Behavioural and neurophysiological correlates of sound segregation in adults and children with and without attention deficit hyperactivity disorder (ADHD)

Dissertation

zur Erlangung des akademischen Grades

des Doktors der Naturwissenschaften (Dr. rer. nat.)

eingereicht an der Mathematisch-Naturwissenschaftlichen

Sektion der Universität Konstanz

Fachbereich Psychologie

vorgelegt von Rosa Lipp

Tag der mündlichen Prüfung: 26.11.2010

Referentin: Dr. Isabella Paul-Jordanov

Referentin: Prof. Dr. Johanna Kißler

Danke

Zuallererst möchte ich mich bei meiner Betreuerin Dr. Isabella Paul-Jordanov für die fachliche Betreuung und Unterstützung während meiner Doktorarbeitszeit und die Begutachtung der Dissertation herzlich bedanken. Danke für Dein Engagement und Deine Zeit. Du hast mir zahlreiche Ratschläge gegeben, fachkundige Vorschläge gemacht und warst selbst während Deiner Babypause immer offen für neue Ideen und Fragen!

Im Hintergrund war Prof. Dr. Brigitte Rockstroh jederzeit verfügbar für Ratschläge und Fragen. Auch dafür ein herzliches Dankeschön! Besonderer Dank gilt Dr. Patrick Berg für Rat und Hilfe hinsichtlich der Auswertung der MEG-Daten, Ursula Lommen für die kompetente Hilfestellung während der MEG-Messungen und Dr. Christian Wienbruch für die Unterstützung bei der Stimuluskorrektur.

Viele weitere Menschen haben auf unterschiedlichste Weise zum Gelingen dieser Doktorarbeit beigetragen, sei es durch fachkräftige Beratung oder Unterstützung bei der Datenerhebung. Herzlichen Dank dem ADHS-Team: Johanna Goepel, Tanja Gihr, Mario Haas, Marieke Bechtold, Stefanie Biehl, Stefanie Kohler und einige mehr für Eure Unterstützung und ein angenehmes Arbeitsklima. Dank auch an Dagmar Moret, Daria Antonenko, und weitere Mitarbeiter des ZPR, sowie an die Kooperationspartner der University of York, vornehmlich Dr. Peter Bailey und Dr. Padraig Kitterick. Außerdem gilt mein besonderer Dank allen Versuchsteilnehmern − ob groß oder klein − ohne die diese Doktorarbeit nicht möglich gewesen wäre und den Eltern der Kinder für ihre Bereitschaft die Forschung über ADHS zu unterstützen.

Bei meinen Eltern, meinem Freund, Freunden und Bekannten möchte ich mich für die moralische Unterstützung und Motivierung bedanken, quasi das A und O einer jeden Arbeit;)

Last but not least danke ich Prof. Dr. Johanna Kißler und Prof. Dr. Harald Schupp für die Bereitschaft, die Dissertation zu begutachten bzw. die mündliche Prüfung abzunehmen.

Contents

List of Abbrevations **III**

Summary **V**

Zusammenfassung **VII**

1 General introduction 1

1.1 Auditory processing 1

1.1.1 Auditory Scene Analysis (ASA) 1

1.1.2 Concurrent sound segregation vs. auditory streaming 3

1.1.3 Cues for concurrent sound segregation 4

1.1.4 Neurophysiological markers for concurrent sound segregation 5

1.1.5 Auditory coding and neural networks involved in concurrent sound perception 6

1.1.6 Influence of attention 8

1.1.7 Developmental aspects of ASA 10

1.2 Attention deficit hyperactivity disorder (ADHD) 11

1.2.1 Diagnostics and prevalence 11

1.2.2 Causal models of ADHD 12

1.2.3 Selected aspects of etiology 13

1.2.4 Event-related potentials (ERPs) 15

1.2.5 ADHD and auditory perception 16

1.2.6 ADHD and (central) auditory processing disorder 17

1.3 The present project 18

2 Studies 22

2.1 Concurrent sound segregation based on inharmonicity and onset asynchrony (Study 1) 22

2.1.1 Introduction (Study 1) 22

2.1.2 Methods (Study 1) 26

2.1.3 Results (Study 1) 32

2.1.4 Discussion (Study 1) 38

2.2 Developmental changes in concurrent sound segregation (Study 2) 44
2.2.1 Introduction (Study 2) 44
2.2.2 Methods (Study 2) 49
2.2.3 Results (Study 2) 54
2.2.4 Discussion (Study 2) 63
2.3 Concurrent sound segregation in children with and without attention deficit hyperactivity disorder (ADHD) (Study 3) 69
2.3.1 Introduction (Study 3) 69
2.3.2 Methods (Study 3) 75
2.3.3 Results (Study 3) 81
2.3.4 Discussion (Study 3) 90

3 General discussion 94
3.1 AEFs associated with concurrent sound segregation 94
3.2 Maturation of concurrent sound segregation 95
3.3 Concurrent sound segregation in children with ADHD 96
3.4 Concluding remark and perspective 98

4 References 99

5 Indices 109
5.1 Figure index 109
5.2 Table index 111

List of Abbrevations

ACPT	auditory continuous performance test
ADD	attention deficit disorder
ADHD	attention deficit hyperactivity disorder
ADHD-C	combined subtype
ADHD-HI	predominantly hyperactive and impulsive subtype
ADHD-I	predominantly inattentive subtype
ADHS	Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung
AEF	auditory evoked field
AEP	auditory evoked potential
AM	amplitude modulation
ANOVA	analyses of variance
ASA	auditory scene analysis
ASHA	american speech-language-hearing association
AUDIVA	questionnaire on auditory development
CAEP	cortical auditory evoked potential
CAP	central auditory processing
(C)APD	(central) auditory processing disorder
CBCL	child behavior checklist
C-DT	choice-delay task
CNS	central nervous system
CPT	continuous performance test
D2	Aufmerksamkeits-Belastungs-Test
DA	dopamine
DAT	dopamine transporter
DRD4	dopamine receptor D4 gene
DSM	diagnostic and statistical manual of mental disorders

DISYPS	diagnostic system of psychiatric disorders in children and adolescents
EEG	electroencephalography/electroencephalogram
EF	executive functions
ERN	error related negativity
ERP	event-related potential
ESPm	early segregation positivity
f_0	fundamental frequency
FD	frequency discrimination
FM	frequency modulation
HRTF	head-related transfer function
ICD	international classification of diseases
ISI	inter-stimulus interval
KITAP	Testbatterie zur Aufmerksamkeitsprüfung für Kinder
KLT-R	Konzentrations-Leistungs-Test
MANOVA	multivariate analysis of variance
MEG	magnetoencephalography
MGN	medial geniculate nuclei
MMN	mismatch negativity
MPH	methylphenidate
NC-ERP	auditory evoked neutral condition event-related potentials
NE	norepinephrine
LSPm	late segregation positivity
ORNm	object related negativity
PN	processing negativity
qEEG	quantitative electroencephalography
RD	reading disability
SST	stop-signal task

Summary

Most of the time, the auditory environment is complex with many simultaneously active sound sources. Yet, listeners are able to assign incoming acoustic elements to separate perceptual auditory objects. The present thesis investigated these concurrent sound segregation skills in adults and children with and without attention deficit hyperactivity disorder (ADHD) by measuring auditory evoked fields (AEFs) using magnetoencephalography (MEG). The aim of the thesis was to learn more about neurological markers of concurrent sound segregation in adults (Study 1), the maturational changes underlying concurrent sound segregation (Study 2) and possible auditory processing deficits in children with ADHD (Study 3), as children with ADHD and children with auditory processing disorders show strikingly similar symptoms. Participants were presented with complex sounds with the third harmonic mistuned by 8% of its original value and the onset of the third harmonic delayed by 160 ms compared to the other harmonics.

Both cues - inharmonicity and onset asynchrony - elicited sound segregation in adults and children with and without ADHD. However, ADHD children performed worse than control children, who in turn performed worse than adults. Three AEFs were identified in adults and children: the *early segregation positivity* (ESPm), the *object related negativity* (ORNm), and the *late segregation positivity* (LSPm), which seem to reflect relatively automatic, bottom-up sound segregation processes. Some differences concerning occurrence and peak latencies of the AEFs between adults and children reflected the behavioral results insofar as automatic sound segregation processes seem to be not fully developed in children. In contrast to adults, in children the ESPm seems to be a more general marker for a bottom-up sound segregation process, whereas in adults the ESPm probably reflects an automatic registration of the *mistuned* harmonic. However, the ORNm was found to be adult-like in children, elicited by inharmonicity and asynchrony, and was therefore interpreted as a general marker of concurrent sound segregation in both groups. The comparison of children with ADHD and control children supports the idea that at least a sub-set of children with ADHD might have some impairment in the pre-attentive processing of sounds, which is indicated by the unstable expression of the ESPm in children with ADHD when inharmonicity serves as cue for concurrent sound segregation.

Together, the results of the present thesis suggest that research on neurological markers of concurrent sound segregation has the potential to serve for clinical diagnostics of genuine auditory processing deficits in children with attention problems in future.

Zusammenfassung

Das uns umgebende auditorische Umfeld ist oftmals komplex im Sinne vieler gleichzeitig aktiver Geräuschquellen. Dennoch ist man als Zuhörer in der Lage, die einströmenden akustischen Elemente einzelnen perzeptuellen auditorischen Objekten zuzuordnen. Im Rahmen der Doktorarbeit wurden diese Fähigkeiten einer simultanen Lautsegregation bei Erwachsenen, sowie Kindern mit und ohne Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung (ADHS) untersucht, indem mittels Magnetenzephalographie (MEG) auditorisch evozierte (Magnet-)Felder (AEF) abgeleitet wurden. Das Ziel dieser Arbeit war es, mehr über neurologische Marker der simultanen Lautsegregation bei Erwachsenen (Studie 1), Reifungsprozesse der simultanen Lautsegregation (Studie 2), sowie über mögliche Defizite in der auditorischen Verarbeitung bei Kindern mit ADHS (Studie 3) zu lernen, da letztere eine zu Kindern mit auditorischer Wahrnehmungs- und Verarbeitungsstörung auffallend ähnliche Symptomatik aufweisen. Den Studienteilnehmern wurden komplexe harmonische Töne präsentiert, bei denen die dritte Harmonische entweder um 8% von ihrem ursprünglichen Wert verstimmt oder der Präsentationsbeginn um 160 ms, verglichen mit den übrigen Harmonischen, verzögert war.

Beide Signalvarianten – verletzte Harmonizität und asynchroner Beginn – riefen Lautsegregation bei Erwachsenen und Kindern mit und ohne ADHS hervor. Allerdings waren ADHS-Kinder schlechter in ihrer Lautsegregationsleistung als Kontrollkinder, und diese waren wiederum schlechter als Erwachsene. Drei AEFs konnten bei Erwachsenen und Kindern identifiziert werden: eine frühe Positivierung (*early segregation positivity*, ESPm), eine sogenannte objekt-bezogene Negativierung (*object related negativity*, ORNm) und eine späte Positivierung (*late segregation positivity*, LSPm), welche alle relativ automatische (*bottom-up*) Prozesse der Lautsegregation zu reflektieren scheinen. Einige Unterschiede bezüglich des Auftretens und der Latenzen der AEFs zwischen Erwachsenen und Kindern, spiegeln die Verhaltensleistungen insofern wieder, als dass automatische Prozesse der Lautsegregation bei Kindern noch nicht voll entwickelt zu sein scheinen. Im Gegensatz zu Erwachsenen scheint die ESPm bei Kindern ein allgemeiner Marker für einen *bottom-up* Lautsegregationsprozess zu sein, wohingegen die ESPm bei Erwachsenen eine automatische Registrierung der *verstimmten* Harmonischen widerspiegelt. Jedoch war die ORNm von Kindern der von Erwachsenen sehr ähnlich, wurde jeweils durch verletzte Harmonizität und asynchronen Beginn hervorgerufen und daher als allgemeiner Marker der

simultanen Lautsegregation für beide Gruppen interpretiert. Der Vergleich zwischen ADHS- und Kontrollkindern spricht dafür, dass ein Teil der ADHS-Kinder Beeinträchtigungen bei aufmerksamkeitsunabhängigen Verarbeitungsprozessen von Tönen haben könnte, was sich durch ein instabiles Auftreten der ESPm in Kindern mit ADHS ausdrückt, wenn verletzte Harmonizität als Signal für simultane Lautsegregation dient.

Zusammenfassend befürworten die Ergebnisse dieser Arbeit eine Intensivierung der Erforschung neurologischer Marker der simultanen Lautsegregation, da letztere das Potential haben könnten, in Zukunft als Instrument in der klinische Diagnostik echter auditorischer Verarbeitungsdefizite bei Kindern mit Aufmerksamkeitsproblemen zu dienen.

1 General introduction

1.1 Auditory processing

1.1.1 Auditory Scene Analysis (ASA)

In most everyday situations, more than one sound source is active at any given moment. From birth on we are surrounded by a noisy environment posing a high demand on the auditory system. Sounds originating from different sources (e.g. radio, television, traffic) and voices of different speakers need to be distinguished, but the sound that reaches the ear is a summation of the pressure waves from the individual sources. The computational problem confronting the auditory system is to interpret this complex waveform as distinct auditory events. How does the brain manage to do so? More precisely, how does the auditory system assort which elements of the mixture belong to a particular sound source and which originate from a different sound source?

The processing stages, which sort the incoming acoustic information into one or more neural representations of auditory objects, can be summarized under the term "Auditory Scene Analysis" (ASA, Bregman, 1990). A segmentation process decomposes a complex acoustic scene into a collection of distinct sensory elements, whereas a grouping process combines these elements into an auditory object according to some principles. Motivated by Gestalt psychology, Bregman (1990) has suggested several grouping principles (Haykin & Chen, 2005):

1. *Proximity* – if two tones are closer in time, frequency and intensity, then it is more likely that they should be grouped together
2. *Similarity* – depends on sound properties, such as timbre
3. *Continuity* – states that an object`s sound does not make rapid jumps, but instead continues smoothly (closely related to proximity)
4. *Common fate* – groups together activities (onsets, glides, or vibrato) that are synchronous
5. *Closure* – leads to percepts of completion. For example, listeners may hear a tone continuing through noise, even though the tone is not present during the noise.

Two mechanisms are essentially involved in auditory perception: sound localization ("where") and sound recognition ("what"). Interaural time difference (for low frequencies) and interaural level difference (for high frequencies) are the main cues for localising sounds in the azimuthal plane, whereas the main cues for vertical localization are provided by diffraction and reflection properties of the head, pinna, and torso, described by the head-related transfer function (HRTF). Intensity (loudness) and early reflections are probable cues for localization as a function of distance (Haykin & Chen, 2005). Sound recognition – the question of "what" – mainly concerns the processes of sound segregation and identification. Acoustic features (cues) used for sound recognition are for example onsets/offsets, amplitude modulation (AM), frequency modulation (FM), pitch, timbre (tone color) and harmonicity (Haykin & Chen, 2005). Although it is important in sound localisation, spatial separation is not considered a strong cue for sound segregation (Bregman, 1990; Darwin & Carlyon, 1995). Similarly, AM and FM seem to be negligible cues for sound segregation (Grossberg, Govindarajan, Wyse, & Cohen, 2004).

