks as the maximal local positivity between 250-450 ms after motion onset on the difference waves for the electrodes F1, Fz, F2, FC1, FC2, C1, Cz and C2. The amplitude and latency of these peaks was extracted and analysed. Both extractions were done automatically.

To assess the dimension of the MOR for each condition, the peak-to-peak calculation of cP2 minus cN1 for each electrode in each condition was extracted and analysed. To investigate the effects of each of the peaks separately, repeated-measure (rm) ANOVAs were carried out at the extracted values (latency and amplitude) for both peaks, at each electrode separately with the factors stationary stimulus time (brief is 500ms and long time is 700ms, BT and LT, respectively) and velocity (slow was up to 60° and fast was up to 90°, SV and FV, respectively), with p values set at .05. To assess the laterality of the effect, the electrodes over the left hemisphere (F1, FC1, C1) were averaged and compared with the ones averaged over the right hemisphere (F2, FC2, C2). If the assumption of sphericity was violated, then Greenhouse-Geisser correction was applied. Multiple comparisons were reported with Bonferroni correction.

The behavioural task was used for ensure the constant attention from the participants to the stimuli, therefore the recorded behavioural measures were not analysed.

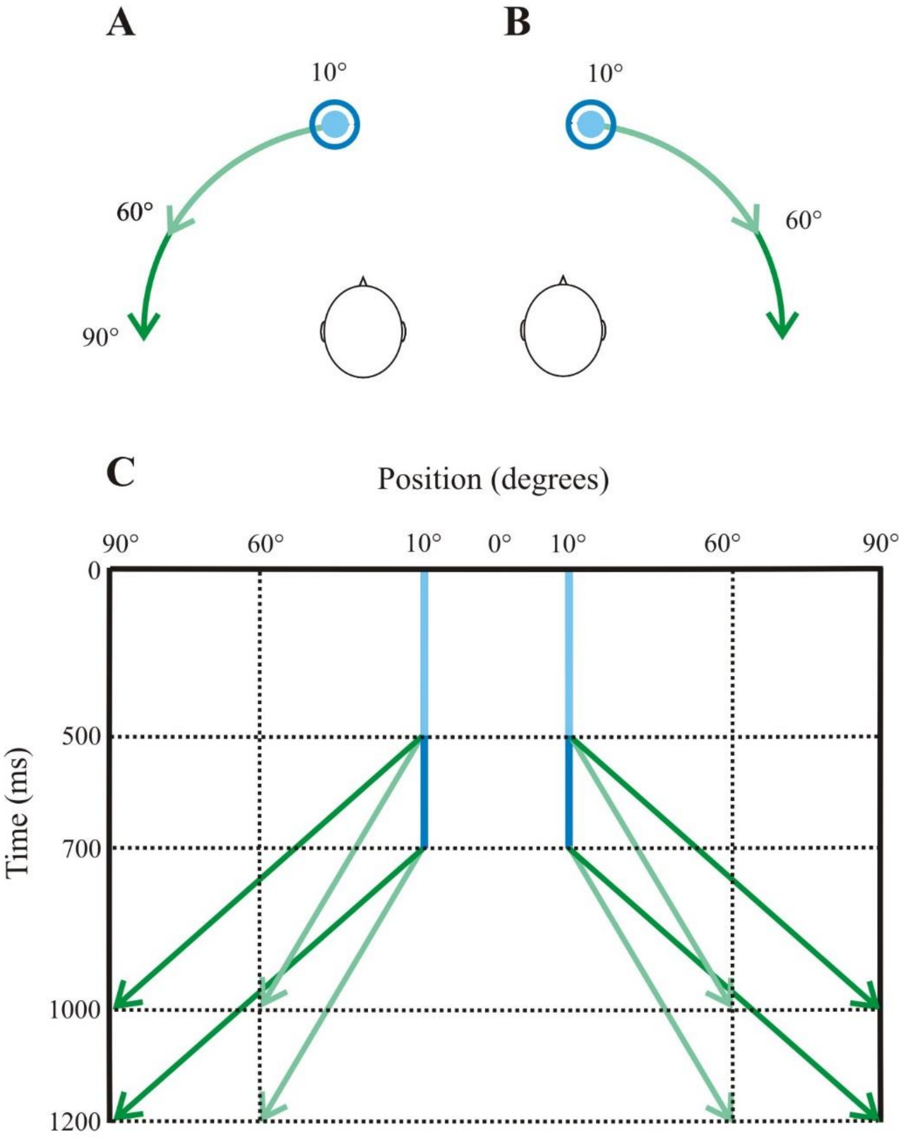


Fig. 1 Graphic and schematic representation of stimuli conditions. In each hemifield, there were 2 types of StMo (A, B) stimuli: one started from 10° with a stationary phase of 500ms (cyan circles at A and B, cyan lines as C) and the other started at the same positions but with a 700ms stationary phase (blue circles at A and B, blue lines at C). After the end of the stationary phase the sound stimuli moved from the 10° to either 60° (mint green arrows at A, B and C) or to 90° (dark green arrows at A, B and C).

6.3 Results

The onset of the sound stimuli elicited the usual peaks of the auditory system, P1, N1 and P2 that are usually referred to as the energy-onset response (EOR) between 50-250ms in all conditions StSt and StMo (Chait et al. 2004) (Fig. 2). In the StSt conditions the signal remained at the same location throughout the whole stimuli duration, either 1000ms for the BT conditions or 1200ms for the LT conditions. Therefore, the signal of the StSt after the initial EOR, followed a patterned of sustained negativity with distinct offset effects after the end of the stimuli, with higher amplitudes for the BT conditions. The StMo conditions, with the onset of motion either at 500ms or at 700ms elicited the usual motion-onset response (MOR). This response was evident with prominent cN1 and cP2 peaks. In the raw waveforms this response in located under the baseline due to the sustained negativity. Higher amplitudes at the MOR were elicited from the LT and FV conditions.

The difference waves resulted from the calculation of StMo minus StSt in each. The EOR is not evident since no differences existed between the beginning of the signal of the StMo and StSt conditions. The MOR in the difference waves is very prominent with apparent cN1 and cP2 peaks.

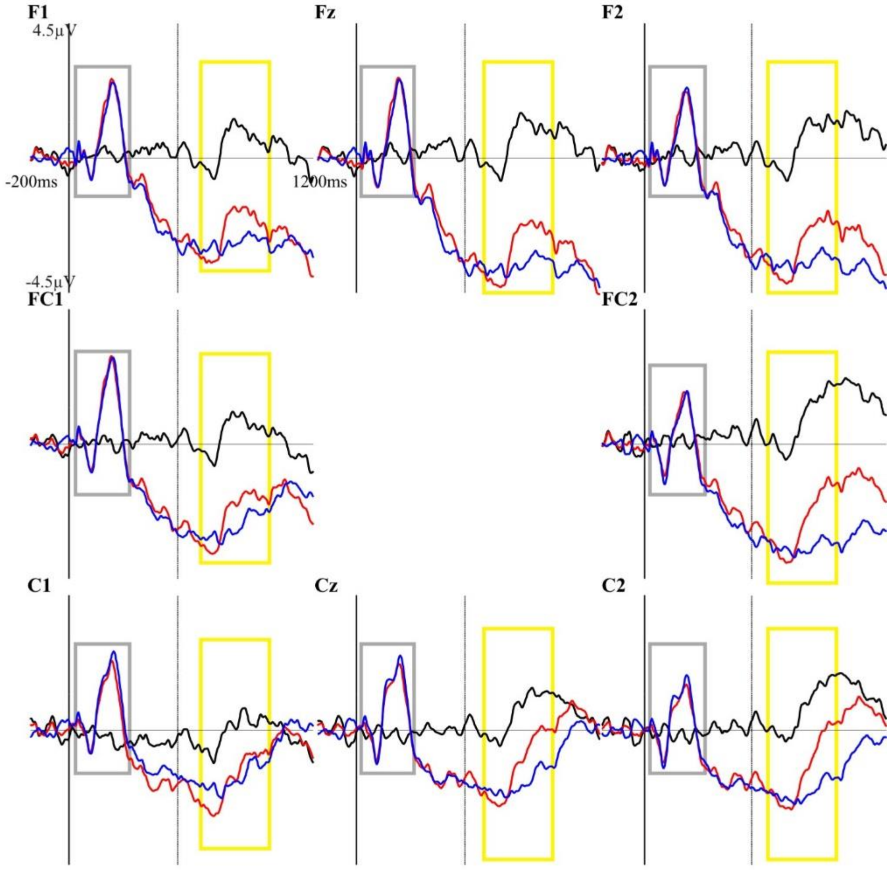
The MOR appeared more pronounced in some conditions. To assess this, rm ANOVA was performed on the peak-to-peak values for each electrode and condition on the difference waves. The results showed a main effect of velocity, $p<.05$, $F(1, 8) = 6.149$, with the biggest peak-amplitude difference for the fast ($M=4.382$) than the slow velocity ($M=3.446$). To assess peak contribution to the previous effect, rm ANOVAs were performed at the extracted latencies and amplitudes of the cN1 and cP2, separately. The results showed a main effect of velocity on the cN1 amplitude, $p<.05$, $F(1, 8) = 6.36$, with higher amplitudes elicited as a response of the fast ($M=-1.453$) than the slow velocity ($M=-.976$). In addition, there was a main effect of electrodes on the latency of cP2, $p<.05$, $F(1.966, 15.728) = 7.397$ (Greenhouse-Geisser reported). Pairwise comparisons showed overall faster times from the cP2 recorded at the electrodes at the left than the right hemisphere (Table 1). The cN1 latency and cP2 amplitude did not show any significance.

To measure the laterality of the MOR, analysis was carried out at the peak-to-peak amplitude differences from the averaged electrodes over the left hemisphere and compared with those over the right hemisphere (Fig. 3). The results showed a main effect of hemisphere, p<.05, F(1, 8) = 6.766, with larger amplitude differences between the peaks from the right (M=3.973) than the left hemisphere (M=3.313). Separate analyses on the peak amplitude and latencies of the cN1 and cP2, showed a main effect of velocity on cN1 amplitude, p<.05, F(1,8) = 7.303, with higher amplitudes elicited after the presentation of the fast (M=-1.394) than the slow stimuli (M=-.921). Also, for the cP2 latency there was a main effect of hemisphere, p<.05, F(1,8) = 20.69, with faster times recorded from the electrodes over the left (M=346.722) than the right (M=380.944) hemisphere. On the other hand, for the cP2 amplitude there was a main effect of hemisphere, p<.05, F(1,8) = 6.96, with higher amplitudes recorded from the electrodes over the right (M=2.76) than over the left (M=2.212) hemisphere. In addition, for the cP2 amplitude there was a weak interaction, p=.039 between the stationary time and laterality, F(1,8) = 6.042. The cN1 latency did not show any significance.