In the model described by Bregman (1990), ASA is broadly proposed to be a two-stage process that helps to decide which components belong to a particular sound source. First, the acoustic signal is analysed pre-attentively by "primitive grouping mechanisms", which are thought to be largely independent of the listener`s attention focus or prior experience, being driven by the physical properties of the stimulus (bottom-up). Primitive grouping mechanisms seem to mature rapidly in humans (Demany, 1982; McAdams & Bertoncini, 1997; Winkler et al., 2003) and are common to both humans and animals (Hulse, MacDougall-Shackleton, & Wisniewski, 1997; Izumi, 2002). Second, the outcome of the primitive grouping mechanism is subjected to "schema-governed mechanisms" that allow the listener to selectively attend to specific auditory objects in the signal. Schema-governed mechanisms are thought to be learnt, and to use prior knowledge to group auditory objects in the acoustic data (top-down). According to Bregman`s model, primitive grouping mechanisms would therefore be largely responsible for the automatic segregation of a mistuned harmonic from an otherwise harmonically-synchronous complex of pure tones (for a detailed description of harmonic complex tones, see section *1.1.3. Cues for concurrent sound segregation*) (e.g. Moore, Glasberg, & Peters, 1986), whereas schema-governed mechanisms are thought to be responsible for results of a study from Bey & McAdams (2002). They showed that the ability to recognize a target melody, which was interleaved with distractor tones, improved when participants had prior knowledge of the

target melody by having heard it before. The authors concluded that top-down processes allowed listeners to select information from a mixture by a matching process between schemas stored in memory and a sensory representation. The combination of bottom-up and top-down processes is thought to be an optimal adaptation of the cognitive system to its environment: a system that takes into account the sensory input and previously acquired knowledge (Bey & McAdams, 2002).

1.1.2 Concurrent sound segregation vs. auditory streaming

When regarding the spectrogram´s representation of a mixture of sounds, the brain has to solve the problem of perceptually separating the various components, which can be seen to have two dimensions, the vertical and the horizontal. Thus, when discussing auditory sound segregation one has to mention two broad classes of phenomenon. One, known as "auditory streaming", concerns the perceptual organization of sounds over time (horizontal), whereas the other, known as "concurrent sound segregation", concern the segregation of sounds that overlap in time (vertical) (Carlyon, 2004). That is, in the horizontal dimension, the brain has to group those frequency components that have come from the same auditory object over time, whereas in the vertical dimension it has to group together the particular frequency components that stem from the same auditory object, from among all the ones present at the same time. For example, the simultaneous grouping process will group sounds together if they have synchronous onsets and offsets, or if they are harmonically related. In contrast, auditory stream segregation takes several seconds to build up and can be induced by a difference in the frequency range of two sets of interleaved pure tone patterns such as an "ABA—ABA—" pattern in which "A" and "B" are tones of different frequencies separated by a silent interval. The greater the frequency separation and the faster the triplets "ABA" are repeated the quicker listeners reach an asymptote for reporting to have heard two separate streams of sounds (Fig. 1) (for a review see Alain & Bernstein, 2008).

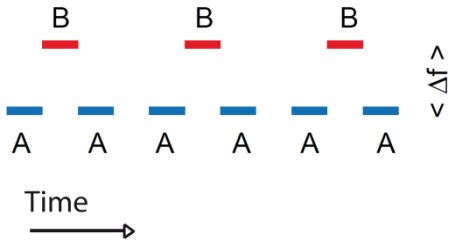

Figure 1: Schematic spectrogram of galloping rhythm sequence used to study streaming. [Adapted from Carylon (2004)].

1.1.3 Cues for concurrent sound segregation

Concurrent sound segregation is mainly based on two cues: harmonicity and onset/offset asynchrony (Darwin & Carlyon, 1995). Periodic sounds (e.g. speech and music) consist of a fundamental frequency (f_0) and partials with frequencies that are integer multiples of that fundamental, called harmonics. This harmonicity gives rise to a highly integrated percept. A complex tone consisting of multiple harmonics is heard as a single, buzz-like sound. However, Moore, Glasberg, & Peters (1986) showed that when the frequency of a single partial was mistuned between 1 and 3% (depending on the participant), it "popped out" from the complex and was audible as a separate tone. The processing mechanisms underlying the perception of the mistuned harmonic as a separate sound are not fully understood but likely involve low-level processes, which take place along the ascending auditory pathway including the cochlear nucleus, the inferior colliculus, and the primary auditory cortex (for a review see Alain & Bernstein, 2008). The concept of a harmonic sieve or template can be used to model the detection of mistuning. Harmonics that are integer multiples of the f_0 would "pass" through the sieve and be grouped into one sound object, while the mistuned harmonic would be attributed to another object. The activation of such a template could involve neurons sensitive to equal spacing between tonal elements (e.g., frequency periodicity), which could act as a series of filters that allow harmonically related partials to group together with the f_0 (for a review see Alain, 2007).

The other stimulus property, with particular relevance to the present thesis, is onset asynchrony. That is, components that start or stop at the same time are more likely to have originated from the same sound source than components that start or stop at different times. For example, if a single frequency component starts sufficiently earlier or ends later than

the other components of the harmonic complex it is perceived as a separate auditory object, being more easily pulled into an auditory stream (Bregman & Pinker, 1978; Dannenbring & Bregman, 1978). Similarly, Hill and Darwin (1996) showed that when the onset of a harmonic component is delayed, it is perceived as separate from the other harmonics.

1.1.4 Neurophysiological markers for concurrent sound segregation

Evidence from scalp recordings of event-related brain potentials (ERPs) (Alain, Arnott, & Picton, 2001; Alain & Izenberg, 2003; Alain, Schuler, & McDonald, 2002) suggests that concurrent sound segregation occurs independently of attention to the sounds. Participants were presented with auditory stimuli, which either promoted sound segregation leading to the perception of two rather than one sound, or did not promote sound segregation leading to the perception of one sound. The presented sounds comprised multiple harmonics of 200 Hz, one of which could be mistuned. In one condition the third harmonic was mistuned by 1-16% (Alain et al., 2001; Alain et al., 2002) causing the third harmonic to stand out perceptually from the sound complex formed by the other harmonics (Mistuning condition), whereas in another condition it was not manipulated (Baseline condition). Increasing the amount of mistuning increased the likelihood that participants would report the perception of two auditory objects. The authors identified a negative deflection in the difference waveform (Mistuning – Baseline), called the object-related negativity (ORN) and another later, positive deflection around 400 ms, labelled P400. The ORN was present when participants attended to the sounds and also when their attention was distracted. Therefore, the authors concluded that the ORN reflects automatic, bottom-up sound segregation and that its generation is minimally affected by attention load (Alain & Izenberg, 2003). Moreover, the ORN was not affected by manipulations of visual attention load (Dyson, Alain, & He, 2005) and only little affected by sound duration (Alain et al., 2002). In contrast, the P400 was only present in the active listening condition. The authors therefore concluded that the P400 reflects awareness-related top-down processes of sound segregation.

The first published study to directly report a magnetic equivalent of the ORN was done by Alain and McDonald (2007), who measured auditory evoked fields (AEFs) while participants were presented with complex sounds – again harmonics were either in tune or the third partial was mistuned. Participants were not required to pay attention to the sounds.

The source difference waveform (Mistuned (16%) – Tuned) generated an early positivity P80 (~80 ms), an ORN (~160 ms) and a positive displacement P230 (~230 ms) after stimulus onset in young adults. The ORN and P230 amplitude correlated with the perceptual judgement of hearing one or two sounds, which was measured separately after the MEG experiment. The P80 was not correlated with perceptual judgement and was thought to reflect an early registration of inharmonicity in primary auditory cortex. The authors concluded that all three components likely reflect automatic sound segregation that may occur independently of listeners´ attention.

It is important to mention that the ORN was elicited by stimuli, which were presented equiprobably (e.g. there was no standard or deviant per se) (Alain et al., 2001). This differentiates the ORN from another component called the mismatch negativity, or MMN, which is similar in peak latency (170 ms) and can also be recorded even if participants ignore the sounds (Sams, Paavilainen, Alho, & Naatanen, 1985). In contrast to the ORN, it is only elicited by rare deviant sounds embedded in a sequence of homogeneous standard stimuli. Thus, the MMN indexes a mismatch between the expected stimulus based on the previously occurring stimuli and the actual incoming stimulus. On the contrary, the ORN generation reflects a discrepancy between the mistuned harmonic and the harmonic template that is presumably extrapolated from the incoming stimulus (Alain, 2007).

1.1.5 Auditory coding and neural networks involved in concurrent sound perception

The initial tonotopic representation (spatial mapping of sound frequency) in the inner ear, namely the basilar membrane of the cochlea, is the basis of many aspects of auditory coding. Via hair cells and auditory nerve fibers, auditory activation reaches the first stage of central processing, the cochlear nucleus. From the cochlear nucleus, information is sent to other subcortical areas prior to reaching the auditory cortex. These subcortical auditory nuclei preserve the tonotopic mapping and integrate information across the two ears and across different frequency regions (for a review see Snyder & Alain, 2007). In monkeys in has been found that the auditory cortex is organized in three distinct auditory cortical regions (Kaas & Hackett, 2000): (a) a core region consisting of three primary auditory areas in the medial portion of the superior temporal plane; (b) a belt region consisting of two secondary areas surrounding the core; and (c) a parabelt region consisting of two secondary

auditory areas. The core regions receive separate inputs from the auditory thalamic nucleus (medial geniculate body), demonstrating parallel processing of acoustic information at the earliest stage of auditory cortex and project to the belt areas. The belt areas receive also input from the auditory thalamus and project to the parabelt areas. Both project to the frontal lobe and modality nonspecific temporal regions. Numerous feed-forward and feedback connections within and between auditory cortex and other cortical and subcortical brain regions have been found. In humans the organization is similar to that of old-world monkeys, with distinct core, belt and parabelt regions, found in Heschl`s gyrus (primary auditory cortex) on the superior temporal plane and planum temporale (secondary auditory cortex) (Snyder & Alain, 2007).

It is very likely that concurrent sound perception concerns a widely distributed network of brain areas. However, due to some recent research, which delineates more precisely the brain regions involved in concurrent sound segregation, the view evolved that auditory cortex in or near Heschl`s gyrus as well as the planum temporale play an important role in sound segregation (for a review see Alain, 2007). Neural generators of concurrent acoustic objects were found, being located along the superior temporal plane in auditory cortices (i.e., primary auditory cortex) (Alain & McDonald, 2007; Dyson & Alain, 2004). For example, dipole source modelling of electroencephalography (EEG) data suggests that the ORN source is consistent with generators in auditory cortices within the Sylvian fissure (Alain et al., 2001). Furthermore, Dyson & Alain (2004) found that the Pa, a middle latency auditory evoked potential at 30 ms after stimulus onset, was significantly larger when the third harmonic was mistuned compared to when it was in tune. The enhanced Pa amplitude was related to an increased likelihood in reporting two concurrent auditory objects. These findings are consistent with an early stage of concurrent sound perception in which acoustic features (e.g. mistuning) act as pre-attentive segregation cues. Thus, the primary auditory cortex (main source of the Pa wave) seems to play an important role in sound segregation.

What is the code for auditory objects? For forming an auditory object, the neural activity of separate processing units has to be combined otherwise auditory objects with multiple stimulus dimensions cannot be established. Precise temporal coherence of neural activity in different neural assemblies might be one cue utilized by the auditory system to encode auditory objects (Klump, van Hemmen, & Sejnowski, 2006).

1.1.6 Influence of attention

In the auditory system the effect of top-down feedback can go down all the way to the outer hair cells in the cochlea via the midbrain structure. The bottom-up signals received from the hair cells are sent to medial geniculate nuclei (MGN) in the thalamus and farther up to the auditory cortex through the thalamocortical pathways. Top-down signals from the cortex are sent back to the hair cells through the corticothalamic pathways to strengthen the signal stream of interest (Haykin & Chen, 2005; Wood & Cowan, 1995). Therefore, the potential for attention processes to influence processing of a sound signal of interest in audition is large.

Single-neuron level and population level

The effect of the efferent (top-down) neural system on frequency tuning has been investigated at single-neuron level. Polley, Steinberg, & Merzenic (2006) examined, if bottom-up sensory inputs or top-down task-dependent mechanisms control cortical reorganization in adult rats. The authors used identical auditory stimuli, but different attention tasks (attention on frequency or intensity) in two groups of rats. The results demonstrated that top-down signals played an important role in the reorganization of primary and secondary auditory cortex. Attention focused on frequency cues seemed to intensify efferent neural inputs and expand the representation of the target frequency range within the cortical tonotopic map. The influence of attention on auditory processing might be tuned by both enhancement of neural responses corresponding to task-relevant stimuli (gain) and suppression of task-irrelevant neural activities (sharpening), contributing to finer neural population-level coding for attended auditory signals. Concerning the human auditory cortex a MEG-study of Okamoto et al. (2007) could show that focused auditory attention cannot only amplify neural activity in general, but can also sharpen frequency tuning at population-level in the lateral auditory cortex, possibly via the inhibitory system by suppressing surrounding neural activity resulting in improved spectral contrast. Enhancement and sharpening of population-level neural responses was reflected by significantly larger N1m responses during focused auditory attention compared with distracted listening. The authors concluded that auditory cortical neurons seem to be influenced by both: bottom-up physical sound features and top-down attention influences. Bottom-up and top-down neural processes seem to affect excitatory and inhibitory neural networks within the human auditory cortex.

Attention and Auditory Scene Analysis

In the context of ASA two kinds of attention processes are often involved. Selective attention, in which the listener attends to one particular sound source and ignores the others and divided attention, in which the listener attends to more than one sound source (Haykin & Chen, 2005). Originally, auditory streaming was thought to be innate and largely independent of attention because it was found in infants (Demany, 1982), newborns (McAdams & Bertoncini, 1997; Winkler et al., 2003), birds (Hulse et al., 1997) and monkeys (Izumi, 2002). However, recent evidence suggests that the process of streaming is affected by attention. Cusack and colleagues (2004) showed that when attention is focused on distracting sounds, build-up in auditory streaming is reduced and that this effect likely varies according to the demands of the task, with inhibition of build-up being greater the more additional demands the distracting task required. It was further found that there was little effect of whether the tone sequence and the distracting noises were in the same frequency region and whether they were in the same or in the different ears, showing that inhibition is not due to a general suppression of one side of the ascending auditory pathway. When attention was switched to a different object for a brief period, the streaming of the unattended streams seemed to be reset. The authors proposed a hierarchical decomposition model, implying that there is some automatic segregation, and then the further build-up of streaming is prevented outside of the stream that is the current focus of selective attention. For example, when a listener is attending to a speaker against a background of music and traffic noise, then the auditory system is not fragmenting the sounds of different car engines.

Although, there is evidence that stimulus-driven sensory mechanisms can be modified under certain circumstances (e.g. depending on task demand and acoustic characteristics of the sounds) by attention processes (Sussman, Winkler, Huotilainen, Ritter, & Naatanen, 2002), which suggests that top-down and bottom-up effects of sound organization may interact at an early stage of auditory processing, there is further evidence that attention is not always required for the formation of auditory streams (Sussman, Horvath, Winkler, & Orr, 2007) and that at least some aspects of auditory stream segregation do not require focused attention (Macken, Tremblay, Houghton, Nicholls, & Jones, 2003).

Consistent with results concerning auditory streaming, there is evidence for a pre-attentive mechanism of concurrent sound segregation. Most importantly, ORN generation is

little affected by attention demands, thus probably reflecting automatic, bottom-up sound segregation (Alain & Izenberg, 2003; Dyson et al., 2005). Furthermore, additional neuromagnetic markers, the P80 and P230, were elicited in the mistuned harmonic paradigm, when participants were not required to pay attention or respond to the stimuli, again indicating a bottom-up sound segregation process. This pre-attentive analysis of stimuli may help to focus attention to a subset of stimuli (Alain & Arnott, 2000). For example, when the outcome of the pre-attentive system reveals more than one sound source then attention can be efficiently allocated to only one of these sources. However, under certain circumstances the pre-attentive process indexed by the ORN may be facilitated by top-down controlled processes. For example, listening situations that promoted selective attention to the frequency region of the mistuned harmonic (active listening) generated a larger ORN than during passive listening (see experiments 1 and 3 from Alain et al., 2001). Accordingly, the implication of attention in bottom-up segregation mechanisms of an auditory object cannot be ruled out entirely, emphasizing the intimate link between perception and attention. Moreover, the influence of attention on concurrent sound segregation has been documented by the P400, which was only present when participants actively listened to sounds (Alain et al., 2001). This is consistent with Bregman`s theoretical framework of a top-down mechanism in ASA.

1.1.7 Developmental aspects of ASA

Changes in the morphology, scalp topography, latency, amplitude and stimulus-dependence of cortical auditory evoked potentials (CAEP), P1, N1, P2 and N2 from the newborn period to adolescence reflect the long-lasting development of low-level cortical processing of various auditory stimuli in humans (for a review see Wunderlich & Cone-Wesson, 2006) and many studies further demonstrated maturational development of specific auditory processes. For example, regarding temporal encoding of information in auditory cortex, which is crucial for speech perception, Trehub, Schneider, & Henderson (1995) showed that infants had significantly higher gap detection thresholds (11 ms) than 5-year-old children (5.6 ms), and children had significantly higher thresholds than adults (5.2 ms).

Demany et al. (1982) demonstrated that already in infants (7-15-weeks) stream segregation processes based on spectral similarity are functioning and that young infants are able to detect a change in the order of tones of certain melodic sequences. Although the

ability of streaming was found in further studies in newborns (McAdams & Bertoncini, 1997; Winkler et al., 2003), indicating that the basic mechanisms of auditory stream segregation seem to be innate, the babies required larger separations between acoustic dimensions and slower paced stimuli to perceive separated sound sources than adults do (Demany, 1982; McAdams & Bertoncini, 1997). Therefore, maturational refinements of neural processes necessary for adult-like sound segregation seem to be ongoing during childhood. For example, Sussman & Steinschneider (2009) demonstrated that 9 – 12-year-old children required a larger frequency separation than adults to perceive two streams. Results suggested that differences in stream segregation between children and adults reflect an under-development of basic auditory processing mechanisms.

Alain and colleagues investigated concurrent sound segregation, measuring the electroencephalogram (EEG) of 8 – 12-year-old children (Alain, Theunissen, Chevalier, Batty, & Taylor, 2003), while they were presented with complex sounds. The mechanisms underlying concurrent sound segregation seemed to be adult-like, as children as well as adults were more likely to perceive the mistuned harmonic as a separate sound when the level of mistuning was increased (0%, 2%, 4%, 8%, and 16%). But children performed worse in detecting mistuning across all levels than adults. The authors found an ORN but no P400 in children and concluded that the bottom-up segregation of concurrent sounds based on mistuning is adult-like in pre-adolescent children, but children are less efficient than adults in processing the information following the detection of mistuning. These results further indicate that sound segregation matures during adolescence.

1.2 Attention deficit hyperactivity disorder (ADHD)

1.2.1 Diagnostics and prevalence

Children diagnosed having attention deficit hyperactivity disorder (ADHD) are characterized by behavioural symptoms such as a deficient ability to sustain attention and concentration, an overly hyperactive pattern of behaviour and impaired control of impulses. For example, they cannot sit still and often interrupt other people. Diagnosis is based on two international classification instruments, namely the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV, APA, 1994) and the International Classification of Diseases (ICD-10, 1996). Both classifications include symptoms of inattention, hyperactivity, and impulsivity that cause impairment to school performance, intellectual functioning, social

skills, and occupational functioning (Biederman & Faraone, 2005). It is specified in the DSM-IV that the core symptoms have to be apparent before the age of seven, need to be present in more than one setting (e.g. school and home) and have to be prevalent for at least six months (DSM-IV, APA, 1994). Criterions for exclusion are the concomitance of pervasive developmental disorders, schizophrenia, or other psychotic disorders. The DSM-IV specifies three subtypes of ADHD, namely the predominantly inattentive subtype (ADHD-I, DSM-IV code 314.00), the predominantly hyperactive and impulsive subtype (ADHD-HI, DSM-IV code 314.01) and the combined subtype (ADHD-C, DSM-IV code 314.01).

Estimates of prevalence rates for ADHD vary between 1.9% and 14.4% (Scahill & Schwab-Stone, 2000), depending for example on the sample selection (clinical vs. community) or diagnostic criteria, as the ICD-10 criteria are more restrictive than the DSM-IV diagnosis. According to Biederman & Faraone (2005) worldwide prevalence rates range within 8% to 12%. Scahill & Schwab-Stone (2000) reviewed the prevalence of ADHD in school-age children based on 19 studies and concluded that the best estimate appears to be 5% to 10% . Epidemiologic studies estimate the prevalence of adult ADHD to be between 3% and 5%. Prevalence rates differ between genders. Boys are usually more affected than girls with a male/female ratio between 10:1 (clinic-referred) and 3:1 (community). Girls with ADHD might be under-identified because they are at less risk for comorbid disruptive behaviour disorder than boys with ADHD (for a review see Biederman, 2005).

1.2.2 Causal models of ADHD

Among others, two causal models have been influential in ADHD research: (1) executive dysfunction due to deficient inhibitory control, and (2) impaired signalling of delayed rewards arising from disturbances in motivational processes. Executive functions (EFs) are neurocognitive processes that maintain an appropriate problem solving set to attain a future goal and are controlled by frontal-subcortical circuits. EFs include inhibition, working memory, set-shifting, planning and sustained attention (Biederman & Faraone, 2005). Behavioural performance of ADHD children on tasks engaging EFs is often impaired compared to control children. ADHD children show difficulty in sustained attention in the Continuous Performance Test (CPT), in response inhibition measured by the Stop-Signal Task (SST) and chose the larger delayed reward less often than control children in the

Choice-Delay Task (C-DT), indicating delay aversion in ADHD children (Nichols & Waschbusch, 2004).

According to the causal model of Barkley (1997) the core symptoms of ADHD, namely hyperactivity and impulsivity, are due to an inefficient executive inhibition system. Barkley postulates that executive dysfunctions in ADHD are a consequence of a higher order deficit in response inhibition. Four other executive functions, (a) working memory, (b) internalisation of speech, (c) self-regulation of affect, motivation and arousal, and (d) reconstitution (analysis and synthesis of behaviour) depend on a functioning behavioural inhibition. Alternatively, a motivationally-based model focuses on altered reward processes, as there is evidence linking ADHD with hypersensitivity to delay and difficulties in waiting for desired outcomes. Delay aversion seems to be independent of inhibitory deficits (Sonuga-Barke, 2003; Sonuga-Barke, Dalen, & Remington, 2003). Recent data suggests that both models can be seen as complementary accounts of two psycho-patho-physiological subtypes of ADHD with different developmental pathways. At a neurobiological level, alterations within the executive circuit modulated by mesocortical dopamine and the reward circuit modulated by mesolimbic dopamine constitute more or less discrete neuropsychologic bases for executive/inhibitory deficits and delay aversion, respectively (Sonuga-Barke, 2003, 2005).

Although ADHD is associated with significant weaknesses in several EF domains, the lack of universality of EF deficits among children with ADHD suggests that EF weaknesses are neither necessary nor sufficient to cause all cases of ADHD (for an overview see Willcutt, Doyle, Nigg, Faraone, & Pennington, 2005). The current view is that ADHD is a complex, multifactoral disorder and that subtypes of ADHD might not have the same etiology (Nigg, 2005; Sonuga-Barke, 2005).

1.2.3 Selected aspects of etiology

There is consistent evidence that genetic factors contribute to the etiology of ADHD. Heritability is estimated 0.76 (Biederman & Faraone, 2005). Most family studies have identified a two- to eightfold increase in the risk for ADHD in parents and siblings of ADHD children (Biederman, 2005). Latest molecular-genetic findings of ADHD suggest that dopaminergic, serotonergic, and noradrenergic neurotransmission pathways account for the etiology of this complex disorder. Neuropharmacological studies further support the

hypothesis of catecholamine dysfunctions in ADHD. Stimulants (e.g. atomoxetin and methylphenidate), used for the treatment of ADHD, act on the norepinephrine (NE) and dopamine (DA) systems to enhance EFs, which are often disturbed in persons with ADHD (Pliszka, 2005). The most strongly implicated gene in ADHD is the 7-repeat allele of the *dopamine receptor D4 gene* (DRD4) (Albayrak, Friedel, Schimmelmann, Hinney, & Hebebrand, 2008; Faraone et al., 1999). A particular variant of DRD4, which differed ADHD children from control children, was found to mediate a blunted response to dopamine (LaHoste et al., 1996). Furthermore, the rationale to consider the *dopamine transporter* (DAT) involved in the pathogenesis of ADHD arises from findings, which showed that ADHD patients revealed a higher striatal DAT density (Krause, Dresel, Krause, la Fougere, & Ackenheil, 2003), resulting in lower dopamine levels by more dopamine reuptake through the transporters. About two thirds of ADHD children respond to medications (e.g. methylphenidate (MPH)) that inhibit the dopamine transporter (Greenhill, Findling, & Swanson, 2002). Blocking the transporter may increase the amount of extracellular dopamine in the neuronal synapse and possibly increases the inhibitory influence of frontal cortical activity on subcortical structures (Albayrak et al., 2008).

Neuroimaging studies found smaller volumes in frontal cortex, cerebellum, and subcortical structures in ADHD patients. The subcortical structures implicated by the imaging studies (caudate, puntamen, and globus pallidus) are part of the neural circuitry underlying motor control, executive functions, inhibition of behavior, and the reward pathways. Aberrances in these structures may therefore contribute to core deficits in ADHD (for an overview see Biederman, 2005). Support for a "lazy frontal lobe" hypothesis – stating that the prefrontal cortex is hypoactive in ADHD children – came from electrophysiological findings. Quantitative electroencephalography (qEEG) studies demonstrated *greater* levels of slow-wave (theta) activity and deficiencies of relative alpha and beta power in ADHD children compared to the control group. These findings, especially lower beta activity during cognitive tasks, were interpreted as cortical underarousal (Barry, Clarke, & Johnstone, 2003; Clarke, Barry, McCarthy, & Selikowitz, 1998; Clarke et al., 2003). It is unclear whether the observed group differences reflect delayed brain maturation or developmental deviation. Furthermore, although some findings do raise implications about possible underarousal in some ADHD subgroups, others argue for a different profile of cortical aberration in others (for an overview see Tannock, 1998). For example, theta band activity in ADHD children was found to be *increased* after MPH

application and correlated with improvement in D2 test (testing short term attention) (Wienbruch, Paul, Bauer, & Kivelitz, 2005).

In summary, recent findings indicate that prefrontal cortex functions, such as executive abilities (e.g. working memory, attention regulation, behavioral inhibition, planning, and organization) are weaker in patients with ADHD and there is strong evidence that genetic factors contribute to the onset of the disorder (Arnsten & Li, 2005; Biederman, 2005; Chamberlain, Robbins, & Sahakian, 2007; Gerlach, Deckert, Rothenberger, & Warnke, 2008). But, although neuropharmacological studies, as well as structural and functional imaging studies consistently suggest dysfunction in prefrontal cortex and imbalances in dopaminergic and noradrenergic systems, a pathophysiologic profile of ADHD has not been fully characterized (di Michele, Prichep, John, & Chabot, 2005).

1.2.4 Event-related potentials (ERPs)

Compared to controls, children and adolescents with ADHD generally perform more poorly on tasks that measure sustained and/or selective attention. These behavioral impairments were mirrored by the P3 wave (also known as P300), which is generated when subjects attend to and discriminate events. Several ERP studies demonstrated that ADHD children exhibit smaller P3 amplitudes and longer P3 latencies compared to controls (for an overview see Tannock, 1998). Furthermore, the N2 amplitude (reflecting inhibitory processes) has been found to be markedly reduced in ADHD children in response to Stop signals (Go/NoGo task) compared to control children and significantly correlated with response-inhibition performance (Pliszka, Liotti, & Woldorff, 2000). The error related negativity (ERN), a negative wave that is present selectively on error trials, was also markedly reduced in ADHD children, reflecting impaired error monitoring (error detection and correction) in ADHD children (Liotti, Pliszka, Perez, Kothmann, & Woldorff, 2005).

Overall, ERPs, which indicate deficits concerning attention, inhibition of behavior and error monitoring in ADHD children, have been reported. Additionally, ERP components, possibly indicating aberrations in auditory processing in ADHD children were found (see next section).

1.2.5 ADHD and auditory perception

Psychoacoustic studies demonstrated that ADHD is a significant factor in children´s performance on auditory tasks. Breier, Fletcher, Foorman, Klaas, & Gray (2003) assessed the perception of auditory temporal and nontemporal cues in children with (a) reading disability (RD)[1], (b) ADHD, (c) RD and ADHD, and (d) no impairments. RD was associated with impairment in detection of a tone onset time asynchrony, whereas the presence of ADHD resulted in a general reduction of performance across tasks (tasks assessing perception of auditory temporal and nontemporal cues). Further, children with ADHD showed poorer speech discrimination abilities than control children when background noise was introduced and dysfunctions in suprathreshold loudness perception (Geffner, Lucker, & Koch, 1996; Lucker, Geffner, & Koch, 1996).