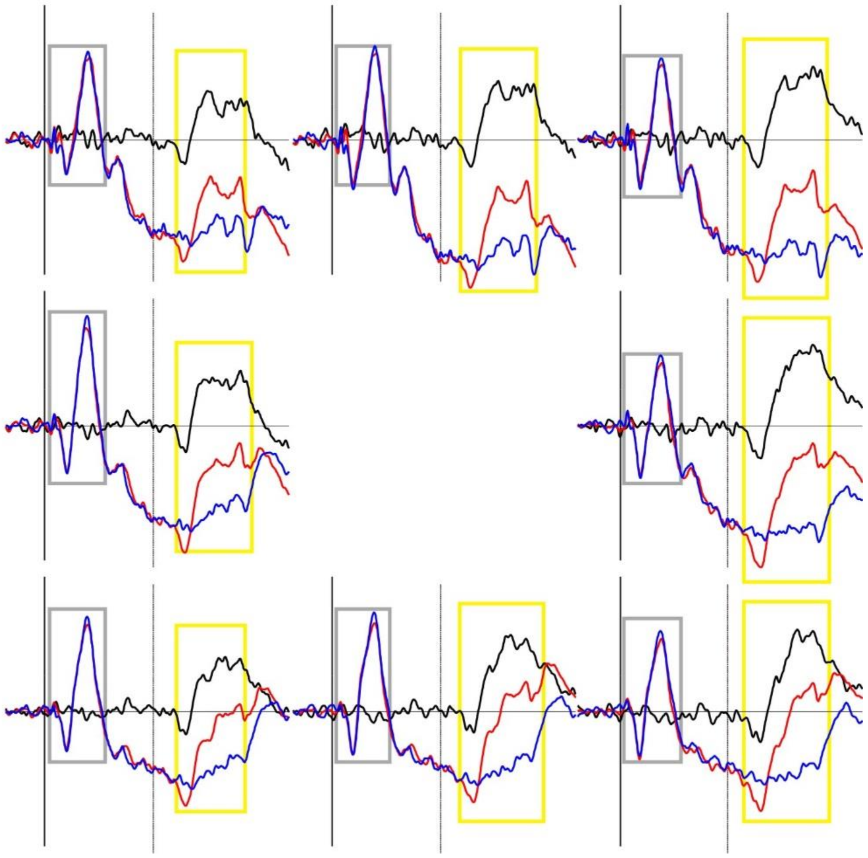
Table 1 Mean times of cP2 for the electrodes used for analysis. Mean shown in parentheses in ms after motion-onset.

Left Hemisphere	**Midline**	**Right Hemisphere**
F1 (338.167)	Fz (353.944)	F2 (360.778)
FC1 (330.944)	Cz (370.222)	FC2 (385.667)
C1 (362.722)		C2 (382.111)

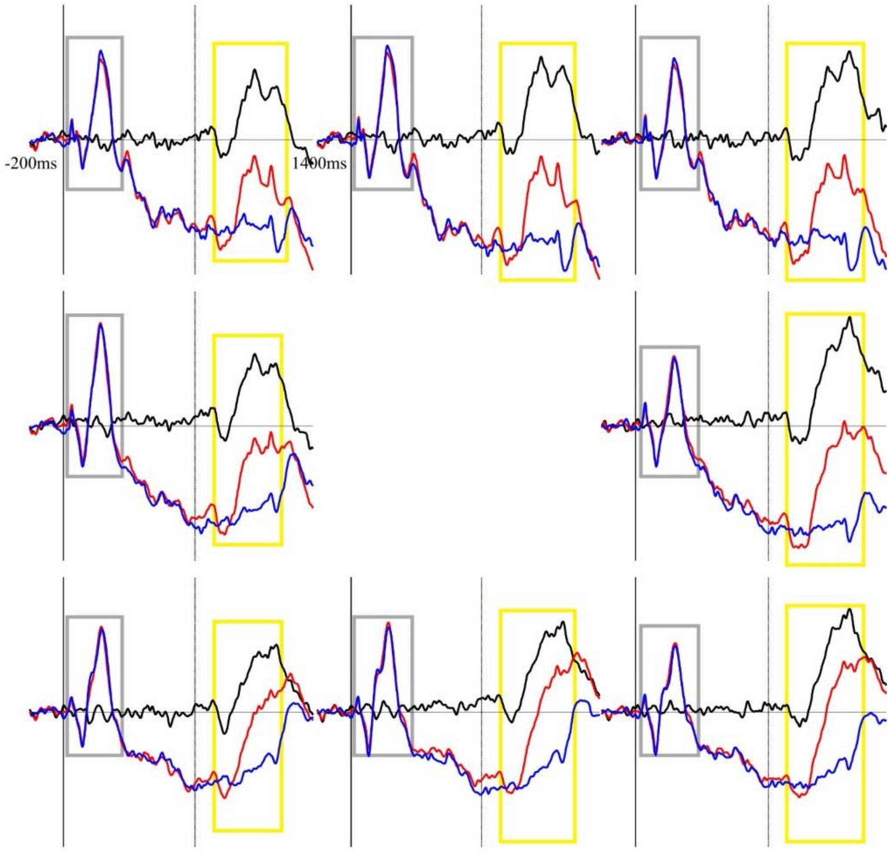
A **Brief time (500ms), slow velocity (100°/s)**



B **Brief time (500ms), fast velocity (160°/s)**



C Long time (700ms), slow velocity (100°/s)



D Long time (700ms), fast velocity (160°/s)

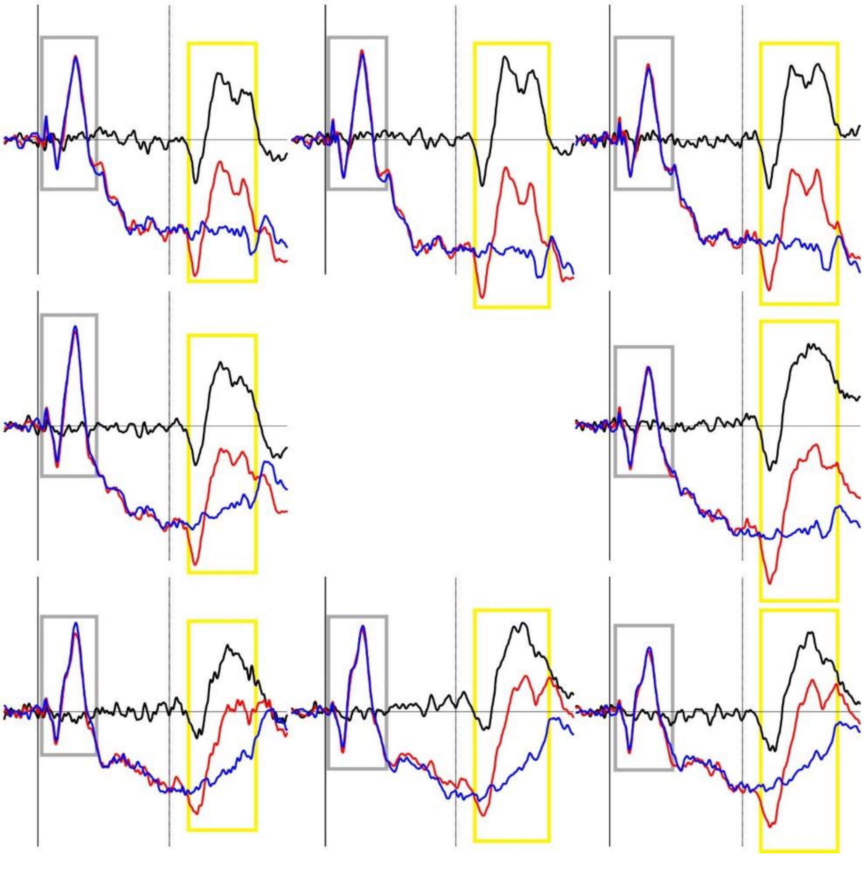


Fig. 2 Raw waveforms and difference waves of the different conditions at the electrodes chosen for analysis. Blue waveforms represent the StSt conditions, red waveforms represent the StMo conditions and the black waveforms are the difference waves resulting from the calculation of StMo minus StSt. The grey boxes highlight the energy-onset response with the peaks N1 and P2. The yellow boxes highlight the motion-onset response time frame in the StMo conditions and the respective response in the difference waves. Vertical black line shows the onset of the stimuli and vertical grey line shows the onset of the motion. All graphs have an amplitude range of -4.5 µV to +4.5 µV. Different time axes apply for the brief time conditions (-200 ms to 1200 ms, A, B) and the long time conditions (-200 ms to 1400 ms, C, D).

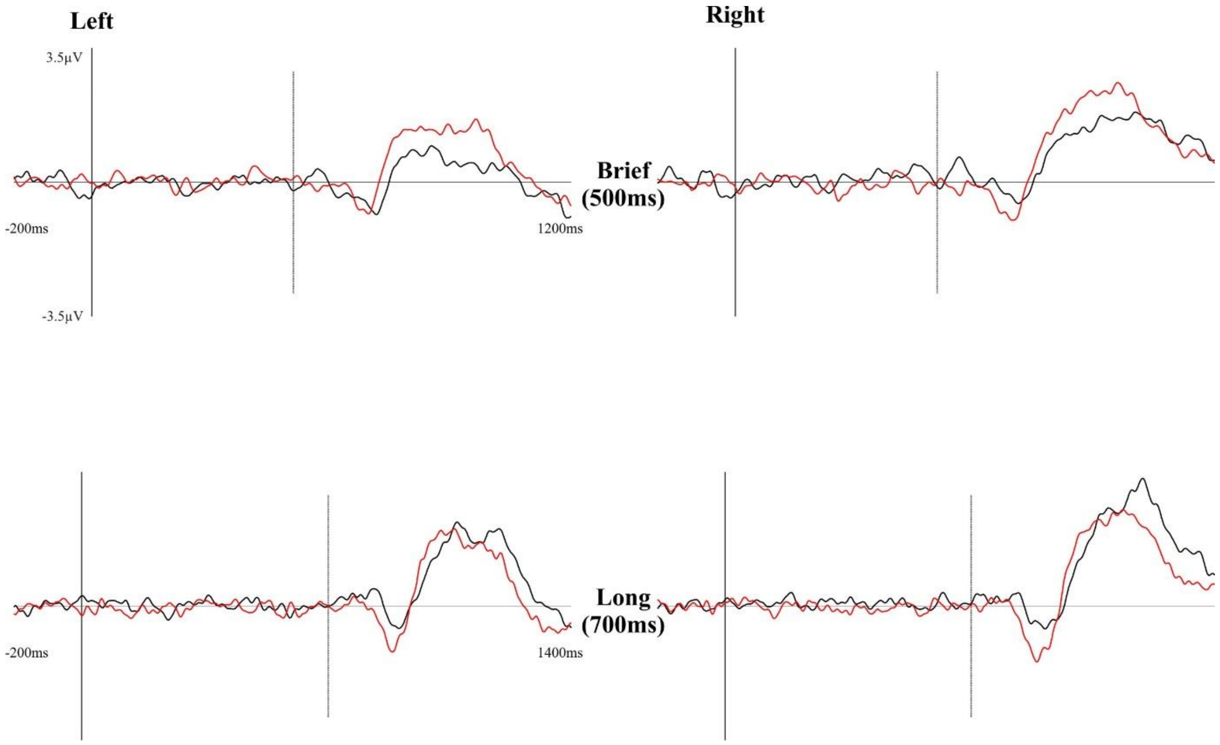


Fig. 3 Difference waves of the different conditions at the electrodes averaged over the left and right hemisphere. Black waveforms represent the slow velocity conditions (100°/s) and red waveforms represent the fast velocity conditions (160°/s). The top raw represents the conditions with the brief initial part (500ms) and the bottom raw represents the conditions with the long initial part (700ms). Vertical black line shows the onset of the stimuli and vertical grey line shows the onset of the motion. All graphs have an amplitude range of -3.5 µV to +3.5 µV. Different time axes apply for the brief time conditions (-200 ms to 1200 ms) and the long time conditions (-200 ms to 1400 ms).

6.4 Discussion

The aim of the current experiment was to investigate the effects of velocity by modifying the distance factor and the effects of prolonged sound adaptation with a stationary sound on the MOR. All stimuli in all conditions elicited the usual triphasic response of the auditory system to the onset of stimulus, P1, N1 and P2 (grey boxes in Fig. 2) known as the EOR. There were no differences among the stimuli used to calculate the difference waves and thus, the signal of the difference waves during the elicitation of the EOR was along the baseline. The onset of motion in all conditions elicited a clear MOR response that was less pronounced at the brief conditions, especially for the slow velocity. The long time conditions elicited in most of the analysed electrode sites two cP2 peaks that were named previously in the literature as cP2a and cP2b (Krumbholz et al. 2007). The analysis of the peak-to-peak of the MOR on the difference waves showed that the peak difference was larger for the fast velocity conditions and analysis on the peaks separately showed that this difference was based on the higher amplitude of the cN1. The cP2 latency differed based on the electrode sites and the subsequent comparison of the left and right electrode sites showed that the cP2 was elicited faster at the left than the right electrode sites. The same comparison of the signal between the left and right electrode sites showed that the peak-to-peak difference was greater for the right than the left electrode sites and separate analyses on each of the peaks showed that this effect was based on higher cP2 amplitudes for the right electrodes than those at the left.

The effect of velocity present in the current data is in line with previous literature (Getzmann 2009). In contrast with the current data, Getzmann (2009) found that both cN1 and cP2 were elicited faster and with higher amplitude with increasing velocity. Although we had higher amplitudes in the current experiment for higher velocities, the latency of the peaks did not follow a similar pattern. A possible explanation for this difference is that in the current study the velocities used are lower than those used in the aforementioned study. In addition, a limitation of the current experiment is the lack of a much faster stimulus velocity. The reason is that the focus was on keeping each stimulus presentation within the same hemifield to avoid crossing the midline and thus keeping the interaural cues underlying the listening of the sounds dominating one ear. Also, if the sounds were to cross the 90° angle, then perhaps would have been a distortion effect on sound perception due to the front-back confusion of sound localization (Makous & Middlebrooks 1990). Finally, maintaining the 500ms duration of sound motion made the result comparison easier with previous data collected from the same lab (Chapter 3, Chapter 4, Chapter 5). Furthermore, previous research on the MOR using step stimuli that employs the sudden displacement of sound location showed higher cN1 and cP2 amplitudes and faster elicitation of the cN1 in comparison to smooth motion, like the one used in the current study (Getzmann & Lewald 2012). The authors placed their results in the frame of the previous data from their lab by quantifying the velocity created by the step stimulus and supporting that the stronger response to the step stimuli was due to the much higher velocity it created. Within this scope, the current study employed two velocities that had 60°/s difference and this difference was enough to elicit a peak-to-peak amplitude difference and stronger cN1 amplitude, but not enough to have an effect on the latency of the peaks.

The effect of the electrode sites showing higher peak-to-peak amplitude difference and higher cP2 amplitude at the right than the left electrode sites is in line with previous evidence showing right-hemispheric dominance in the processing of auditory motion (Griffiths et al. 1998; Kreitewolf et al. 2011; Magezi et al. 2013; Shestopalova et al. 2016; Xiang et al. 2002). Also, it seems that the cP2 was elicited later at the right electrode sites in comparison to the sites at the left. A similar result appeared in the study of Shestopalova and colleagues (2016) that showed higher amplitude and later latency on the N1 elicited at the right electrode sites. In addition, the right-hemispheric lateralization was velocity-independent, which is in line with previous data (Shestopalova et al. 2016). However, Getzmann (2009) reported right-hemispheric dominance for slow velocity conditions and not for the fast ones. It is worth mentioning, however, that the velocities in the current study did not differ greatly from the slow velocity in the previous study. Thus, similar results could be achieved if the current study used a greater range of velocities. Moreover, the current result could be explained by the use of an active task during the experiment. Although, the task was only used to maintain constant attentional level from the participants, it still allocated attention to the presence or absence of a moving part within the stimulus. This is supported by previous data from a passive task (Getzmann 2011) that lacked evidence of right-laterality at the MOR.

Interestingly, there was no effect of the initial stationary phase on the MOR for the peak-to-peak amplitude difference. With visual inspection, however, of the raw waveforms as well as the difference waves, it is clear that the longer duration of the stationary part of the stimuli aids in the elicitation of the two peaks that cumulatively form the cP2: cP2a and cP2b. The observation of these two peaks was initially mentioned by Krumbholz and colleagues (2007) in their description of the MOR and later observed by Getzmann (2009). In the earlier study, the stationary part of the stimulus was 1000ms and it the latter, 700ms. Yet it is lacking from most of MOR literature that used similar durations in the stationary part (Getzmann 2011; Getzmann & Lewald 2012; Grzeschik et al. 2010; 2013; 2016; Kreitewolf et al. 2011). In the current study, the calculation of difference waves, confirms that the second peak of the cP2 is motion-related. Also, it seems that is not related to stimulus offset since its latency is shorter than the end of the stimuli duration. The overlay of the cP2a and the cP2b is such that it creates a wider area for the cP2. This is why, the time frame used to extract cP2 information in the current study is twice in duration as the one used for the cN1. In addition to the previous, the majority of the MOR literature uses similar time frames for the cN1 and cP2 and thus it is possible that the information existing for the cP2 so far, reflects the cP2a. It was mentioned, however, that the cP2 is direction-dependent in contrast with the cN1 that is location-dependent (Getzmann 2011). This was later supported with results from a study using moving sounds and investigating the effects on N1 and P2 (Shestopalova et al. 2016). Despite the differences in the stimuli design for the component elicitation between the previous two studies, the latter study found similarly to the results of the earlier study that the early component N1 encodes the stimulus onset with the basic sound characteristics and the late component P2 encodes the movement of the sound as a secondary characteristic. These results resemble the ones from the previous study. In addition to this, other studies show that the P2 component is the cumulative result of the overlap of several components, which is another similarity to the cP2 (Crowley & Colrain 2004; Garcia-Larrea et al. 1992; Verkindt et al. 1994). In particular, Garcia-Larrea and colleagues (1992) applied an oddball paradigm measuring the cortical response to target and non-target stimuli. They inferred from their results that the robustness of the P250 component throughout their conditions was evidence of the cognitive evaluation of the stimulus, which, if it resulted as the stimulus being a target, it would then lead to the elicitation of the P300. This highlights that the P2 represents secondary cognitive processes, which is also relevant to the conclusions from the motion study of Shestopalova et al (2016). If the relation for the role of the early and late parts of the EOR and MOR holds, then the two peaks of the late part of the MOR might have the same functionality, with cP2a elicited for all motion conditions and cP2b elicited for particular stimuli characteristics. In the current data, this characteristic was associated to the duration of the initial stationary part, although vaguely a second peak seems to be overlaid around the cP2 time frame in the fast velocity after the short stationary part as well. Therefore, the cP2b could be elicited in conditions where more evidence for stimulus motion exist. On the one hand, in the case of the stimuli with the long stationary part, this early phase serves as a sufficient adaptor of the neurons responding to it and thus, the onset of motion creates a more abrupt change in the neurons receptive fields, which could lead to the cP2b. This adaptation effect changes with just 200ms of shorter stationary duration. On the other hand, in the case of the stimuli with the brief stationary part, the faster velocity provides sufficient evidence for the stimulus motion, but without the early strong stationary adaptation, the two peaks of the cP2 are not very clear. Either way, both variables enhance the motion cues needed to be detected from the auditory system by providing sudden angular shifts in location. In order to disambiguate the different stimuli effects on the components contributing to the cP2, future studies could present a larger combination of stimuli with a wide range of velocities and adaptation times.

All in all, the data show that increase in velocity of sound motion by modifying its end position increases the peak-to-peak difference of the MOR, which is in line with previous literature that used modulations of motion duration to increase sound velocity (Getzmann 2009). Interestingly, our data show that this increase is based on cN1 amplitude increase, which contradicts previous research showing that the cN1 is associated with the motion-onset and the cP2 with the secondary motion characteristics (Getzmann 2011). Finally, longer adaptation from the initial part of the delayed motion-onset stimulus provides strong release from adaptation with the motion-onset, leading the cP2 to differentiate into its cumulative parts, showing perhaps stronger evaluation of motion from the auditory system.

Chapter 7. *Study 5*: Integration of binaural and spectral cues for motion detection.

Abstract

The azimuthal localization of moving sounds depends on interaural time and level difference cues (ITD/ILD) that change as a function of spatial location. Due to the position of the ears, these cues have an almost symmetrical pattern between the frontal and rear acoustic fields, which creates ambiguity in the localization of sounds. For stationary sounds, this ambiguity is resolved by spectral localisation cues. But how the binaural and spectral cues interact for the perception of moving sounds is not known. We used an adaptation paradigm to examine the integration of binaural and spectral cues. A moving probe was presented in the left hemifield, following an adaptor that spanned either the same trajectory or a trajectory located in the opposite field (frontal/ rear). Participants had to indicate the direction of the probe. The results showed that performance was worse when adaptor and probe were sharing the same binaural cues, even if they were in different hemifields and their direction was opposite. But the magnitude of the adaptation effect when the pair was in different hemifields was smaller, thus showing motion-direction detection depends on the integration of interaural and spectral cues.