Abnormalities in ERP components, such as brainstem evoked potentials and processing negativity (PN), have been reported, being more prevalent in auditory than in the visual modality. For example, PN during performance of a selective auditory task, being located in the auditory sensory cortex, was reduced in children with ADHD (Kemner et al., 2004). Moreover, children with attention deficits had prolonged latencies of brainstem auditory evoked potentials (Lahat et al., 1995), indicating impairments in early processing of auditory stimuli. Morphologic and structural differences in auditory brain areas, relative to controls, suggest a neuroanatomical basis for the frequently observed auditory performance deficits among ADHD children (for an overview see Chermak, Hall, & Musiek, 1999).

Scientists dispute over the hypothesis that children with ADHD have a genuine auditory processing deficit, because behavioural deficits of ADHD children (e.g. deficits in sustained attention, inhibition and working memory) could potentially affect performance on perceptual tasks (Barkley, 1997; Barkley, Grodzinsky, & DuPaul, 1992). For example, Sutcliffe and colleagues (2006) have reported deficits in frequency discrimination (FD) in children with ADHD compared to controls when off but not on stimulant medication. The authors concluded that poor performance on tasks that involve discrimination of brief auditory stimuli may be due to problems in temporal synchronization of attention, rather than genuine inability to hear differences. Non-medicated ADHD children also responded

[1] The rate of comorbidity of reading disability (RD) and ADHD is substantially greater than predicted by chance, with estimates ranging from 15% to 45% (Purvis & Tannock, 2000).

slower on the FD task, which suggests that impulsivity was not the explanation of poor performance. The other way round, it is also possible that listening difficulties lead to ADHD symptoms. When auditory processing of a particular sound source is impaired, children might be less able to focus their attention on that sound source. Therefore, central auditory performance deficits among children with ADHD may reflect the presence of a (central) auditory processing disorder (C)APD rather than the ADHD per se.

1.2.6 ADHD and (central) auditory processing disorder

Auditory processing disorder (ICD-10, code F80.20, Version 2007) describes a variable set of symptoms that have in common a difficulty listening to sounds in the absence of an audiometric deficit. Prevalence estimates range between 2-3% and 7%, with a 2:1 ratio between boys and girls (Bamiou, Musiek, & Luxon, 2001; Chermak et al., 1999; Nickisch et al., 2007). The American Speech-Language-Hearing Association (ASHA) defined (C)APD in a position statement of 2005 as impairments of auditory information in the central nervous system (CNS) as demonstrated by poor performance in one or more of the following skills: *sound localisation and lateralisation; auditory discrimination; auditory pattern recognition; temporal aspects of audition including temporal integration, temporal discrimination (e.g. temporal gap detection), temporal ordering, and temporal masking; auditory performance with competing acoustic signals; and auditory performance with degraded acoustic signals* (Nickisch et al., 2007). Higher order cognitive-communicative and/or language related functions have been excluded from that definition (Brunner, 2007). However, there is still scientific debate over what symptoms, tests, stimuli and mechanisms are appropriate to define and diagnose APD, as it is a complex and heterogeneous disorder that may occur or be causal to other deficits (Moore, Halliday, & Amitay, 2009).

Recent research suggests that ADHD and (C)APD are often co-morbid, but can occur independently (Keller & Tillery, 2002; Tillery, Katz, & Keller, 2000). Tillery et al. (2000) investigated the effect of MPH on Central Auditory Processing (CAP) tests and Auditory Continuous Performance Test (ACPT) in children diagnosed with both ADHD and (C)APD. Findings revealed a significant improvement of the attention/impulsivity (ACPT) test performance under medication, but no medication effect on the CAP measures (e.g. assessing phonemic skills or the ability to extract speech from background noise). The authors concluded that (a) MPH had a positive effect on sustained attention and

impulsivity, but did not eliminate auditory dysfunctions and that (b) ADHD and (C)APD are independent problems.

The similar symptomatology of children with ADHD, especially those of the predominantly inattentive subtype, and (C)APD render differential diagnosis especially challenging. Common symptoms of both disorders are academic difficulties, distraction, inattention, listening deficits, asking for things to be repeated, auditory divided attention deficit and difficulty hearing in background noise (Chermak et al., 1999; Chermak, Somers, & Seikel, 1998; Chermak, Tucker, & Seikel, 2002). Identifying an objective measure for modality-specific perceptual dysfunctions would help to refine diagnostic criteria for ADHD and (C)APD (Cacace & McFarland, 2005). Late AEPs, "auditory evoked neutral condition event-related potentials" (NC-ERP), were thought to be a possible tool to unveil dysfunctions of auditory processing. In a retrospective analysis of NC-ERP data from children suffering from (C)APD alone and (C)APD-children with ADHD the NC-ERPs did not differentiate between (C)APD and ADHD (Ptok, Blachnik, & Schonweiler, 2004). The authors concluded that the two disorders may indeed be overlapping entities.

In summary, a heterogeneous population of children may be subsumed under the denominations of ADHD-I, ADHD-C and (C)APD. This could be due to mistaken diagnosis, comorbidity, or to a causal role of (C)APD in ADHD (Barkley, 2006; Cacace & McFarland, 2005).

1.3 The present project

In comparison with auditory streaming, less research has focussed on concurrent sound segregation. How does the brain manage to assign *simultaneously* incoming acoustic elements to perceptual auditory objects corresponding to different physical sound sources? Which factors contribute to concurrent sound segregation? How does focused attention influence concurrent sound segregation? These questions have concerned researchers for decades and auditory scientists have identified cues, which listeners can use when performing ASA, but still the mechanisms are not fully understood. The present thesis examined neural processes underlying concurrent sound segregation in adults and children with and without attention deficit hyperactivity disorder (ADHD) by measuring auditory evoked fields (AEFs) using magnetoencephalography (MEG). The aim of the thesis was to learn more about general markers of concurrent sound segregation and the bottom-up and

top-down processing stages of ASA. Further, it aimed to investigate the maturational changes underlying concurrent sound segregation and finally tried to identify markers of concurrent sound segregation that may indicate a genuine auditory processing deficit in children with ADHD. Two distinct cues — inharmonicity and onset asynchrony — were used to invoke sound segregation (Darwin & Carlyon, 1995). More precisely, participants were presented with complex sounds with the third harmonic mistuned by 8% of its original value and the onset of the third harmonic delayed by 160 ms compared to the other harmonics. Theses manipulations of the harmonic complex could promote sound segregation leading to the perception of two rather than one sound: a "buzz" sound along with another sound corresponding to the mistuned/delayed third harmonic, which "pops out" of the harmonic complex. Complex sounds with all harmonics in-tune served as *Baseline* stimuli. Active and passive listening conditions were chosen to evaluate the contribution of attention on sound segregation.

Most research has focused on event related potentials (ERPs) and AEFs, which were generated by sounds that evoked segregation based on inharmonicity (Alain et al., 2001; Alain & McDonald, 2007; Alain et al., 2003). **Study 1** extends prior research by adding an alternative cue to inharmonicity, namely onset asynchrony to measure brain activation, which indexes a more *general* process of sound segregation in adults. Additionally, the influence of focused attention on that activation was investigated in MEG and a sound localization task was introduced, which could only be performed correctly if participants truly heard out the segregated partial. The invention of this task ensured that participants based their judgment on sound segregation rather than other cues. We expected that if AEF components are sensitive to attention to the sounds, they should differ between the active (*Attend)* and passive (*Non-Attend)* listening condition. A further hypothesis was that only AEF components which are elicited by both, mistuned *and* delayed sounds are general markers of sound segregation.

Sound segregation is crucial for children in many listening situations like for example in a typical classroom situation, where children need to parse the acoustic elements that originate from the teacher and distinguish them from other simultaneously present sounds (e.g., other pupils talking). Similarly to adults, most research has investigated the ability to process sequentially presented sounds in children. Thus, it could be demonstrated that basic mechanisms of auditory stream segregation seem to be present already in newborns and

infants (Demany, 1982; McAdams & Bertoncini, 1997; Winkler et al., 2003). However, many studies documented that children differ in auditory perceptual abilities from adults and that the processing of various auditory stimuli has a maturational time course that extends into adolescence (for a review see Wunderlich & Cone-Wesson, 2006). The motivation for **Study 2** was to extend prior research with new findings about the auditory perceptual abilities of children with a focus on developmental changes in concurrent sound segregation. The experimental design of Study 1 was applied to children to benefit from its above mentioned advantages (localisation task, attention manipulation and two distinct cues evoking sound segregation). It was expected that concurrent sound segregation in children may be not fully developed and that they accordingly should perform poorer than adults in the sound localization task. Further, AEF components, which potentially reflect an automatic concurrent sound segregation process in adults, should be elicited less robustly in children, if bottom-up mechanisms of sound segregation are immature.

Study 3 concerned children with attention deficit hyperactivity disorder (ADHD). Psychological theories of ADHD, as well as present psychological research evolved the view that ADHD is a complex, multifactoral disorder (Nigg, 2005; Sonuga-Barke, 2005) probably subsuming an etiologically heterogeneous population of children. Why is the investigation of neural processes involved in concurrent sound segregation based on primitive grouping mechanisms in children with ADHD of special interest? Because recent research indicates that children with ADHD perform poor on psychoacoustic tasks and present strikingly similar symptoms with children suffering from (C)APD. One possibility is that listening difficulties lead to ADHD symptoms, as children might be less able to focus their attention on a particular sound source when other sounds sources are simultaneously active. However, debate continues over the hypothesis that children with ADHD have a genuine auditory processing deficit. For a better understanding of the neural underpinning of auditory processing dysfunctions in ADHD children we investigated the ability of these children to segregate auditory objects in Study 3 using the same experimental paradigm as in the previous studies. If children with ADHD have a genuine auditory processing deficit concerning concurrent sound segregation, they should differ from control children in AEF components, which potentially reflect an automatic, bottom-up concurrent sound segregation process. For accurate differential diagnosis between ADHD and (C)APD it is required to measure modality-specific perceptual dysfunctions

(Cacace & McFarland, 1998). Study 3 of the present thesis aimed to identify a neurological marker that could help to refine diagnostic criteria.

To put it in a nutshell, the scope of the present thesis was to contribute to existent knowledge about concurrent sound segregation in general by methodical ameliorations (Study 1), document children`s abilities of concurrent sound segregation compared to adults (Study 2) and finally shed light on the debate if children with ADHD have a genuine auditory processing deficit (Study 3). The following sections comprise the detailed descriptions of the particular studies.

2 Studies

2.1 Concurrent sound segregation based on inharmonicity and onset asynchrony (Study 1)[2]

Abstract

To explore the neural processes underlying concurrent sound segregation, auditory evoked fields (AEFs) were measured using magnetoencephalography (MEG). To induce the segregation of two auditory objects we manipulated harmonicity and onset synchrony. Participants were presented with complex sounds with (i) all harmonics in-tune (ii) the third harmonic mistuned by 8% of its original value (iii) the onset of the third harmonic delayed by 160 ms compared to the other harmonics. During recording, participants listened to the sounds and performed an auditory localisation task whereas in another session they ignored the sounds and performed a visual localisation task. Active and passive listening was chosen to evaluate the contribution of attention on sound segregation. Both cues - inharmonicity and onset asynchrony - elicited sound segregation, as participants were more likely to report correctly on which side they heard the third harmonic when it was mistuned or delayed compared to being in-tune with all other harmonics. AEF activity associated with concurrent sound segregation was identified over both temporal lobes. We found an early deflection at ~75 ms (P75m) after sound onset, probably reflecting an automatic registration of the *mistuned* harmonic. Subsequent deflections, the object-related negativity (ORNm) and a later displacement (P230m) seem to be more general markers of concurrent sound segregation, as they were elicited by both mistuning *and* delaying the third harmonic. Results indicate that the ORNm reflects relatively automatic, bottom-up sound segregation processes, whereas the P230m is more sensitive to attention, especially with inharmonicity as the cue for concurrent sound segregation.

2.1.1 Introduction (Study 1)

Many social interactions occur in noisy environments posing a high demand on the auditory system: voices of different speakers and sounds originating from other sources (e.g. radio,

[2] *Published*: Lipp, R., Kitterick, P., Summerfield, Q., Bailey, P. J., & Paul-Jordanov, I. (2010). Concurrent sound segregation based on inharmonicity and onset asynchrony. *Neuropsychologia, 48*(5), 1417-1425.

television, traffic) need to be distinguished. This is a complex challenge, since the sound that reaches the ear is a summation of the pressure waves from the individual sources. The ability to assign incoming acoustic elements to perceptual auditory objects corresponding to different physical sound sources is known as auditory scene analysis. Features from the same source are grouped together while parsing the incoming acoustic stream (Bregman, 1990). In typical complex auditory scenes there are multiple cues that can be used to segregate the acoustic input. For example, segregation of concurrent speech signals improves with increasing difference between the fundamental frequencies of the signals (Alain, Reinke, He, Wang, & Lobaugh, 2005; Chalikia & Bregman, 1989; Culling & Darwin, 1993). Furthermore, sounds with different onset times, intensities and timbres are more likely to be considered as originating from different sound sources than sounds with the same onset times, intensities and timbres (Bregman, 1990). Bregman (1990) distinguishes two classes of mechanisms that help to decide which components belong to a particular sound source: "primitive grouping mechanisms" and "schema-governed mechanisms". Whereas schema-governed mechanisms are thought to be learned and dependent on the listener's specific auditory experience, primitive grouping mechanisms do not depend on experience and use stimulus properties dependent on the physical properties of the sound sources — such as frequency, intensity, phase, harmonicity, and temporal coherence — to segregate the incoming signal.

Alain and colleagues investigated sound segregation in a series of experiments measuring the electroencephalogram (EEG) of adults (Alain, Arnott, & Picton, 2001; Alain & Izenberg, 2003; Alain, Schuler, & McDonald, 2002) while they were presented with auditory stimuli. The stimuli either promoted sound segregation leading to the perception of two rather than one sound (Segregation condition), or did not promote sound segregation leading to the perception of one sound (No Segregation condition). The presented sounds comprised multiple harmonics of 200 Hz, one of which could be mistuned so that it was no longer an integer multiple of the fundamental. In the Segregation condition the third harmonic was mistuned by 1-16% (Alain et al., 2001, 2002) causing the third harmonic to stand out perceptually from the sound complex formed by the other harmonics. Increasing the amount of mistuning increased the likelihood that participants would report the perception of two auditory objects. The authors identified a negative deflection in the difference waveform (Segregation – No Segregation) approximately 180 ms after stimulus onset, referred to as the object-related negativity (ORN) and another later, positive

deflection around 400 ms, labelled P400. The ORN was present when participants actively listened to the sounds and indicated via key-press if they heard one or two sounds and when their attention was distracted by reading a book (Alain et al., 2001) or watching a silent movie (Alain et al., 2002). The authors concluded that the ORN reflects automatic, bottom-up sound segregation and that its generation is minimally affected by attentional load (Alain & Izenberg, 2003). In contrast, the P400 was only present in the active listening condition. The authors therefore concluded that the P400 reflects awareness-related top-down processes of sound segregation.