7.1 Introduction

Localization of sound sources is valuable for our navigation and interaction with the environment. On the horizontal plane, sound localization is achieved by extracting the binaural cues, the interaural time and level differences (ITDs and ILDs: Schnupp et al. 2011). The binaural cues vary as a function of spatial location with stronger binaural cues corresponding to far left/right locations (Middlebrooks and Green 1991). The spectral cues, generated by direction-dependent filtering of the pinnae and the upper body, disambiguate locations that have the same binaural cues and consequently allow sound localization on the vertical plane as well as front/back disambiguation (King et al. 2001). Localization of moving sound sources on the horizontal plane is based on the dynamic aspects of the binaural cues, and perception of motion direction is based on the balance of these cues, as the spatial location of sound changes with motion (Middlebrooks and Green 1991). However, whether spectral cues contribute to binaural cues for the perception of motion direction remains unknown.

The adaptation paradigm, in which pairs of adaptor and probe stimuli are presented to participants (Butler 1972) has been widely used in the literature. The probe is presented after a repeated adaptor presentation and the response to the probe is measured. The more similar characteristics the adaptor and probe share, the more diminished the response to the probe would be. In the auditory domain, the adaptation paradigm was applied to study the auditory motion aftereffect (aMAE). For example, Grantham (1989) asked participants to indicate the motion direction of a probe, which was either moving or stationary and had the same or different spectrum as the adaptor. The results showed that the aMAE was stronger when the adaptor and the probe had the same spectrum. This effect was later replicated and extended with the addition of spatial disparity between the adaptor and the probe (Dong et al. 2000). In separate experiments, the results showed that the closer the adaptor and the probe are in frequency spectrum and spatial location, the stronger the aMAE is. Dong and colleagues (2000) suggested that the auditory system contains motion-specific channels which are fed by earlier analyses of sound spectrum.

To study whether the aMAE is direction-sensitive, Magezi and colleagues (2013) applied an adaptation paradigm, while also measuring the cortical response of the participants with electroencephalogram (EEG). They presented adaptors that were unidirectional (leftward or rightward), bidirectional (the direction alternating between leftward and rightward motion with each repetition) or stationary. Following the adaptors, the probes were either moving or stationary. The participants had to indicate the motion direction of the probe. The behavioural data showed that the aMAE is direction-sensitive because participants perceived the stationary probes following a unidirectional adaptor, as moving in the opposite direction. This result was not reflected in the EEG responses, supporting that the cortical processing of auditory motion is not motion-sensitive. The direction-selectivity of the aMAE in the behavioural results was based on ITD manipulation to create the percept of motion, which was delivered through headphones. However, whether the direction-selectivity of aMAE still holds when the sound presentation has all cues ITD, ILD and spectral is an issue open for investigation.

The advantage of free-field is not only the naturalistic sounds in structure but also in location making it possible to present sounds at the rear field as well, an aspect of sound presentation that has been a matter of debate in the literature because of the confusing perceptual effects it creates. Ehrenstein (1978) used an adaptor that spanned the whole azimuthal trajectory around the participants and his results to the probes were that no aMAE was observed. He attributed these effects to the change of binaural cues when the sound crosses the rear part of the midline. However, within the left or right hemifield, as mentioned before, certain locations share the same binaural cues and the variance of activity of these cues, which is accounted for auditory motion perception, is what encodes motion-direction. What distinguishes the difference in localization among these locations is how the binaural cues these locations relate to interact with the spectral cues.

Therefore, to investigate this, we decided to present adaptor-probe pairs that share the same binaural cues and have either the same or opposite direction. We used locations within the frontal and rear fields so that when there is a probe in one and an adaptor in the other, they would share the same balance of binaural cues but have opposite direction. In this way, we wanted to observe: (i) Can there be an adaptation effect when adaptor and probe are in different hemifields?; (ii) Will we observe a direction-adaptation effect or a binaural cue adaptation?; and (iii) Will the adaptation get stronger with stronger binaural cues?

7.2 Methods

Participants

Twenty-six participants agreed to participate in the experiment after being informed about the scope of the experiment and had signed a written consent form for their participation. Three participants were excluded because they did not reach a 75% performance criterion in any condition of the psychometric function test, even after doing the test twice. In the first experiment, a total of 15 healthy volunteers participated in the study (6 females) with a mean age of 27.25 years in exchange of small monetary compensation (13 right-handed, 1 left-handed, 1 ambidextrous). In the second experiment, a total of 8 participants with a mean age of 27.9 years in exchange of small monetary compensation (8 right-handed). Participants had hearing thresholds of 15 dB HL or lower, for octave frequencies between 0.125 and 8 kHz, defined at the initial meeting with a stepwise Bekesy audiogram. The experimental procedures agreed with the Declaration of Helsinki (World Medical Association 2000) and was approved by the Research Ethics Committee of the Faculty for Arts and Sciences (CERAS) of the University of Montreal.

Apparatus

The experiment took part in a semi-anechoic room (2.5 X 5.5 X 2.5 m). A comfortable chair was located in the centre of an 80 loudspeaker spherical array (Orb Audio, New York, NY, USA) with 1.8m diameter. The stimuli were white noise sounds generated and presented through the speakers using TDT System 3 hardware (Tucker Davis Technologies, Alachua, FL, USA). Virtual motion was achieved by customized Matlab scripts (The MathWorks Inc., Natick, Massachusetts, USA). The responses were given with the right hand on an OPTIMUS Maximus customized keyboard. The layout of the buttons representing ‘4’ and ‘6’ at the right part of the keyboard was changed with circular arrows, each representing the clockwise and anticlockwise fashion of the replies.

Stimuli

During the initial familiarization phase, the acoustic motion was presented with full modulation depth, for the first experiment at the trajectories 0° - 45° and 45° - 90°, and for the second experiment at the trajectories 90° - 135° and 135° - 180° to the left of the participants, with both directions, making the detection of motion direction very easy.

Then during the psychometric function test, moving sounds were presented at the same 2 testing trajectories as before, with both directions and presented in 7 amplitude modulation depths. The modulation depths used were: 0 dB, 2.5 dB, 5 dB, 7.5 dB, 10 dB, 12.5 dB and 15 dB, which 15 dB was the full modulation depth and 0 dB the least.

At the motion adaptation test, the stimuli were pairs of adaptor and probe stimuli (Fig. 1). The pairs were presented repetitively 10 times in a row, followed by the presentation of one probe. The average of the 75% of the overall performance measured at the psychometric function test was used for the probe presentation; the adaptors, however, were kept at full modulation depth to confirm the maximum of their direction effect. In the first experiment, the adaptors were presented at the following trajectories 0° - 45°, 45° - 90° and 135° - 180°, followed by probe presentation at 0° - 45°, 45° - 90° and 0° - 45°, respectively. In the second experiment, the adaptors were presented at the following trajectories 0° - 45°, 90° - 135° and 135° - 180°, followed by probe presentation at 135° - 180°, 90° - 135° and 135° - 180°, respectively. The pairs could either have the same or different direction. The duration of the adaptor and probe was 416.7 ms and their motion velocity was 180°/sec. The virtual sound object had a width of 15° and traveled a distance of 75° (+15° before the stimulus beginning and 15° after the stimulus end) but only the speakers corresponding to the 45° trajectories were used. The overall amplitude of the stimuli was gated by a cosine gate (as implemented in the TDT software) with its highest intensity at the middle of the trajectory. The response time was set to 1000 ms. The interstimulus interval (ISI) was set at 100 ms. The overall intensity of the sounds was set on 60 dB sensation level (SL).

Overall procedure

The experiment was completed in three parts within 1 hour. Participants completed initially a familiarization task, then a psychometric function test and then a motion adaptation test.

Familiarization task

During this part, participants familiarized themselves with the task and the equipment. The moving sounds were presented at 2 trajectories (see *Stimuli*) and participants were requested to reply as quickly as possible whether the direction of the sound stimulus was clockwise or anti-clockwise. This set of responses was chosen to avoid confusion in the subsequent experimental parts when the sounds were also presented in the rear hemifield. Each stimulus was presented 20 times. This part lasted ~3 min.

Psychometric function test

During this part, the individual threshold of the amplitude modulation depth was defined. The moving sounds were presented in the same 2 trajectories as before in both directions, in 7 modulation depths. There were 20 repetitions of each stimulus that resulted in 560 trials. We introduced 2 breaks (every ~ 6.5 min) in order to prevent tiredness effects, leading to about 24 min duration.

Motion adaptation test

The purpose of this test was to test effect of adaptation on the detection of motion direction. Here we presented 3 pairs of adaptors and probes that had either the same of different direction (Fig. 2). The pairs could either have the same or different direction. Each adaptor was repeated 10 times and then it was followed by a probe (see *Stimuli*). After the last adaptor, there was a break of 500 ms during which, an LED turned on at 0° to indicate that response will be need after the probe. The response time was set to 1000 ms. If an answer was not given within this time frame, the response would be set as missed. The participants indicated the probe's direction (clockwise/anticlockwise) as fast as possible. To facilitate the detection of the probe from the participants, its onset was paired with the simultaneous onset of an LED in front of the participants. There were 20 repetitions of each adaptor-probe pair resulting in 240 trials. This part lasted ~31min with 3 breaks in between.

Statistical analysis

The recorded behavioural measures were reaction times (RT) and hit rates (HR). Reaction times (RT) were defined as the time from stimulus onset to the press of the button. RT of incorrect and missed trials were excluded from the analysis. For the second part of the experiment, Wilcoxon statistical tests were used to measure the difference in accuracy and reaction time of responses between each location for each modulation depth and to extract the psychometric function of these responses. For the third part of the experiment, Wilcoxon statistical tests were used to measure significance between same-/opposite-direction adaptor-probe pairs.

In the second experiment, full amplitude modulation depth (15 dB) was used for the conditions that did not reach the preset threshold.

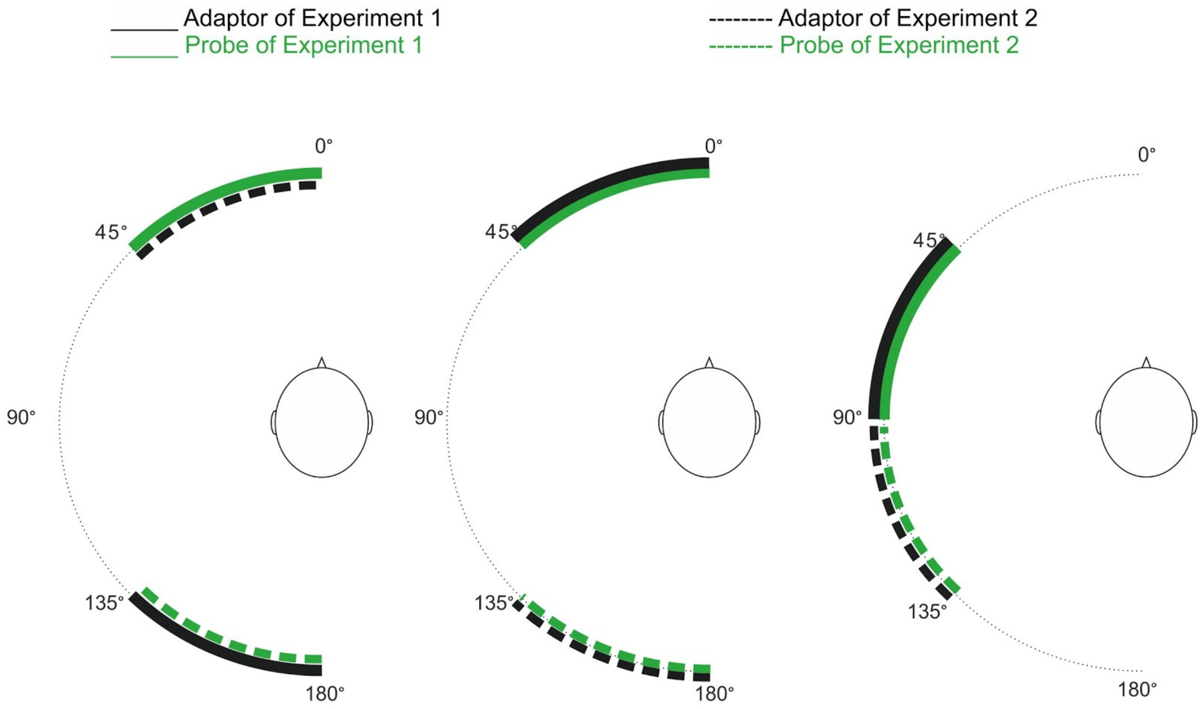


Fig. 1 Graphical representation of the 3 adaptor-probe pair stimuli used for the third part of the first and second experiments. Solid lines represent the stimuli from the first experiment and the dashed lines represent the stimuli from the second experiment. The green trajectories represent the locations of the probe. The black trajectories represent the locations of the adaptors.

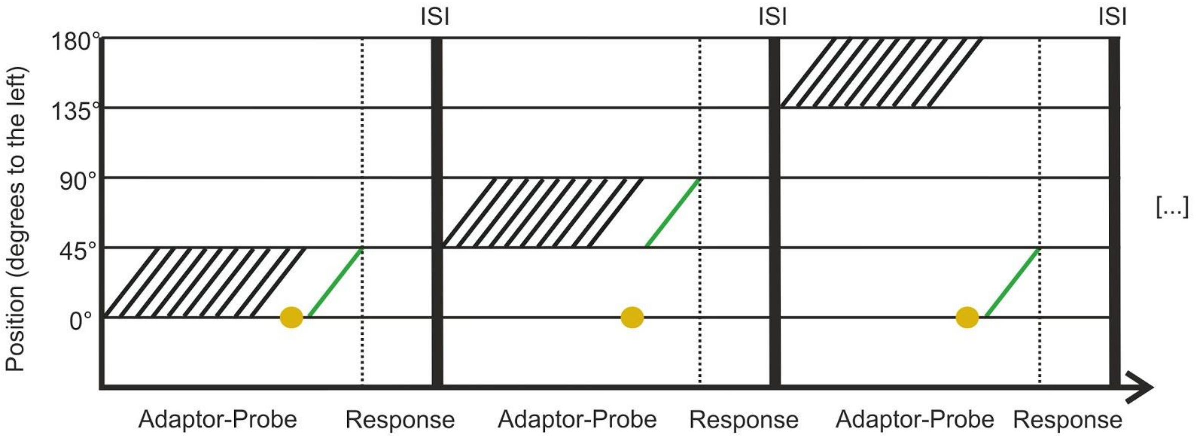


Fig. 2 Graphical representation of the experimental procedure during the third part of the experiment. Here the conditions to the first experiment are depicted. The adaptor (black lines) was presented repeatedly for 10 times in 1 of 3 trajectories (Experiment 1: 0° - 45°, 45° - 90°, 135° - 180°; Experiment 2: 0° - 45°, 90° - 135°, 135° - 180°) followed by a probe (green lines) (Experiment 1: 0° - 45°, 45° - 90°, 0° - 45°, respectively for each adaptor; Experiment 2: 135° - 180°, 90° - 135°, 135° - 180°, respectively for each adaptor). The adaptor-probe pair could have the same or different direction. After the last adaptor, there was a break of 500 ms during which, an LED (yellow circles) turned on at 0° to indicate that response will be need after the probe. Each set of adaptor-probe presentation was followed by the response time (1000 ms) and the ISI (100 ms).

7.3 Results

The results of the participants' performance for the psychometric function are presented in Fig. 3. For the 0° - 45° trajectory the average 75% of amplitude modulation depth was 6.34 dB. For the 45° - 90° trajectory the average 75% of amplitude modulation depth was 7.85 dB. For the 90° - 135° trajectory the average 75% of amplitude modulation depth was 15 dB. For the 135° - 180° trajectory the average 75% of amplitude modulation depth was 5.98 dB. Significance testing with the Wilcoxon test showed that participants' performance differed significantly between locations, $p<.005$, for both the first and the second experiments. RT did not differed significantly between the 2 locations, $p = .095$ and $p = .55$ for the first and second experiments, respectively.

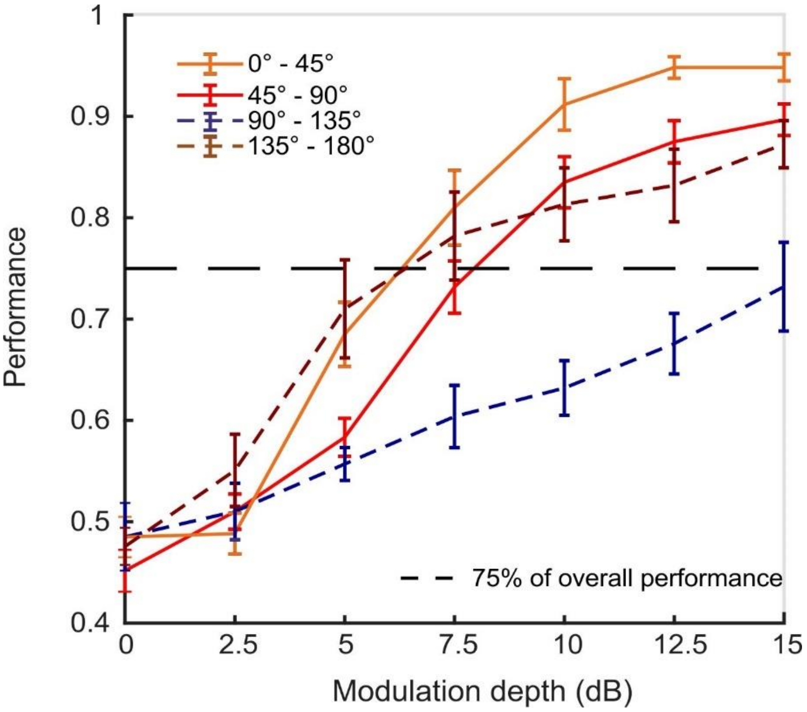


Fig. 3 Line chart presenting the psychometric function of the average of all participants' performance. Orange line shows the mean of the responses to 0° - 45° trajectory. Red line shows the mean of the responses to 45° - 90° trajectory. Blue line shows the mean of the responses to 90° - 135° trajectory. Brown line shows the mean of the responses to 135° - 180° trajectory. Solid lines are the psychometric functions resulting from the first experiment and dashed lines are the psychometric functions resulting from the second experiment. Vertical bars represent standard deviation. The horizontal black dashed line shows the 75% of the performance.

The results of the participants' accuracy performance and RT are shown in Fig. 4. For the first experiment, significance testing with the Wilcoxon test showed that participants' accuracy in all adaptor-probe pairs differed significantly between same-/opposite-direction, p<.001. For the second experiment, significance testing with the Wilcoxon test showed that participants' accuracy differed significantly between same-/opposite-direction at the position 135° - 180°, p<.05. Overall participants' RT differed significantly for the 0° - 45° trajectory, p<.05, and for the 90° - 135° trajectory, p<.05.