Alain and McDonald (2007) also measured auditory evoked fields (AEFs) using magnetoencephalography while participants were presented with complex sounds – again harmonics were either in-tune or the third partial was mistuned. Participants were not required to pay attention or respond to the stimuli. Regional dipole sources were fitted on the N100m peak of the resulting auditory evoked fields and source waveforms were analysed between conditions. In young adults the source difference waveform (Mistuned (16%) – Tuned) generated an early positivity P80 (~80 ms), an ORN (~160 ms) and a positive displacement P230 (~230 ms) after stimulus onset. The ORN and P230 amplitude correlated with the perceptual judgement of hearing one or two sounds (measured separately after the MEG experiment). The early positivity was not correlated with perceptual judgement and was interpreted as reflecting an early registration of inharmonicity in primary auditory cortex. The authors concluded that all three components likely reflect automatic sound segregation that may occur independently of listeners´ attention. However, attention was not manipulated making it difficult to determine whether the early P80, ORN, and P230 are modulated by attention.

The AEFs (P80, ORN and P230) and ERPs (ORN and P400) were generated by sounds that evoked segregation based on inharmonicity (Alain et al., 2001; Alain & McDonald, 2007; Alain, Theunissen, Chevalier, Batty, & Taylor, 2003). However, the question remains, whether these AEFs/ERPs really reflect *segregation*. It could also be the case that they reflect the detection of an over-all change in the frequency pattern of the stimuli in the *Segregation* condition rather than sound segregation itself, since the frequency of the third harmonic was raised in the *Segregation* condition compared to the *No Segregation* condition. Another caveat concerns the task participants performed. For Alain and colleagues (Alain et al., 2001, 2003; Alain & Izenberg, 2003; McDonald &

Alain, 2005) the index of perceptual experience was whether participants reported hearing one sound or two sounds. This measure is problematic, for example because of the possibility that it reflects not the presence or absence of perceptual segregation *per se*, but a tendency for participants to use the "one sound" and "two sound" responses as labels for the different sounds in the *No Segregation* and *Segregation* conditions.

The present study tried to address these problems as follows. To investigate neural processes involved in concurrent sound segregation based on primitive grouping mechanisms, we used two distinct cues — inharmonicity and onset asynchrony — to invoke sound segregation (Darwin & Carlyon, 1995). For this purpose, we included a second type of stimulus that evokes segregation without changing the harmonicity of the stimulus by delaying the onset of one harmonic relative to the other harmonics. If the same components (ORN and P230/P400) occur using this kind of stimulus, it is safer to assume that ORN and P230/P400 really do reflect sound segregation. To ensure that participants based their judgment on sound segregation rather than other cues, we used a sound localisation task that could only be performed correctly if participants truly heard out the segregated partial, which was presented either left or right from straight ahead (*Attend* condition). In order to investigate the susceptibility of the ORN and P230/P400 to attentional focus, an additional task was included that distracted the participants' attention away from the sounds, but kept the over-all attentional load comparable (*Non-Attend* condition).

The present MEG-study aimed to investigate brain activation which indexes a general process of sound segregation, and if there are aspects of that activation that are affected by attention. We hypothesized that: (1) only AEF components which are generated by both, mistuned *and* delayed sounds are general markers of sound segregation, and (2) if an AEF component depends on attention to the sounds, it should differ between the *Attend* and *Non-Attend* condition. Lastly, the present study extends prior research by examining whether the AEFs can generalize to concurrent sound segregation and whether they are sensitive to attention. This was realized by a new variant of the mistuned harmonic paradigm that allows for a more objective assessment of concurrent sound segregation. We introduced a sound localisation task, added an alternative cue (onset asynchrony) and measured the influence of attention on concurrent sound segregation in MEG.

2.1.2 Methods (Study 1)

Participants

Fifteen adults (mean age 24.2 years, range 20-28 years, 7 female, 1 left-handed) and one adolescent (15 years, right-handed) took part in the experiment. Participants were contacted through advertisements at the University of Konstanz. All participants had pure-tone audiometric thresholds in the normal range in both ears for frequencies 500 Hz, 1 kHz, 2 kHz and 4 kHz.

Procedure

In session one, hearing thresholds were determined and participants were familiarised with the experiment. Two attention tests, one testing short-term attention, the other one long-term attention (D2 and KLT-R) were administered (Brickenkamp, 2002; Dünker, Lienert, Lukesch, & Mayrhofer, 2001). In session two, MEG data were collected in two conditions: (a) participants listened to sounds and performed an auditory localisation task (*Attend* condition), and (b) participants were played the same sounds while performing a visual localisation task (*Non-attend* condition). *Attend* and *Non-Attend* conditions were presented in a counterbalanced order. Ethical approval of the study was granted by the Ethical Review Board of the University of Konstanz.

Acoustic stimuli

The stimuli were created at the Department of Psychology of the University of York, UK (Summerfield, Paul, Kitterick, Briley, & Bailey, 2007). *Baseline* stimuli consisted of 24 sinusoidal partials which were synthesised digitally (44,100 samples/s, 16-bit amplitude quantization) with equal amplitudes. The frequencies of the sinusoids were integer multiples of 200 Hz, from 200 Hz to 4800 Hz inclusive. Odd-numbered partials started in sine phase and even-numbered partials started in cosine phase, giving a composite waveform with a low peak factor. The duration of each partial was 500 ms, including 10-ms ramps shaped by raised cosines to smooth the onset and offset. Two variants of the *Baseline* stimulus were formed by perturbing the third partial at 600 Hz. In *Mistuned* stimuli, the frequency of the partial was raised by 8% to 648 Hz. In *Delayed* stimuli, the partial started 160 ms after the other partials, but ended at the same time as them (Fig. 1).

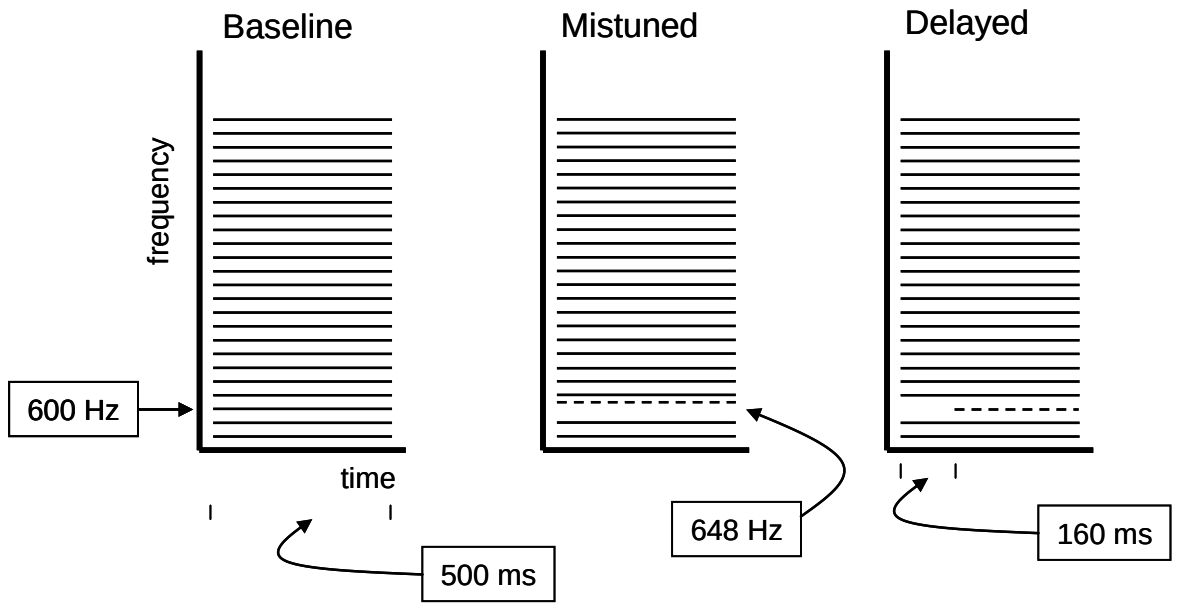

Figure 1: Schematic spectrograms of Baseline, Mistuned and Delayed stimuli.

The 24 partials were presented simultaneously, each through one member of a set of 24 Bose Acoustimas Cube Loudspeakers mounted on 135-cm poles spaced at 15-degree intervals around the edge of a circular stage. The stage had a diameter of 3.3 m and was located in a 6 m × 4 m × 2 m sound-treated room (Industrial Acoustics Company). One hundred exemplars of each type of stimulus were created by randomly reassigning the partials to the loudspeakers, with the exception of the third partial which always came from either 45° to the left or 45° to the right of straight ahead, with equal probability. Additionally, 22 *Baseline*, 22 *Mistuned* and 20 *Delayed* stimuli were created in which the third partial was 10 dB more intense than the other partials. The purpose of the latter stimuli was to provide occasional reminders of the sound of the third partial.

A Brüel & Kjær (B&K) Head and Torso Simulator (HATS, Type 4100D) was placed in the centre of the ring of loudspeakers with its ear canals at the same height as the loudspeakers. A B&K 0.5-in. microphone (Type 4189-A-002) was located in each of the ear canals of the HATS. The 364 stimuli were presented while the outputs of the microphones were recorded digitally (44,100 samples/s, 16-bit amplitude quantisation). The binaural recording of each stimulus was edited to a duration of 540 ms to include low-amplitude echoes resulting from the room acoustics.

The resulting stimuli were presented through headphones (Dynamic Stereo Headphones MDR-CD470, Sony) for familiarising the participants with the experiment in session one. For the MEG recordings the stimuli were presented through tube-phones (Etymotic Research, ER30). For each ear, a 6.1-m length of 4-mm diameter tube terminated in a right-angle adapter (ER30-9) ending with a disposable 13-mm foam eartip (ER13-14). A digital filter was designed to compensate for the low-pass frequency response of this sound delivery system, resulting in a frequency response that was flat to within -4–0 dB from 100–4800 Hz when measured with a B&K Ear Simulator (Type 4157) with 0.5-inch microphone (Type 4134), Microphone Preamplifier (ZC 0032) and B&K Hand-held Analyzer (Type 2250). Using the same measurement system, the presentation level of *Baseline* stimuli was set to 70 dB (A).

Training and testing with headphone presentation

The aim of this phase of the experiment was to give participants experience of the task that they would be asked to perform later during MEG imaging. Participants sat in front of a 19-inch visual display unit displaying a fixation cross and practiced in two conditions. In the *Attend* condition, a sequence of 364 stimuli was presented through headphones (Dynamic Stereo Headphones MDR-CD470, Sony) and a Digital Audio Interface (Lynx L22, 192 kHz PCI). The sequence contained 100 instances of each type of stimulus (*Baseline, Mistuned, Delayed*) plus 64 reminder stimuli in which the third partial was 10 dB more intense than the other partials.

On each trial of the *Attend* condition, participants reported whether they heard the third partial on the left or right using two keys on a keyboard. Feedback on the accuracy of the response was given by displaying a coloured rectangle on the side where the key had been pressed. It was green if the response was correct or red if the response was incorrect. The inter-stimulus interval (ISI) was the response latency plus 500 ms.

In the *Non-attend* condition 20 stimuli were presented to make the participant familiar with the visual localisation task. On each trial of the *Non-attend* condition, a faint grey circle subtending a visual angle of 1° appeared for 50 ms at a randomly chosen location within 8° right or left of the fixation cross. Participants were instructed to report whether the circle appeared on the left or the right by making a key-press. Feedback was provided in the same format as used in the *Attend* condition. The ISI was the response

latency plus a random value in the range from 500 ms to 1500 ms. At the same time, the acoustic stimuli were presented asynchronously with an ISI that was chosen randomly from the range 500–1500 ms. Participants were instructed to concentrate on the visual stimuli and to ignore acoustic stimuli.

MEG recordings

Recording was done with a 148-channel magnetometer (MAGNESTM 2500 WH, 4D Neuroimaging, San Diego, USA). Participants practised the *Attend* and *Non-Attend* condition. In each condition, the 364 stimuli as described above were presented. Acoustic stimuli were delivered with an analog-to-digital converter (Motu 2408) and amplifier (Servo 200, Samson) through tube-phones (Etymotic Research, ER30). The foam tips were introduced carefully into the ear canals of participants and were additionally secured with medical tape. Participants lay on a height-adjustable bed in a magnetically shielded room (Vakuumschmelze Hanau). For artefact control, eye movements (EOG) were recorded from four electrodes attached to the left and right temple and above and below the right eye, as well as cardiac activity (ECG) via two electrodes, one on each forearm. A Neurofax amplifier (EEG-11006, Nihon Kohden) served for the recording of EOG and ECG. A video camera installed inside the chamber allowed monitoring the subject's behaviour and compliance throughout the experiment. Visual stimuli and feedback, as used during the practice phase, were projected (D-ILA Projector, JVC™, DLA-G11E) onto a screen placed ~40cm away from the participants' eyes through a mirror system. Data from each channel were band-pass filtered between 0.1 Hz and 200 Hz and sampled at a rate of 678.17 Hz. Recording was continuous.

MEG data analysis

Data were corrected for ocular and cardiac artefacts using BESA™ 5.2.4 and were averaged separately for *Baseline*, *Delayed*, and *Mistuned* stimuli in *Attend* and *Non-attend* conditions using a pre-stimulus baseline of 100 ms and a post-stimulus time-window of 800 ms. Epochs were excluded if the signal amplitude in any channel exceeded 3.5 pT or if the signal amplitude differed by more than 2.5 pT between adjacent sampling points. Reminder stimuli were excluded from analyses. The minimum number of averaged epochs was 76 in the *Attend* condition and 65 in the *Non-attend* condition. The average number of averaged epochs was 92.9.

Sensor groups that showed significant differences between conditions were identified objectively. First, a t-test was computed at each time point across the averaged epoch between two stimulus-types of interest (e.g. *Baseline* vs. *Mistuned*) for each sensor. Then an algorithm identified groups of a minimum of 10 adjacent sensors that showed significant differences between conditions ($p<.05$) for at least 20 ms. In this way, two groups of 6 sensors, located bilaterally over the temporal lobes, were identified which showed consistent differences between conditions (Fig. 2). The left channel group covered the in-going magnetic field (positive values) and the right channel group the out-going magnetic field (negative values) of the underlying auditory sources. Therefore, we reversed the sign of the signals over the left hemisphere to facilitate comparisons with right-hemisphere signals.

We computed mean amplitudes in time-windows of interest for each sensor group, participant, stimulus, and condition. Time-windows were chosen to embrace prominent deflections from zero in either direction in difference waveforms (*Mistuned minus Baseline* and *Delayed minus Baseline*). Mean amplitudes were analysed statistically with ANOVAs. Condition *(Baseline, Mistuned, Delayed),* Hemisphere (*Left, Right*), and Attention (*Attend, Non-attend*) were within-group variables. Additionally, paired t-tests and effect sizes (r) were computed for all relevant contrasts using the following formulae:[3]

$$\text{Cohen's } d = M_1 - M_2 \,/\, \sigma_{pooled}$$

$$\text{where } \sigma_{pooled} = \sqrt{[(\sigma_1^2 + \sigma_2^2) / 2]}$$

$$r_{Y\lambda} = d / \sqrt{(d^2 + 4)}$$

The differences of the mean amplitudes (*Mistuned–Baseline, Delayed–Baseline*) characterising the components (P75m, ORNm, P230m) were analysed statistically with ANOVAs.