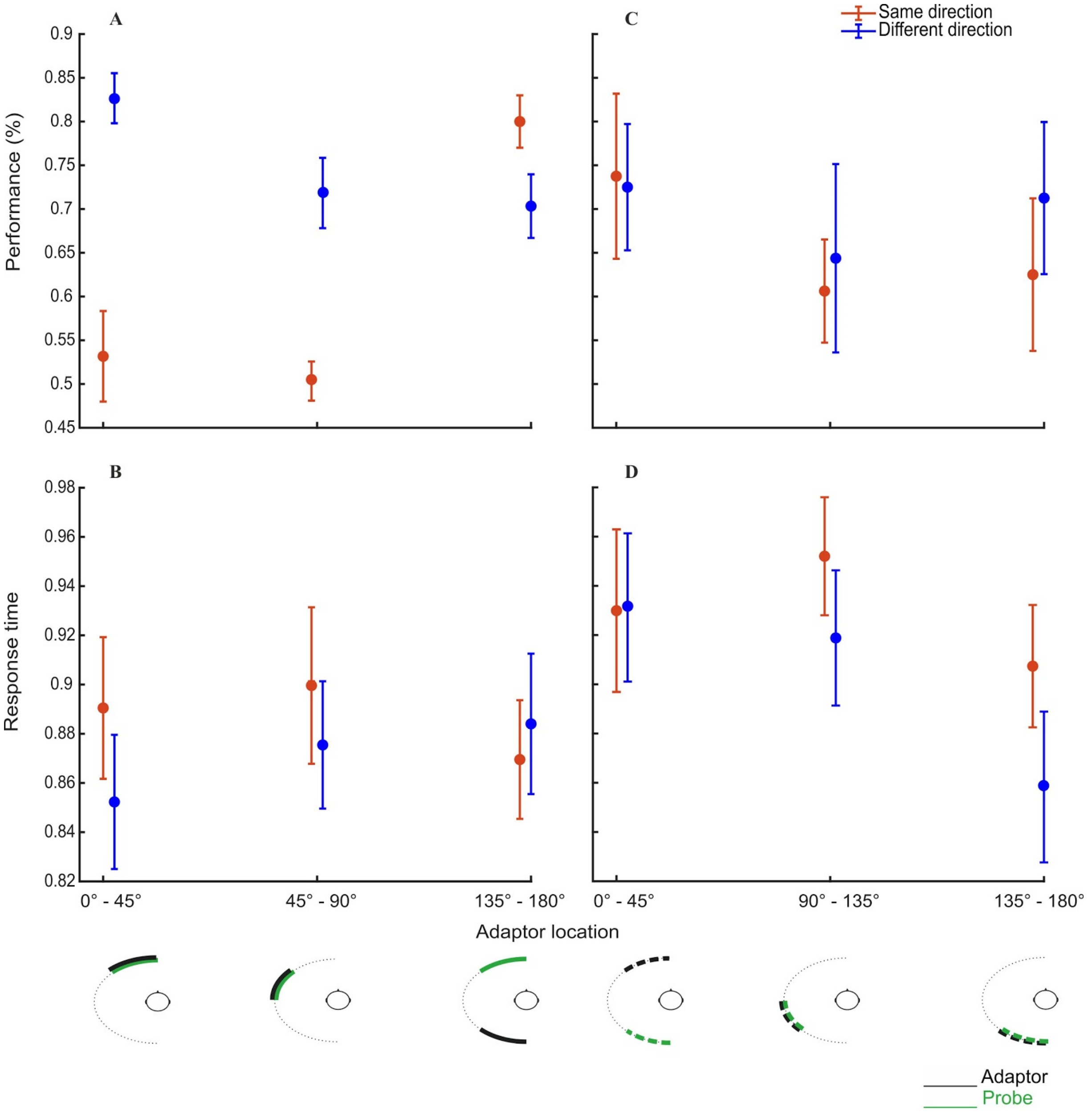


Fig. 4 Line charts presenting the results of the adaptation task after each adaptor for both experiments. The accuracy performance (**A**, **C**) and the reaction time (**B**, **D**) of the first (A, B) and second (C, D) experiments, respectively. Red lines represent trials when adaptors and probes had the same direction, i.e. both moving leftward or both moving rightward. Blue lines represent trials when adaptors and probes had opposite direction, i.e. one moving leftward and the other rightward. Vertical bars represent standard deviation. The semicircular icons at the bottom depict the location of the probe and adaptor (see Fig. 1 for larger version).

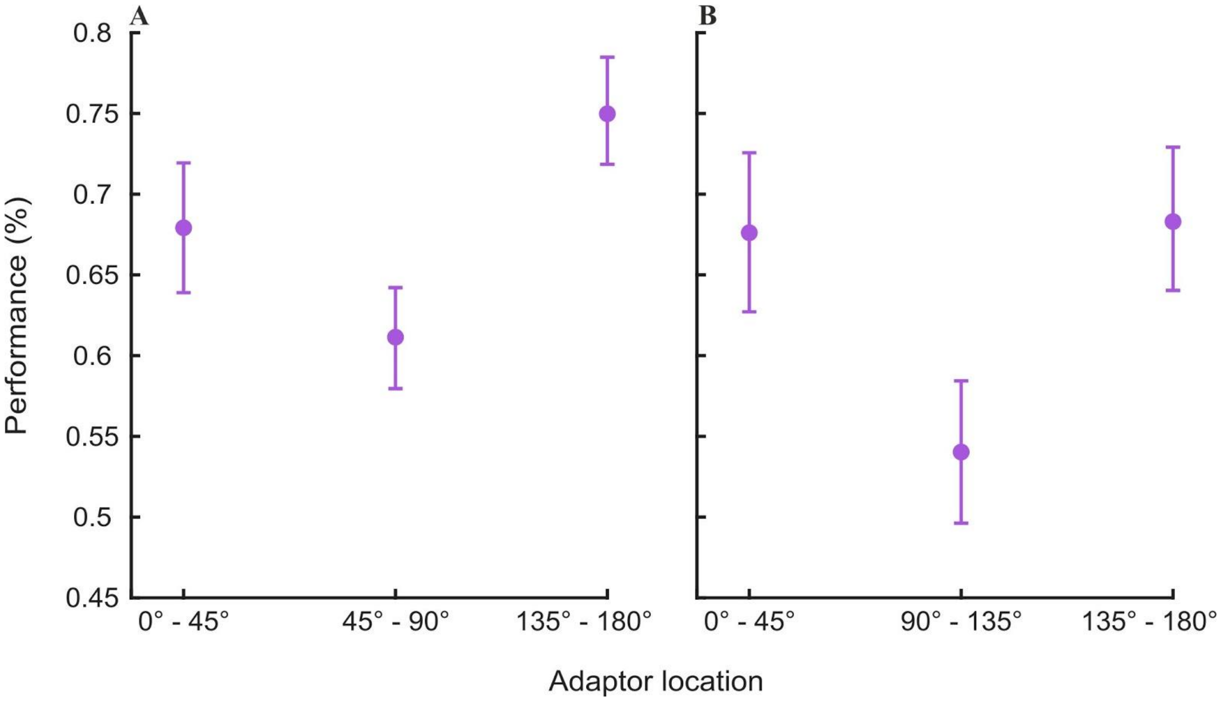


Fig. 5 Line charts showing the magnitude of the adaptation effect on the performance of both experiments by calculating the mean adaptation between adaptor-probe pairs that had the same direction (i.e. both moving leftward or both moving rightward) and adaptor-probe pairs that had different direction (i.e. one moving leftward and the other rightward) per location. **A**: The performance of the first experiment. **B**: The performance of the second experiment. Vertical bars represent standard deviation.

7.4 Discussion

In the current study, we used an adaptation paradigm to study the interaction between binaural and spectral cues for the detection of motion-direction. We used locations at the frontal and rear fields to present adaptor-probe pairs and the results showed that there is an adaptation effect on the binaural cues, even if the pair does not have the same spatial location. In addition, the adaptation effect gets stronger the highest the interaural differences get and this effect is intensified when the level of change of these cues follows the same pattern for the adaptor and the probe. In both experiments, the results showed that if the adaptor and probe span the same binaural cues, there will be an adaptation effect, reflected in diminished performance or slower reaction times, even if they are in different hemifields. The effect was stronger in the first experiment. In both experiments, though, the results highlight that adaptation is stronger when the adaptor and probe modulate the interaural cues in a similar fashion. For instance, clockwise motion at the frontal field corresponds to the same modulation of interaural cues with anticlockwise motion at the rear field and vice versa.

In both experiments, the conditions having stronger adaptation effect reflected on diminished performance were those located around the far left, corresponding to the position of the ears and the opening of the pinna. The effect was stronger at the frontal field and when adaptor and probe had the same motion-direction. The difference in these conditions effects' magnitude between the frontal and rear fields could be because of the decreased performance of the participants in the second experiment, but also, because of the interaction of spectral cues. Perception of motion at the rear field is more difficult to achieve than in the frontal field, as it is also evident from the psychometric function results (Ehrenstein 1978; Makous and Middlebrooks 1990). This result was also shown to be reflected in cortical responses, in a magnetoencephalography (MEG) study about the encoding of auditory space with stationary sounds (Salminen et al. 2009). In the aforementioned study, the researchers presented adaptor-probe pairs that differed in their spatial disparity. They compared the results from a probe presented at the frontal left hemifield and another presented at the rear left hemifield. The researchers concluded that localization of sounds is based on the hemifield code that says that the balance of activity between two sets of neurons, one tuned to the left and another tuned to the right hemifields, both found in both cortical hemispheres, aids in sound source localization.

The significance of the study is emphasized by the difference in magnitude of the adaptation effect when adaptor and probe share both the same binaural cues and the same location, versus when they share only the same binaural cues and not the location. In the first case, adaptation is based mainly on the level of binaural cues that changes with the sound's motion, while the spectral cues are similar. In the second case, adaptation is based on the integration of binaural cues with spectral cues that are different for the adaptor and the probe, because they are located in different hemifields. Should there be no effect of the spectral cues, then the magnitude of the effect in both conditions would have been the same. This effect is mirrored between the two experiments showing that adaptation is hemifield-specific.

Studies investigating the direction-specificity of the neuronal MOR at the EEG, found that there are direction encoding neurons contributing to the signal (Grzeschik et al. 2013) that are modality-specific (Grzeschik et al. 2016). In the former study, however, the adaptor-probe pairs were presented at the frontal field with the adaptor spanning a frontal trajectory including both left and right hemifields and the probe spanning at half of the trajectory of the adaptor to the left. This way, the second half of the adaptor had the same variance of binaural cues with the probe. If we employ EEG to the current experimental design, we expect that the pattern of results will also appear at the EEG response. Specifically, adaptation is expected when adaptor and probe share the same locations and interaural cues with maximal adaptation when their direction is the same, replicating the previous results. In addition, when adaptor and probe have the same binaural cues but are in different hemifields, then maximal adaptation is expected when the modulation of their binaural cues is the same, even if the direction would be opposite.

Chapter 8. General discussion and outlook

The current thesis investigated the processing and the perception of auditory motion and how this relates to dynamically changing ITDs and ILDs, as well as, their position on the azimuth. I hypothesized that since ITDs and ILDs follow separate parallel processing pathways in the brainstem and a large number of studies show that these pathways might remain separated at the cortex, it is likely that this separation is also represented in the processing of dynamically changing interaural differences. Therefore, Study 2 (and Study 1) aimed at examining this in the freefield with the presentation of moving acoustic signals of different frequency ranges. An additional analysis with the data of Study 2 aimed at examining the level of dependency of the motion-onset response on the modality of the initial part of the late-motion-onset stimulus by comparing the response to the onset of motion following a stationary part with the onset of motion following a moving part (Study 3). The results of the two previous studies showed that the components of the motion-onset response were elicited later than those reported in the literature and I wanted to confirm that this was a true result from having narrow-spectrum instead of broad-spectrum frequency for the stimuli and not based on other stimulus attributes. Thus, in Study 4, I compared responses to the onset of motion elicited from sound stimuli that had a long or brief duration and that were either moving fast or slow. Finally, I was interested to study how the binaural cues interact with the spectral cues for motion direction detection in front-back sound localization. This was investigated by employing an adaptation paradigm and by varying the direction of motion between the adaptor and the probe as well as their azimuthal location, in two studies. The behavioural effects of adaptation were presented in Study 5.