To specify *individual* peak latencies, individual difference waveforms (*Mistuned – Baseline* and *Delayed – Baseline*) for both conditions (*Attend, Non-attend*) were band-pass filtered from 1 Hz to 25 Hz for detecting P75m and ORNm, and from 1 Hz to 8 Hz for detecting P230m. In order to improve signal-to-noise ratios in individual subjects, left- and

[3] M_1, $M_{2\ =\ means}$; σ_1,$\sigma_{2\ =\ standard\ deviations}$

right-hemispheric channels were averaged together. Peak latencies were compared between conditions with paired t-tests.

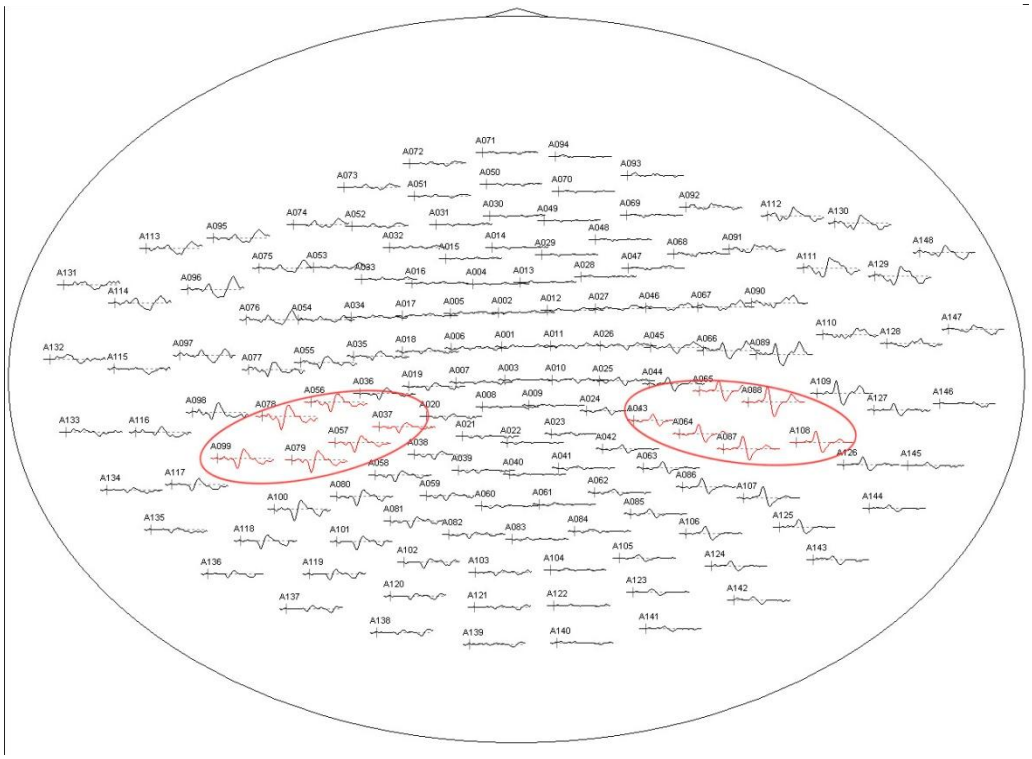

Figure 2: Difference waveforms (Delayed – Baseline) for each sensor averaged over 16 participants. Channel groups used for data analysis are circled (6 channels for each hemisphere).

Behavioural data

The percentages of correct responses to *Baseline*, *Mistuned,* and *Delayed* stimuli were compared in a multivariate analysis of variance (MANOVA). Condition (*Baseline*, *Mistuned*, *Delayed*) and Session (headphone presentation, MEG) were within-group variables. Further, Spearman`s rank correlations were calculated between behavioural results in MEG, attention test scores and amplitudes/latencies of MEG components (P75m, ORNm, P230m). P-Values were corrected for multiple tests using the Bonferroni-Holm method.

All analyses were conducted with Statistica (StatSoft, Inc., Version 6, 2003). *Post-hoc* analyses were performed with Tukey's HSD test.

2.1.3 Results (Study 1)

Hearing thresholds

All participants had normal hearing levels (<20 dB HL) for both ears for all tested frequencies. Mean hearing levels were lower than 5 dB HL in both ears.

Attention test scores

Table 1 shows the results of short- and long-term attention testing. Results indicate that participants performed slightly above average on both tests (percentile ranks > 60).

Table 1: Attention test scores of the long (KLT-R) and short term (D2) attention test.

Attention test	Value	Mean (percentile rank)	Lower Boundary (percentile rank)	Upper Boundary (percentile rank)
Short term *(D2)*	Overall performance *(GZ-F)*	63.4	13	99
	Concentration performance *(KL)*	64.4	24	99
Long-term *(KLT-R)*	Overall performance *(GL)*	60.6	4	100
	%Errors *(F%)*	43.5	4	93

Behavioural results

Participants performed well above chance localising the target stimulus partial, indicating that they segregated it successfully from the harmonic background. Overall, participants made more errors in the MEG condition than during headphone presentation ($F(1,15)=36.8$, $p=.00002$, see Fig. 3). Further, a significant main effect was found for Condition (*Baseline, Mistuned, Delayed*) ($F(2,30)=54.44$, $p<.001$), indicating that localisation performance was better in Delayed and Mistuned conditions compared to the Baseline condition irrespective of the session (*Post-hoc* testing, Headphone presentation *Delayed>Baseline* $p=.0001$, *Mistuned>Baseline* $p=.0002$; MEG: *Delayed>Baseline* $p=.0001$, *Mistuned>Baseline* $p=.0001$). A significant interaction Session × Condition was revealed ($F(2,30)=21.38$,

p<.001). Post-hoc testing showed that localisation performance in MEG was lower than in headphone presentation for *Delayed* (p<.001) and *Baseline* (p<.001) stimuli.

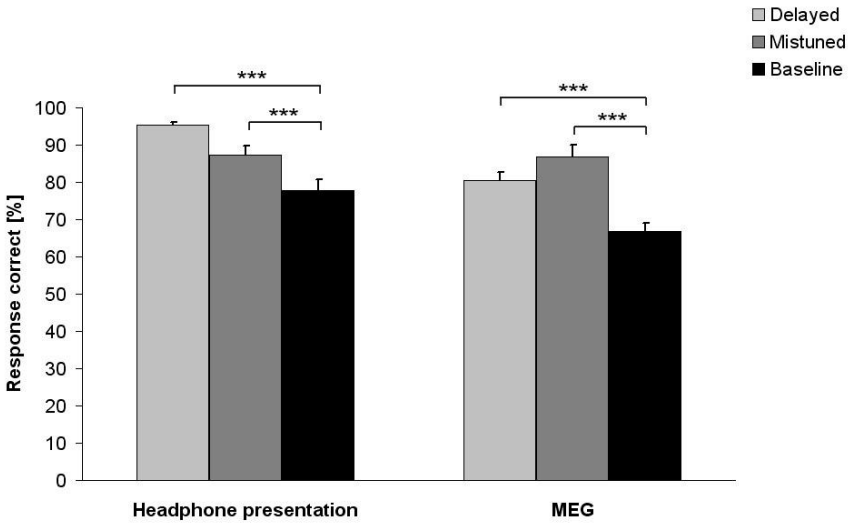

Figure 3: Localisation performance during the MEG scan and during headphone presentation. Error bars indicate standard errors.

Cortical response elicited by Baseline stimuli

Grand average waveforms for *Baseline* stimuli were characterised by an N100m with a mean peak latency of 86 ms and mean peak amplitude of -170 fT followed by a later sustained field, ca. 300–400 ms after stimulus onset (see Fig. 4). The N100m mean latencies for *Mistuned* and *Delayed* stimuli were 92 ms and 86 ms.

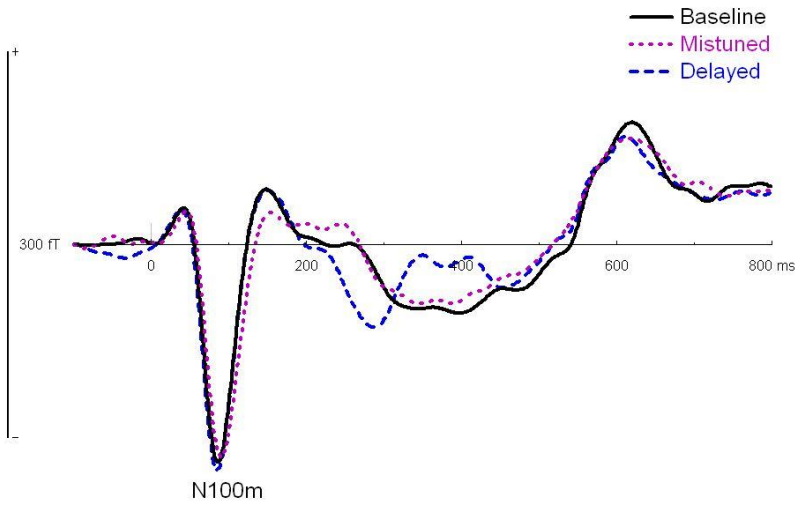

Figure 4: Waveform averaged over all participants (16) in the right-hemisphere (6 channels), evoked by Baseline, Mistuned and Delayed stimuli in the Attend condition; band-pass filtered from 1 Hz to 25 Hz.

Mistuning the third harmonic

Figure 5 shows the difference waveforms *Mistuned* minus *Baseline*. The difference waveforms were characterised by distinct components peaking at ~74 ms (P75m) and ~132 ms (ORNm) (also see Table 2). This likely corresponds to the early positivity (~80 ms), and the object-related negativity (ORN, ~160 ms) originally identified by Alain and McDonald (2007). The ORNm was followed by a broader component that peaked at ~256 ms (Fig. 5 and Table 2). A similar component was also found by Alain and colleagues, which was labelled P400 (Alain et al., 2001) or P230 (Alain & McDonald, 2007). Here is it referred to as the P230m.

Table 2: Peak latencies for components P75m, ORNm and P230m in attended and non-attended conditions for left- and right-hemispheric channel groups averaged over all participants. Time-ranges used for calculating mean amplitudes and further statistical analyses are displayed.

Component	Comparison	Attention	Hemisphere	Peak Latency (ms)	Time-Range (ms)	Mean Peak Latency (ms)
ORNm	*Mistuned* vs. *Baseline*	Attend	Left	126	106-146	132
			Right	118	98-138	
		Non-attend	Left	152	132-172	
			Right	130	110-150	
	Delayed vs. *Baseline*	Attend	Left	283	263-303	277
			Right	275	255-295	
		Non-attend	Left	277	257-297	
			Right	273	253-293	
P230m	*Mistuned* vs. *Baseline*	Attend	Left	251	216-286	256
			Right	239	204-274	
		Non-attend	Left	306	271-341	
			Right	229	194-264	
	Delayed vs. *Baseline*	Attend	Left	376	341-411	380
			Right	393	358-428	
		Non-attend	Left	381	346-416	
			Right	370	335-405	
P75m	*Mistuned* vs. *Baseline*	Attend	Left	75	60-90	74
			Right	71	56-86	
		Non-attend	Left	75	60-90	
			Right	73	58-88	

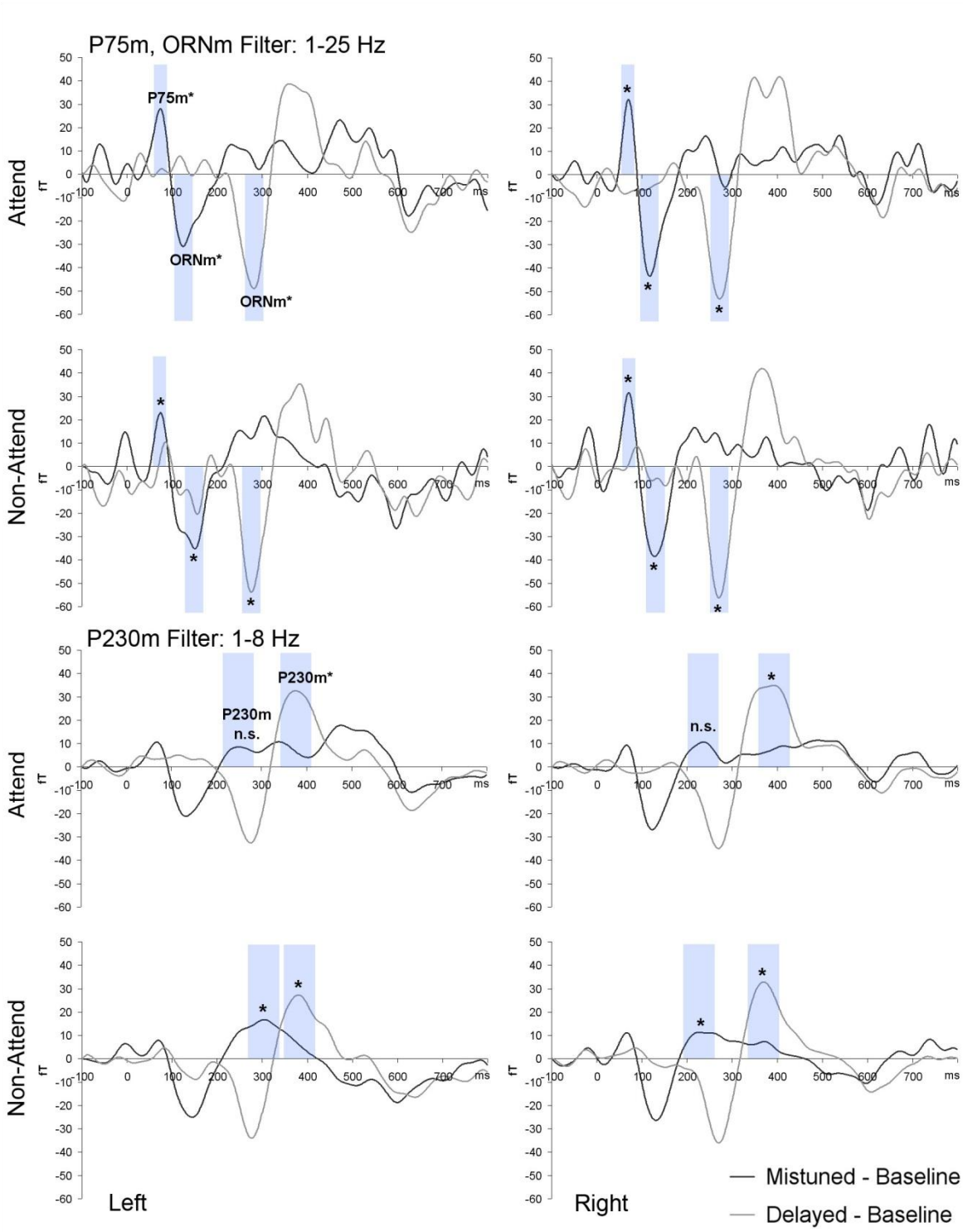

Figure 5: Difference waveforms for components P75m, ORNm and P230m in attended and non-attended conditions for left- and right-hemispheric channel groups averaged over all participants. The polarity of the signal is reversed in the left hemisphere. Grey boxes outline time-windows used for statistical analyses. P75m and ORNm are displayed in the upper four graphs (signals were band-pass filtered from 1 Hz to 25 Hz). P230m waveforms are displayed in the lower four graphs (band-pass filter between 1Hz and 8 Hz).

In the P75m time-ranges, mean amplitudes in the two temporal channel groups were computed for attended and non-attended *Baseline* and *Mistuned* conditions. P75m time-ranges covered ±15 ms around P75m peak latency. Mean amplitudes in the ORNm time-ranges covered ±20 ms around ORNm-peak latency, and for P230m time-ranges ±35 ms around P230m peak latency (Table 2). Mean amplitudes were retrieved per participant. Time-windows were set to be shorter for the P75m and ORNm responses than the P230m, to reflect their inherent durations; the P75m and ORNm responses also involved higher frequencies than the P230m, motivating a choice of band-pass filter from 1 Hz to 25 Hz for P75m and ORNm and 1 Hz and 8 Hz for the P230m.

The P75m was significant, i.e. *Baseline* stimuli led to higher amplitudes than *Mistuned* stimuli. Likewise, the ORNm was significant, i.e. amplitudes in *Mistuned* conditions were higher than in *Baseline* conditions for both hemispheres and both attention conditions (*Attend* and *Non-attend*). The P230m was significant, i.e. amplitudes in *Baseline* conditions were higher than in *Mistuned* conditions for both hemispheres but only for the *Non-attend* condition (Table 4). Main effects for condition (irrespective of attention) are summarised in Table 3.

Table 3: Main Effect of Sound Condition. F- and p-values of ANOVAs.

Component	Main Effect Sound Condition	F(1,15)	p
ORNm	*Mistuned* vs. *Baseline*	81.15	<.001
	Delayed vs. *Baseline*	63.86	<.001
P230m	*Mistuned* vs. *Baseline*	15.79	<.01
	Delayed vs. *Baseline*	36.92	<.001
P75m	*Mistuned* vs. *Baseline*	53.33	<.001

Table 4: Paired t-test results of the comparison between conditions (Baseline, Mistuned and Delayed) for each component (P75m, ORNm, P230m) displayed along with mean amplitudes and standard errors of the difference waveforms (Mistuned–Baseline and Delayed–Baseline).

Component	Comparison	Attention	Hemisphere	t(15)	P	Effect Size r	Mean Amplitude (fT)	Standard Error (fT)
ORNm	*Mistuned* vs. *Baseline*	Attend	Left	-5.3	<.001	-0.20	-26.31	4.96
			Right	-6.0	<.001	-0.25	-36.92	6.20
		Non-attend	Left	-5.4	<.001	-0.20	-30.97	5.63
			Right	-7.2	<.001	-0.26	-35.35	4.92
	Delayed vs. *Baseline*	Attend	Left	-6.7	<.001	-0.57	-44.07	6.61
			Right	-6.8	<.001	-0.61	-48.10	7.07
		Non-attend	Left	-5.8	<.001	-0.62	-47.20	8.14
			Right	-8.8	<.001	-0.78	-49.76	5.68
P230m	*Mistuned* vs. *Baseline*	Attend	Left	1.7	**0.1 n.s.**	0.19	7.36	4.37
			Right	1.9	**0.07 n.s.**	0.14	8.48	4.40
		Non-attend	Left	3.9	0.002	0.31	15.05	3.91
			Right	2.5	0.03	0.19	10.16	4.11
	Delayed vs. *Baseline*	Attend	Left	4.4	<.001	0.36	29.78	6.76
			Right	5.9	<.001	0.52	32.73	5.57
		Non-attend	Left	4.8	<.001	0.39	23.59	4.90
			Right	6.4	<.001	0.57	28.30	4.42
P75m	*Mistuned* vs. *Baseline*	Attend	Left	5.4	<.001	0.14	23.08	4.30
			Right	4.3	<.001	0.17	24.46	5.71
		Non-attend	Left	3.5	0.003	0.12	17.92	5.08
			Right	5.6	<.001	0.18	25.77	4.63

Delaying the third harmonic

The difference waveforms *Delayed* minus *Baseline* are also shown in Fig. 5. Three important differences between the morphology of *Delayed* grand average difference waveforms and *Mistuned* difference waveforms were observable. First, there was no P75m in the *Delayed* condition. Second, the ORNm-peak occurred later than in the *Mistuned* condition (~277 ms, see Table 2). This was expected due to the fact that the onset of the third harmonic was delayed and sound segregation thus occurred later. Third, the P230m component was more distinct than in mistuned conditions with a clear peak at ~380 ms (Table 2).

In the ORNm and P230m time-ranges, mean amplitudes in the two temporal channel groups were computed for attended and non-attended *Baseline* and *Delayed* conditions as described above. Mean amplitudes were retrieved per participant. Filter settings were also chosen as described above.

As summarised in Table 3, the ORNm was significant, i.e. amplitudes in *Delayed* conditions were higher than in *Baseline* conditions. Likewise, the P230m was significant, i.e. amplitudes in *Baseline* conditions were higher than in *Delayed* conditions. The ORNm and P230m were significant in both temporal channel groups and in both conditions (*Attend* and *Non-attend*, see Table 4).

Comparison between Delayed and Mistuned components regarding amplitudes and peak latencies

Overall, ORNm amplitudes (absolute value) in the *Delayed* conditions were larger than in *Mistuned* conditions ($F(1,15)=10.5$, $p=.005$), as were the P230m amplitudes ($F(1,15)=31.7$, $p=.00005$, see Fig. 5). P75m was only analysed in *Mistuned* conditions, as it was not present in *Delayed* conditions. No further significant main effects or interactions with Hemisphere (*Left, Right*), and Attention (*Attend, Non-attend*) were found.

Peak latencies of P75m and P230m did not differ between *Attend* and *Non-attend* conditions for *Delayed or Mistuned stimuli*. This was also true for ORNm-peak latencies in the *Delayed* condition. However, ORNm-peak latencies for *Mistuned* stimuli were longer in the *Non-Attend* condition than in the *Attend* condition ($t=2.2$, $p=.04$, $r=0.35$, mean latency *Non-attend* $=142.7 \pm 4.4$ ms, mean latency *Attend* $=129.1 \pm 4.9$ ms).

Correlations between localisation performance, attention test results, and MEG components (P75m, ORNm, P230m)

No significant correlations were found between localisation performance in MEG, performance in attention tests (D2, KLT-R), and MEG P75m/ORNm/P230m latencies, or amplitudes.

2.1.4 Discussion (Study 1)

ORNm and P230m as markers of sound segregation

In line with Alain and colleagues (2001; Alain & McDonald, 2007), we identified a magnetic equivalent of the ORN and P400/P230 in response to a harmonic complex with a mistuned partial (inharmonicity). Additionally, we identified an ORNm and P230m following the presentation of a harmonic complex, where the onset of one partial was delayed rather than mistuned (asynchrony). Harmonicity and onset synchrony are both

thought to be indicators for primitive grouping mechanisms that different sounds "belong together" and are thus perceived as one auditory object. Inharmonicity and asynchrony in turn evoke sound segregation (Alain et al., 2001; Bidet-Caulet, Fischer, Bauchet, Aguera, & Bertrand, 2008; Hartmann, McAdams, & Smith, 1990; Moore, Glasberg, & Peters, 1986). Accordingly, participants were more likely to report correctly on which side they had heard the third partial when it was mistuned or delayed compared to being in tune and synchronous with all other harmonics of the complex sound. This suggests that in those conditions, the harmonic was perceived as a separate auditory object.

The morphology of the ORNm and P230m in the MEG waveform was very similar in both *Delayed* and *Mistuned* conditions, with the exception that consistent with the delayed onset of the third harmonic in the *Delayed* condition, ORNm and P230m in that condition also occurred with a delay relative to *Mistuned* components. The latency shift between *Mistuned* and *Delayed* ORNm was 145 ms, this being roughly the same extent of delay of the third partial relative to the onset of the harmonic complex in *Delayed* stimuli (160 ms). The latency shift between *Mistuned* and *Delayed* P230m was on average shorter and more variable than for the ORNm. This can probably be explained by the P230m as a "late", top-down modulated component reflecting a summation of several underlying processes, whereas the ORNm is a more bottom-up, stimulus-driven component.

The fact that ORNm and P230m were evoked by mistuning *and* delaying the third partial of a harmonic complex supports Alain's interpretation that ORNm and P230m are markers of sound segregation in the brain signal. Nevertheless, the *Delayed* ORNm/P230m might solely represent the onset response elicited by the delayed third partial. But as the *Delayed* ORNm/P230m latencies were significantly[4] longer than the expected N1m/P2m latencies of the onset response of the delayed third partial (N1m/P2m of Baseline plus delay of 160 ms), we concluded that the AEFs in the *Delayed* condition indeed index concurrent sound segregation.

Alain et al. (2001, 2002) identified ORN and P400 at fronto-central electrodes. In our data, the components were most prominent at bilateral temporal channels reflecting generators in auditory cortices along the Sylvian fissure. Due to differences between EEG and MEG measurements, auditory activation is maximally picked up by fronto-central

[4] *Right-hemisphere*, *Attend*: t=5.1, p<.001, r=0.67, mean latency ORNm =274.1 ± 4.8 ms, mean latency N1m =247.4 ± 1.8 ms; t=7.6, p<.001, r=0.81, mean latency P230m =375.6 ± 7.6 ms, mean latency P2m =306.8 ± 4.2 ms

electrodes in EEG, whereas it is strongest at temporal sensors in MEG (N'Diaye, Ragot, Garnero, & Pouthas, 2004). Accordingly, Alain and McDonald (2007) showed an ORN and P230 in the source waveforms of bilateral regional dipoles in primary auditory cortex.

A concern was raised in the introduction that in the procedure used by Alain et al. (2001, 2002) participants' responses might not have reflected sound segregation but simply the labelling of two distinct sounds. Our participants responded to the *location* of the mistuned/delayed harmonic, which required them to segregate the mistuned/delayed harmonic as a separate perceptual event; listeners who made their judgements only on the basis of cues such as unevenness or roughness would have been expected to perform at chance levels in the location task. As participants were correct in 80–95% of *Mistuned/Delayed* trials, it is safe to assume that the mistuned/delayed harmonic was heard as a separate entity. In contrast to Alain et al. (2001), who reported that the larger the P400 and ORN amplitude, the more likely participants were to perceive two auditory objects, we did not find correlations between any component amplitude and localisation performance. This may be due to our experimental design, which allowed aspects of the MEG data to be correlated with the perceptual segregation of two auditory objects and not simply with acoustic differences between stimuli. Our findings support the idea that the amplitude of P75m/ORNm/P230m alone does not necessarily reflect the ability to segregate concurrent sounds at an individual level. It seems that further cognitive processes are involved before the conscious decision about the location of an auditory object is made.

Attention modulation of ORNm and P230m

The ORNm was significant in all conditions irrespective of attention. This is consistent with the proposal that concurrent sound segregation may occur independently of listeners` attention. We tried to limit attention to the sounds in the *non-attend* condition by presenting participants a visual localisation task. We used comparable localisation tasks in the two conditions in the hope that the attentional load would be similar in the *Attend* and *Non-attend* conditions, but focused on acoustic stimuli in one case and on visual in the other. Therefore it is likely that attention to the sounds only contributed minimally to the generation of the ORNm during the *Non-attend* task.

In the present study, the third harmonic was the only harmonic that was manipulated. Participants may thus have realized that the only changing partial was always in the same

frequency region and therefore listened more carefully to this particular frequency. For example, individuals are able to identify a single harmonic in a complex sound if they have previously listened to that harmonic presented alone (Bregman, 1990). In line with this finding, participants in the present study were also able to localise the third harmonic in *Baseline* stimuli above chance. However, localisation performance was significantly better in *Mistuned* and *Delayed* conditions. This was true for the headphone presentation and presentation of the sounds through tube-phones during MEG testing. I.e. although participants were able to even hear out the third harmonic *without* perturbation after a training phase, they still significantly gained from inharmonicity and onset delay cues. Further, P75m and ORNm amplitudes were unaffected by attention to the sounds. Thus, automatic registration of the perturbed harmonic even occurred when participants ignored the sounds.

The P230m was significant in *Delayed* conditions, whether attention was paid to the sounds or not. In *Mistuned* conditions, however, the P230m was only significant when participants did *not* attend to the sounds. Interestingly, Alain and McDonald (2007) also reported a P230m in an MEG-study where participants were not attending to the stimuli. It reached significance for 16% mistuning of the third harmonic, but not for 4% mistuning and was reduced in older adults compared to younger adults. In addition, they mentioned a "positive wave that peaked at 245 ms following sound onset that was present during passive listening" in an EEG experiment (Alain et al., 2002). It was affected by sound duration, being larger for medium durations (400 ms) than for shorter (100 ms) or longer (1000 ms) duration sounds. The finding in the present study that the P230m was significant for non-attended *Mistuned* as well as attended and non-attended *Delayed* stimuli implies that the P230m is a general marker of sound segregation. It likely reflects an automatic process, as it was recorded even when participants did not attend to the stimuli. However, P230m amplitude was more sensitive to attentional manipulation than the ORNm. P230m may partly index top-down influences, but seems also to be affected by physical stimulus properties (e.g. duration, frequency periodicity, onset asynchrony) and age. The effect of attention on the P230m is probably weak, which is supported by the finding that the P230m mean amplitudes of the *Attend* and *Non-attend* conditions did not differ.

Differences between inharmonicity and asynchrony

Mistuned stimuli led to an early deflection around 75 ms in the difference waveform Mistuned – Baseline that was similar in both hemispheres. The P75m was significant, whether attention was paid to the sounds or not. This was similar to an early response found in a recent study by Alain and McDonald (2007), which involved greater mistuning (16%) than the present study (8%). In *Delayed* conditions, by contrast, there was no comparable P75m component. Given that ORNm and P230m were present for both *Mistuned* and *Delayed* stimuli, it might be assumed that the P75m reflects an early registration of inharmonicity rather than sound segregation in general. The latency of ~75 ms implies a rapid registration of frequency periodicity which may be taking place in primary auditory cortex. Furthermore, we showed that its amplitude was not affected by attention and therefore seems to represent an automatic registration of the mistuned partial.