To begin with, my research supports the view that auditory motion processing is frequency-specific. The results from Study 2 and Study 3 clearly showed overall higher amplitudes on the components of the MOR as a response to high frequency than low frequency stimuli. On the one hand, this could be explained from a mechanism of stronger adaptation of neurons that encode low frequency. This adaptation is embedded within the type of stimulus used to elicit the MOR because the initial stimulus part is used as an adaptor and the release from adaptation gives rise to the MOR. This is in agreement with a previous study on aMAE (Dong et al. 2000) that showed stronger adaptation effect when the adaptor-probe pair had low than high frequency. On the other hand, these results could be explained by stronger contribution of high frequency encoding neurons to the MOR. This would suggest that motion-specific cortical areas are tonotopically organized. If, however, these results do not represent activity from motion-specific areas, then the same effect would be expected from the EOR. Yet, this was not present in the current data. Its absence is in line with previous data that show stronger amplitudes as a response to stationary sounds of low frequency for N1 (Jacobson et al. 1992), MMN (Wunderlich and Cone-Wesson 2001) and P150 (Verkindt et al. 1994) than those as a response to high frequency. It is important to mention though, that the two explanations for the frequency-specificity of the MOR are not mutually exclusive. Whereas the neurons that respond maximally to low frequency sounds might be adapted stronger, the motion-specific cortical areas might be constituted from neurons responding maximally to high frequency sounds. Also, since the stimuli had narrow band frequency and they were presented in the freefield, their processing was based on the interaural differences that form the basis of frequency processing according to the duplex theory (Rayleigh 1907). Based on this, the results of the current thesis support the separate representation of ITDs and ILDs at the cortex, which agrees with previous literature (Schröger 1996). If they were integrated, then the response to sounds with low frequency spectrum would be similar to the response to sounds with high frequency spectrum.

Moreover, my research suggests that the temporal window of the MOR component elicitation is not dependent on the adaptation phase created by the initial part of the delayed-motion-onset stimulus, nor on the velocity of the sound motion. The main result of Study 4 showed higher peak-to-peak amplitudes following the presentation of faster velocities, something that agrees with previous studies (Getzmann 2009; Shestopalova et al. 2016). Getzmann (2009) studied the effect of faster velocity on the MOR by changing the duration of the stimuli but keeping their travelled trajectory constant in the freefield and Shestopalova and colleagues (2016) studied the effect of fast angular change on the EOR by presenting slow and fast sounds and step sounds with abrupt displacements, with ITD transitions via earphones. Despite their methodological differences, both studies found stronger amplitudes as a result of faster angular shift; for the first study (Getzmann 2009), this was the very fast sound, for the second study (Shestopalova et al. 2016), this was the step stimulus. The result of Study 4 added the missing information from the literature on how the modulation of distance travelled while keeping the duration of sounds constant will affect the MOR. The surprising exploratory result that came out of Study 4 is the effect of the duration of the stationary part of the stimulus. The long duration of this part resulted in the differentiation of the cP2 peak of the MOR into two peaks, the cP2a and the cP2b. This differentiation is not usually made in the literature. The differences between stimulus and experimental design characteristics among the literature made it difficult to address its dependency. The differentiation of the cP2 was also not the focus of Study 4. However, the appearance of the two peaks led to its comparison with results from previous literature showing that the presence of a component depended on the stimulus characteristics (Garcia-Larrea et al. 1992). In the aforementioned study, if the stimulus was a target, it would lead to the elicitation of P300 after P250. Based on this exploratory result, the assumption that certain components appear after specific cognitive evaluation of the stimulus, seems to be justified, yet more research is needed to assess it.

Last but not least, my research showed that binaural and spectral cues integrate for motion direction detection and this is hemifield-specific. It was known from the literature that spectral cues are used to disambiguate locations on the azimuth that share the same ITDs and ILDs (Schnupp et al. 2011). The results from Study 5 added to this by showing that motion perception is based on the integration of the two and that adaptation of the binaural cues is possible, even if adaptor and probe do not share the same locations. The important implication from this study arose when comparing the data between the two behavioural experiments, the one that had the probes at the front and the other that had the probes at the rear hemifield. The effect was mirrored showing that motion adaptation is hemifield-specific. These data are strong support of the hemifield code as the model for spatial representation (see Introduction for details) since the data point to the existence of neuronal populations that are maximally activated from sounds presented at the locations corresponding to the ear’s pinna openings and additionally, these populations are also maximally activated when the direction of sound motion is towards these locations. These are supported from data with stationary sound adaptation (Salminen et al. 2009) and also, from data with moving sounds (Magezi and Krumbholz 2010). In addition to this, support for the hemifield code comes also from data from Study 2, since the outward movement of the stimuli elicited higher amplitudes on the MOR than the inward movement.

One last point, it is worth noting that since the data support the hemifield code with data collected with the MOR, the current thesis inevitably supports the existence of motion-specific cortical areas. If ‘snapshot’ theory would hold true, then the data from Study 2 would not show any differentiation at the MOR between the inward and outward moving stimuli. Also, the frequency specificity of the MOR was not generalized to the EOR, which points to a different mechanism that elicits the two responses. In addition, the results from Study 5 would not show differentiation between the adaptor-probe pairs that had the same motion direction and those that had different motion direction. The results would be based merely on whether they are spanning the same trajectory and not the direction of motion.

8.1 Limitations and future suggestions

Despite the novelty of the results presented in the current thesis, there are several directions that can be followed in the future in order to clarify their implication.

To begin with, the results from Study 2 that show higher MOR amplitude as a response to high than low frequency sounds, led to the understanding that this phenomenon could be either due to stronger adaptation of neurons than encode low frequency or due to high frequency tuning of motion-specific cortical areas. The earlier could be addressed with an invasive study in mammals that would adapt neurons whose best frequency varies. Then, the recovery rate of these neurons would be measured and compared among the different frequencies. A study of this kind, could also clarify to which point the neuronal recovery-from-adaptation follows a linear trend. For example, do neurons that encode middle frequencies recover faster than those that encode low frequencies and slower than those that encode high frequencies? Also, would a broadband stimulus have the same adaptation effect on these neuronal populations? The latter explanation on the results, the frequency-specificity of motion-specific cortical areas, is a topic that has not been addressed in the literature due to the long held debate of whether motion-specific areas exist. But perhaps the results of the study from Poirier and colleagues (2017) are those that the field has been waiting to proceed further. This study showed very elegantly that motion-specific areas do exist in the primate cortex and that these are the posterior belt and parabelt areas. The next step, could be the tonotopical mapping of these specific areas. Previous research showed that most neurons in the macaque parabelt area respond to high frequencies (Rauschecker et al. 1997). Also, another study that documented the tonotopicity of the human auditory cortex (Humphries et al. 2010) showed that high frequency regions are located in the posterior primary cortex, which corresponds to the macaque posterior belt area. All these results together, point to motion-specific areas in the auditory cortex that could be mostly high frequency tuned. If a study provided evidence in support of this hypothesis, then it would inevitably support the current data as well.

Moreover, the results from Study 4 that showed the differentiation of a particular peak of the MOR into two peaks, indicate that certain stimulus characteristics are represented in the MOR form. Although the study’s purpose was not to investigate this peak differentiation, it still provided enough evidence to show its comparison with earlier cortical peaks and how they relate with the processing of primary and secondary stimulus characteristics (Garcia-Larrea et al. 1992; Shestopalova et al. 2016). Since the distinction, in the current thesis, was related to how ‘quick and easy’ the difference between stationary and moving stimuli was, future studies can explore to which extend each of the two phases affect this peak differentiation. For instance, the data of Study 4 showed this differentiation as a result of longer stationary stimulus adaptation and as a result of faster motion velocities after short stationary adaptation. One the one hand, the stationary part was highlighted, and on the other hand, the motion part was highlighted. Both provided a more distinct change from the stationary to the moving part or in other words, both provided enough evidence for faster cognitive evaluation of the change between stationary and moving parts. If this distinction is indeed true, the further studies could address how motion characteristics such as velocity and direction relate to this differentiation.

8.2 Conclusion

The thesis builds on and contributes to work in the field of auditory motion perception and auditory motion-onset processing. Although a number of studies have examined its velocity- (Getzmann 2009), attention- (Kreitewolf et al. 2011), direction- (Grzeschik et al. 2013; Magezi and Krumbholz 2010) and non-crossmodal-specificity (Grzeschik et al. 2016), there has not been a strong focus on its frequency-specificity and whether it integrates with spectral cues. As such, this thesis provides additional insights about these gaps by presenting moving sound stimuli in the freefield while using EEG. This research differs from previous studies in its implications by documenting the cortical responses to narrow-band stimuli and thus basing the frequency-specificity of the responses on the ITD and ILD processing pathways (Rayleigh 1907) and also by showing that auditory motion perception is based on the activity of the neurons encoding the interaural differences and not on direction-specific neurons. In doing this it draws strongly on the work of Salminen and colleagues (2009) who found evidence of hemifield code for the encoding of auditory space and Poirier and colleagues (2017) who showed that there are motion-specific cortical areas in the primate brain.

References

Baumgart F, Gaschler-Markefski B, Woldorff MG, Heinze HJ, Scheich H (1999) A movement-sensitive area in the auditory cortex. Nat 400:724-726. doi: 10.1038/23390

Butler RA (1968) Effect of changes in stimulus frequency and intensity on habituation of the human vertex potential. The Journal of the Acoustical Society of America 44:945-950

Butler RA (1972) The influence of spatial separation of sound sources on the auditory evoked response. Neuropsychologia 10:219-225

Chait M, Simon JZ, Poeppel D (2004) Auditory M50 and M100 responses to broadband noise: functional implications. Neuroreport 15:2455-8.

Chase SM, Young ED (2005) Limited segregation of different types of sound localization information among classes of units in the inferior colliculus. J Neurosci. 25:7575-7585. doi: 10.1523/jneurosci.0915.05.2005

Crowley KE, Colrain IM (2004) A review of the evidence for P2 being an independent component process: age, sleep and modality. Clin Neurophysiol 115:732-744. doi:

10.1016/j.clinph.2003.11.021

Dong CJ, Swindale NV, Zakarauskas P, Hayward V, Cynader MS (2000) The auditory motion aftereffect: Its tunning and specificity in the spatial and frequency domains. Percept Psychophys, 62:1099-1111. doi: 10.3758/BF03212091

Ehrenstein, WH (1978) Direction-specific acoustical aftereffects. J Acoust Soc Am., 64(Suppl. 1), S35.

Freeman TCA, Leung J, Wufong E, Orchard-Mills E, Carlile S, Alais D (2014) Discrimination contours for moving sounds reveal duration and distance cues dominate auditory speed perception. PloS one 9(7): e102864. https://doi.org/10.1371/journal.pone.0102864

Garcia-Larrea L, Lukaszewicz AC, Mauguiere F (1992) Revisiting the oddball paradigm. Non-target vs neutral stimuli and the evaluation of ERP attentional effects. Neuropsychologia 30(8):723-741.

Getzmann S (2008) Effects of velocity and motion-onset delay on detection and discrimination of sound motion. Hear Res. 246:44-51. doi: 10.1016/j.heares.2008.09.007

Getzmann S (2009) Effect of auditory motion velocity on reaction time and cortical processes. Neuropsychologia 47:2625-33. doi: 10.1016/j.neuropsychologia.2009.05.012

Getzmann S (2011) Auditory motion perception: onset position and motion direction are encoded in discrete processing stages. Eur J Neurosci. 33:1339-50. doi: 10.1111/j.1460-9568.2011.07617.x

Getzmann S, Lewald J (2010) Effects of natural versus artificial spatial cues on electrophysiological correlates of auditory motion. Hear Res 259:44-54. doi:10.1016/j.heares.2009.09.021

Getzmann S, Lewald J (2011) The effect of spatial adaptation on auditory motion processing. Hear Res. 272:21-9. doi: 10.1016/j.heares.2010.11.00

Getzmann S, Lewald J (2012) Cortical processing of change in sound location: smooth motion versus discontinuous displacement. Brain Res. 1466:119-27. doi: 10.1016/j.brainres.2012.05.033

Getzmann S, Lewald J (2014) Modulation of auditory motion processing by visual motion. J Psychophysiol, 28(2):82-100. Doi: 10.1027/0269-8803/a000113

Grantham, DW (1989) Motion aftereffects with horizontally moving sound sources in the free field. Perception & Psychophysics, 45:129-136.

Gratton G, Coles MG, Donchin E (1983) A new method for off-line removal of ocular artefact. Electroencephalogr Clin Neurophysiol. 55:468-484

Griffiths TD, Rees G, Rees A, Green GGR, Witton C, Rowe D, Büchel C, Turner R, Frackowiak RSJ (1998) Right parietal cortex is involved in the perception of sound movement in humans. Nat Neurosci 1:74-79. doi: 10.1038/276

Grothe B, Pecka M, McAlpine D (2010) Mechanisms of sound localization in mammals. Physiol Rev 90:983-1012. doi: 10.1152/physrev.00026.2009

Gruber T, Grandjean D (2017) A comparative neurological approach to emotional expressions in primate vocalizations. Neurosci Biobehav Rev, 73:182-190. doi: 10.1016/j.neubiorev.2016.12.004

Grzeschik R, Böckmann-Barthel M, Mühler R, Verhey JL, Hoffmann MB (2013) Direction-specific adaptation of motion-onset auditory evoked potentials. Eur J Neurosci. 38:2557-65. doi: 10.1111/ejn.12264

Grzeschik R, Lewald J, Verhey JL, Hoffmann MB, Getzmann S (2016) Absence of dierection-specific cross-modal visual-auditory adaptation in motion-onset event-related potentials. Eur J Neurosci. 43:66-77. doi: 10.1111/ejn.13102

Haumann M, Corballis MC, Fabri M, Paggi A, Lewald J (2005) Sound lateralization in subjects with callosotomy, callosal agenesis, or hemispheroctomy. Brain Res Cogn Brain Res 25:537-546. doi: 10.1016/j.cogbrainres.2005.08.008

Hillyard SA, Picton TW (1978). On and off components in the auditory evoked potential. Perc Psychoph 24:391-398.

Jacobson GP, Lombardi DM, Gibbens ND, Ahmad BK, Newman CW (1992) The effects of stimulus frequency and recording site on the amplitude and latency of multichannel cortical auditory evoked potential (CAEP) component N1. Ear Hear 13:300-306.

Jerger J, Martin J (2004) Hemispheric asymmetry of the right ear advantage in dichotic listening. Hear Res 198:125-136. doi: 10.1016/j.heares.2004.07.019

King AJ, Schnupp JWH, Doubell TP (2001) The shape of ears to come: dynamic coding of auditory space. Trends Cogn Sci 5:261-270. doi: 10.1016/S1364-6613(00)01660-0

Kreitewolf J, Lewald J, Getzmann S (2011) Effect of attention on cortical processing of sound motion: an EEG study. Neuroimage 54:2340-9. doi: 10.1016/j.neuroimage.2010.10.031

Krumbholz K, Hewson-Stoate N, Schönwiesner M (2007) Cortical response to auditory motion suggests an asymmetry in the reliance on inter-hemispheric connections between the left and right auditory cortices. J Neurophysiol 97:1649-55. doi: 10.1152/jn.00560.2006

Langers DRM, Krumbholz K, Bowtell RW, Hall DA (2014) Neuroimaging paradigms for tonotopic mapping (I): The influence of sound stimulus type. Neuroimage 100:650-662. doi: 10.1016/j.neuroimage.2014.07.044

Lord Rayleigh OM Pres. RS (1907) XII. On our perception of sound direction. Philos. Mag. Ser. 6 74:214-232. doi: 10.1080/14786440709463595

Luck SJ (2014) Statistical analysis. In: Luck SJ (ed) An introduction to the event-related potential technique, 2nd edn. MIT Press, pp309-340

Magezi DA, Buetler KA, Chouiter L, Annoni JM, Spierer L (2013) Electrical neuroimaging during auditory motion aftereffects reveals that auditory motion processing is motion sensitive but not direction selective. J Neurophysiol 109:321-331. doi: 10.1152/jn.00625.2012

Magezi DA, Krumbholz K (2010) Evidence for opponent-channel coding of interaural time differences in human auditory cortex. J Neurophysiol 104:1997-2007. doi: 10.1152/jn.00424.2009

Makous, JC, Middlebrooks, JC (1990) Two-dimensional sound localization by human listeners. J Acoust Soc Am., 87:2188-2200.

McAlpine D, Jiang D, Paler AR (2001) A neural code for low-frequency sound localization in mammals. Nat Neurosci 4:396-401. doi: 10.1038/86049

McFadden D (1993) A speculation about the parallel ear asymmetries and sex differences in hearing sensitivity and optoacoustic emissions. Hear Res 68:143-151.

Middlebrooks JC, Green DM (1991) Sound localization by human listeners. Annu Rev Psychol 42:135-159. doi: 10.1146/annurev.ps.42.020191.001031

Ordonez-Gomez JD, Dunn JC, Arroyo-Rodriguez V, Mendez-Cardenas MG, Marquez-Arias A, Santillan-Doherty AM (2015) Role of emitter and severity of aggression influence the agonistic vocalizations of Geoffroy’s spider monkeys (Ateles geoffroyi). Int J Primatol, 36:429-440. doi: 10.1007/s10764-015-9833-5

Poirier C, Baumann S, Dheerendra P, Joly O, Hunter D, Balezeau F, Sun L, Rees A, Petkov CI, Thiele A, Griffiths TD (2017) Auditory motion-specific mechanisms in the primate brain. PLos Biol 15(5): e2001379. https://doi.org/10.1371/journal.pbio.2001379

Poirier C, Collignon O, DeVolder AG, Renier L, Vanlierde A, Tranduy D, Scheiber C (2005) Specific activation of the V5 brain area by auditory motion processing: An fMRI study. Cog Brain Res 25:650-658. doi: 10.1016/j.cogbrainres.2005.08.015

Polich J (1989) Habituation of P300 from auditory stimuli. Psychobiology 17(1):19-28

Ritter W, Vaughan HG Jr, Costa LD (1968) Orienting and habituation to auditory stimuli: a study of short term changes in averaged evoked responses. Electroencephalography and Clinical Neurophysiology 25:550-556

Salminen, NH, May, PJC, Alku, P, Tiitinen, H (2009) A population rate code of auditory space in the human cortex. PloS One, 4:e7600.

Schnupp, JW, Nelken, I, Kind A (2011) Auditory neuroscience: making sense of sound. MIT Press. Cambridge, Massachusetts. London, England.

Schröger E (1996) Interaural time and level differences: integrated or separated processing? Hear Res 96:191-198

Shestopalova LB, Petropavlovskaia EA, Vaitulevich SP, Nikitin NI (2016) Hemispheric asymmetry of ERPs and MMNs evoked by slow, fast and abrupt auditory motion. Neuropsychologia 91:465-479. https://doi.org/10.1016/j.neuropsychologia.2016.09.011

Smulders FTY, Miller JO (2012) The lateralized readiness potential. In Luck SJ, Kappenman ES (Eds), The Oxford handbook of event-related potential components (pp. 209-229). New York: Oxford University Press.

Stevens SS, Newman BB (1936) The localization of actual sources of sound. Amer J Psychol 48:297-306

Sugg MJ, Polich J (1995) P300 from auditory stimuli: intensity and frequency effects. Biol Psychol, 41:255-269.

Tian B, Rauschecker JP (2004) Processing of frequency-modulated sounds in the lateral auditory belt cortex of the rhesus monkey. J Neurophysiol 92:2993-3013. doi: 10.1152/jn.00472.2003

Teshiba TM, Ling J, Ruhl DA, Bedrick BS, Pena A, Mayer AR (2013) Evoked and intrinsic asymmetries during auditory attention: implications for the contralateral and neglect models of functioning. Cereb Cortex, 23(3):560-569. doi: http://dx.doi.org/10.1093/cercor/bhs039

Tollin DJ (2003) The lateral superior olive: a functional role in sound source localization. Neuroscientist 9:127-143. doi: 10.1177/1073858403252228

Towle VL, Bolanos J, Suarez K, Tan K, Grzeszczuk R, Levin DN, Cakmur R, Frank SA, Spire JP (1993) The spatial location of EEG electrodes: locating the best-fitting sphere relative to cortical anatomy. Electroencephalogr Clin Neurophysiol. 86:1-6

Verkindt C, Bertrand O, Thevenet M, Pernier J (1994) Two auditory components in the 130-230 ms range disclosed by their stimulus frequency dependence. Neuroreport 5:1189-92.

Warren RL, Ramamoorthy S, Ciganovic N, Zhang Y, Wilson TM, Petrie T, Wang RK, Jacques SL, Reichenbach T, Nuttall AL, Fridberger A (2016) Minimal basilar membrane motion in low-frequency hearing. Proc Natl Acad Sci U S A 113:4304-4310. doi: 10.1073/pnas.1606317113

Wightman FL, Kistler DJ (1993) Sound localization. In: Yost WA, Popper AN, Fay RR (ed) Human Psychophysics, 1st edn. Springer-Verlag, New York, pp 155-192

Woods DL, Alain C, Covarrubias D, Zaidel O (1993) Frequency-related differences in the speed of human auditory processing. Hear Res, 66:46-52

Woods DL, Alain C, Covarrubias D, Zaidel O (1995) Middle latency auditory evoked potentials to tones of different frequency. Hear Res, 85:69-75

Wunderlich JL, Cone-Wesson BK (2001) Effects of stimulus frequency and complexity on the mismatch negativity and other components of the cortical auditory-evoked potential. J Acoust Soc Am, 109(4):1526-1537. Doi: http://dx.doi.org/10.1121/1.1349184

Xiang J, Chuang S, Wilson D, Otsubo H, Pang E, Holowka S, Sharma R, Ochi A, Chitoku S (2002) Sound motion evoked magnetic fields. Clin Neurophysiol 113:1-9. https://doi.org/10.1016/S1388-2457(01)00709-X