ORNm and P230m amplitudes were larger in the *Delayed* than in the *Mistuned* condition. This amplitude difference might be explained by frequency-selective habituation mechanisms. Bidet-Caulet et al. (2008) used stimuli that were each composed of two sounds which could be perceived as one or two streams. Each sound was divided into two halves of which the last half was identical for all stimuli. In the two-stream condition, stimuli started with two partials that were amplitude-modulated at 21 Hz and were joined by another partial that was amplitude-modulated at 29 Hz during the second half of the stimulus. The pitch of the first 21 Hz-modulated partials did not change throughout the stimulus. Thus, the 29 Hz-modulated partial was perceived as a separate sound as it joined the other partials later. In the one-stream condition, the pitch of the 21 Hz-modulated partials changed up or down with the onset of the 29 Hz-modulated partial during the second part of the stimulus. This led to grouping of the 21 Hz and 29 Hz partials. Steady-state activity in response to the two-stream condition was smaller compared to the one-stream condition. The authors explained this reduction by a decrease of steady-state activity in response to the 21 Hz partial from the beginning to the end of the stimulus. This reduction is likely related to frequency-selective habituation (Thompson & Spencer, 1966). Habituation in the one-stream condition was probably less pronounced, as the 21 Hz-modulated partials changed pitch along with the appearance of the 29 Hz partial in the second half of the stimulus. In the *Mistuned* condition of the present study, the habituation process could have developed for the whole stimulus duration and thus led to attenuated activity of all 24 frequencies. In the *Delayed* condition, the third harmonic response might

have been less prone to habituation than the activity of the remaining harmonics, as its onset occurred later. This might have led to over-all larger signals in response to *Delayed* than *Mistuned* tones, where all harmonics (including the third partial) were subject to the same habituation mechanisms. An alternative explanation might be that the third harmonic in the *Delayed* condition gets an extra pre-stimulus-interval, which might have been advantageous in terms of refractory period in the tonotopically organized neurons tuned to that frequency.

It is likely that the processing of *Mistuned* stimuli is different from that of *Delayed* stimuli in several respects, and the differences in the early P75m may be one index of these differences. Equally, grouping mechanisms based on frequency are likely to lead to different patterns of cortical activity from grouping mechanisms based on timing. These factors, which may include frequency-selective habituation mechanisms, are likely to influence the morphology of responses like the ORNm and P230m, without necessarily having consequences on the perceived quality of sound segregation. This idea is supported by the fact that we did not find significant correlations between ORNm or P230m amplitudes and localisation performance in MEG, for both *Delayed* and *Mistuned* conditions.

Conclusion

The present results indicate three MEG components associated with concurrent sound segregation. Whereas the early component (P75m) reflects an automatic registration of the *mistuned* harmonic, later deflections (ORNm, P230m) seem to be more general markers of concurrent sound segregation, as they were elicited by both inharmonicity *and* onset asynchrony. In contrary to the ORNm, which reflects relatively automatic, bottom-up sound segregation processes, the P230m seems to be more sensitive to selective attention, especially when a mistuned harmonic serves as a cue for concurrent sound segregation.

2.2 Developmental changes in concurrent sound segregation (Study 2)

Abstract

It is known that the processing of various auditory stimuli has a maturational time course that extends into adolescence. To learn more about the auditory perceptual abilities of children, maturational changes in concurrent sound segregation - a key function of the auditory system - were investigated. Auditory evoked fields (AEFs) were measured using magnetoencephalography (MEG). Children and adults were presented with harmonic complex tones with (i) all harmonics in-tune, (ii) the third harmonic mistuned by 8% of its original value, or (iii) the onset of the third harmonic delayed by 160 ms compared to the other harmonics. Active and passive listening conditions were compared to evaluate the contribution of attention on sound segregation. Both cues - inharmonicity and onset asynchrony - elicited sound segregation in both groups. However children performed worse than adults. AEFs demonstrate that automatic (bottom-up) sound segregation processes are immature in children for both cues. However, especially for delaying AEF morphology differs conspicuously between children and adults. An early segregation positivity (ESPm) was found to be a more general marker for a bottom-up sound segregation process in children, whereas in adults it probably reflects an automatic registration of the mistuned harmonic. The object-related negativity (ORNm) was adult-like in children and was interpreted as a general marker of sound segregation for both groups.

2.2.1 Introduction (Study 2)

From infancy on we live in a complex acoustic environment. The information that reaches the ear at nearly any moment is a composition of many simultaneously active sounds from different sources, which overlap in many acoustic parameters such as frequency or intensity. A key function of the auditory system is to disentangle the incoming sound stream by assigning acoustic elements to perceptual auditory objects, which correspond to different physical sound sources. For example, in the classroom, children have to identify and parse the acoustic elements that originate from the teacher and ignore other simultaneously present sounds (e.g., noise from outdoors, other pupils talking). The processing stages, which sort the incoming acoustic information into one or more neural representations of auditory objects, can be summarized under the term Auditory Scene Analysis (ASA, Bregman, 1990). Bregman (1990) proposed two processing stages – a

primitive (bottom-up) and a schema-based (top-down) mechanism – that help to decide which components belong to a particular sound source. Schema-governed mechanisms are thought be learnt and to involve active attention. They may use prior knowledge to extract meaning from the acoustic information, for example, for the recognition of a familiar voice. On the other hand, low-level sound segregation is stimulus driven. It involves the grouping of sounds according to frequency, spatial, and temporal aspects. For example, sounds that have a similar stimulus onset, intensity, amplitude modulation, and/or frequency periodicity are more likely to be coming from the same source than sounds that differ in these parameters. This stage is thought to be automatic, innate, and largely independent of listeners` attention. Even though there is evidence that attention is not always required for the formation of auditory streams (= sequential sound segregation), it has been found that in certain circumstances, e.g. depending on task demand and acoustic characteristics of the sounds, attention can affect early stages of auditory processing (Carlyon, Cusack, Foxton, & Robertson, 2001; Cusack, Deeks, Aikman, & Carlyon, 2004; Sussman, Winkler, Huotilainen, Ritter, & Naatanen, 2002; Sussman, Horvath, Winkler, & Orr, 2007). However, support for primitive aspects of sound segregation, especially sequential sound segregation, occurring without much auditory experience comes from studies of 7 – 15-week-old infants (Demany, 1982) and newborns (McAdams & Bertoncini, 1997; Winkler et al., 2003). The basic mechanisms of auditory stream segregation seem to be present and functioning very early in life. Even if not completely developed, the studies showed that babies could organize sound sequences on the basis of frequency proximity (Demany, 1982; Winkler et al., 2003), timbre, and spatial position (McAdams & Bertoncini, 1997), although infants required larger separations between acoustic dimensions and slower paced stimuli to perceive separated sound sources than adults do (Demany, 1982; McAdams & Bertoncini, 1997). Many further studies have also documented that children do not have the same auditory perceptual abilities as adults. For example, children reach adult like levels of the identification of speech in noise at different ages for varying components of speech (e.g., vowels and consonants), some extending into teenage years (Johnson, 2000). Fallon et al. (2000) showed that 5–11 year old children required more favorable signal-to-noise ratios (SNR) than young adults to identify the final word of a sentence presented in background babble. The sentences did not provide contextual clues to the final word, which could have been especially beneficial for adults. Furthermore, equivalent increases in noise level led to similar performance decrements in both groups. Although group differences

between children and adults in speech tasks are likely to be the result of differences in the amount of experience and thus schema-based mechanisms according to Bregman`s logic, it is also possible that children have more difficulties than adults because of immature bottom-up segregation of concurrent acoustic elements. In line with this assumption evidence from auditory evoked potentials (AEP) suggests that low-level cortical processing of various auditory stimuli has a maturational time course that extends into adolescence (for a review see Wunderlich & Cone-Wesson, 2006).

Therefore, it seems plausible that the mechanisms for sound segregation are set from birth but follow the developmental course of the auditory system such that larger differences in cues (e.g., spatial location, frequency, or onset time) might be necessary for automatic sound segregation in children. In this regard, Sussman & Steinschneider (2009) demonstrated, by comparing behavioral measures of frequency separation with passively and actively obtained electrophysiological indices (mismatch negativity (MMN) and P3b), that 9 – 12-year-old children required larger frequency separations to perceive two streams compared to adults. This suggests that differences in stream segregation between children and adults reflect an under-development of basic auditory processing mechanisms.

Regarding the development of ASA, most research has focused on the ability to process sequentially presented sounds. However, Alain and colleagues investigated concurrent sound segregation, measuring the electroencephalogram (EEG) of adults (Alain, Arnott, & Picton, 2001; Alain & Izenberg, 2003; Alain, Schuler, & McDonald, 2002) and 8 – 12-year-old children (Alain, Theunissen, Chevalier, Batty, & Taylor, 2003), while they were presented with complex sounds. The sounds contained multiple harmonics, one of which was mistuned so that it was no longer an integer multiple of the fundamental. The stimuli either promoted sound segregation leading to the perception of two rather than one sound (Segregation condition), or did not promote sound segregation leading to the perception of one sound (No Segregation condition). In the Segregation condition the third harmonic was mistuned causing the third harmonic to stand out perceptually from the sound complex formed by the other harmonics. The authors identified a negative deflection in the difference waveform (Segregation – No Segregation) approximately 180 ms after stimulus onset, referred to as the object-related negativity (ORN) and another later, positive deflection around 400 ms (P400). In adults the ORN was found to be little affected by attention, thus probably reflecting automatic, bottom-up sound segregation (Alain et al.,

2001, 2002; Alain & Izenberg, 2003). In contrast, the P400 was only present in the active listening condition and thus was thought to reflect top-down processes of sound segregation. Concurrent sound segregation in children was found to be adult-like (Alain et al., 2003). Children as well as adults were more likely to perceive the mistuned harmonic as a separate sound when the level of mistuning was increased (0%, 2%, 4%, 8%, and 16%). However, the authors found a larger ORN amplitude and no P400 in children. Additionally, children were less likely to detect mistuning across all levels than adults. The authors concluded that the bottom-up segregation of concurrent sounds based on inharmonicity is adult-like in pre-adolescent children, but children are less efficient than adults in processing the information following the detection of mistuning. Thus it appears that concurrent sound segregation matures during adolescence. However, Alain and colleagues did not manipulate attention during the EEG measurements making it difficult to determine whether the ORN in children reflects automatic concurrent sound segregation occurring independently of attention to the sounds.

Therefore, the present study extended prior research of developmental changes of concurrent sound segregation in two ways. First, active and passive listening was chosen to evaluate the contribution of attention on sound segregation making it possible to determine to which degree bottom-up sound segregation is developed in children. Second, an additional type of stimulus was included that evokes segregation by delaying the onset of one harmonic relative to the other harmonics. Choosing a temporal cue to invoke concurrent sound segregation was motivated by previous research. More precisely, it has been demonstrated that infants and children have higher auditory thresholds than adults for temporal resolution, which is crucial for speech perception and probably for concurrent sound segregation based on onset asynchrony (Morrongiello & Trehub, 1987; Wang, Datta, & Sussman, 2005; Werner, Marean, Halpin, Spetner, & Gillenwater, 1992). Accordingly, for a better general understanding of the development of sound segregation in children, drawing a comparison between concurrent sound segregation based on two distinct cues — inharmonicity and onset asynchrony — was considered to be of great interest.

The same experiment, previously reported by Lipp and colleagues (2010), was chosen to measured auditory evoked fields (AEFs) using magnetoencephalography. Participants were presented with complex sounds with (i) all harmonics in-tune (= *Baseline*) (ii) the third harmonic mistuned by 8% of its original value (iii) the onset of the third harmonic

delayed by 160 ms compared to the other harmonics. During recording, participants listened to the sounds and performed an auditory localisation task (*Attend* condition) whereas in another session they ignored the sounds and performed a visual localisation task (*Non-attend* condition). Previous analysis found an early deflection at ~75 ms (P75m) after sound onset in the difference waveforms (Mistuned – Baseline), probably reflecting an automatic registration of the mistuned harmonic (Lipp et al., 2010). Subsequent deflections, the object-related negativity (ORNm) and a later displacement (P230m) seem to be more general markers of concurrent sound segregation, as they were elicited by both mistuning and delaying the third harmonic. Results indicate that the ORNm reflects relatively automatic, bottom-up sound segregation processes, whereas the P230m may be more sensitive to attention and may partly index top-down influences. But as the effect of attention on the P230m was weak and detectable only when a mistuned harmonic had to be segregated, the P230m was thought to likely reflect an automatic process as well. To ease comparisons of the components between children and adults the P75m is referred to as the *early segregation positivity* (ESPm) and the P230m is referred to as the *late segregation positivity* (LSPm) in the following.

The present MEG-study aimed to investigate differences in the process of concurrent sound segregation between 9 – 13-year-old children and young adults. To distinguish between bottom-up (stimulus driven) and top-down (schema-driven) mechanisms of concurrent sound segregation the attention load on the sounds was manipulated. Additionally, two cues invoking sound segregation (inharmonicity and onset asynchrony) were used to investigate, if the time course of maturation of general markers of concurrent sound segregation (ORNm, LSPm) is comparable for both cues. We hypothesized that if concurrent sound segregation in children is not fully developed, they should perform poorer than adults in the sound localization task. Further, if bottom-up mechanisms of sound segregation are immature in children, AEF components, which potentially reflect an automatic concurrent sound segregation process in adults, should be elicited less robustly, especially, when children do not attend to the sounds. Learning more about maturational changes in processing acoustic temporal and spectral characteristics may also contribute to a better understanding of the development of speech perception in children (Fitch, Miller, & Tallal, 1997).

2.2.2 Methods (Study 2)

Participants

Fifteen children (mean age 11.0 years, range 9-13 years, 5 female, 13 right-handed, 2 bimanual) and fifteen adults (mean age 24.2 years, range 20-28 years, 7 female, 14 right-handed, 1 left-handed) took part in the experiment. Handedness was assessed with the Edinburgh handedness inventory (Oldfield, 1971). Adults were contacted through advertisements at the University of Konstanz. Children were recruited at local schools. Ethical approval of the study was granted by the Ethical Review Board of the University of Konstanz. Participants´ pure-tone audiometric thresholds were determined with clinical audiometer (AC40).

MEG recordings

In session one, hearing thresholds were determined and participants were familiarised with the experiment. In session two, MEG data were collected in two conditions: a) participants listened to sounds and performed an auditory localisation task (*Attend* condition), and b) participants were played the same sounds while performing a visual localisation task (*Non-attend* condition). *Attend* and *Non-Attend* conditions were presented in a counterbalanced order.

Acoustic Stimuli

Stimuli were created at the Department of Psychology of the University of York, UK (Summerfield, Paul, Kitterick, Briley, & Bailey, 2007). *Baseline* stimuli were complex tones consisting of 24 sinusoidal partials (fundamental frequency = 200 Hz), which were synthesised digitally (44100 samples/s, 16-bit amplitude quantization) with equal amplitudes. Two variants of the *Baseline* stimulus were formed by perturbing the third partial at 600 Hz. In *Mistuned* stimuli, the frequency of the third partial was raised by 8% to 648 Hz. In *Delayed* stimuli, the third partial started 160 ms after the other partials, but ended simultaneously (Fig. 1). One hundred samples of each stimulus condition (*Baseline, Mistuned, Delayed*) were created; stimuli lasted 500 ms.

Stimuli were created so that the third partial was always perceived 45° left or right of straight ahead, with equal probability. For specifics of stimulus generation, please see (Lipp et al., 2010).

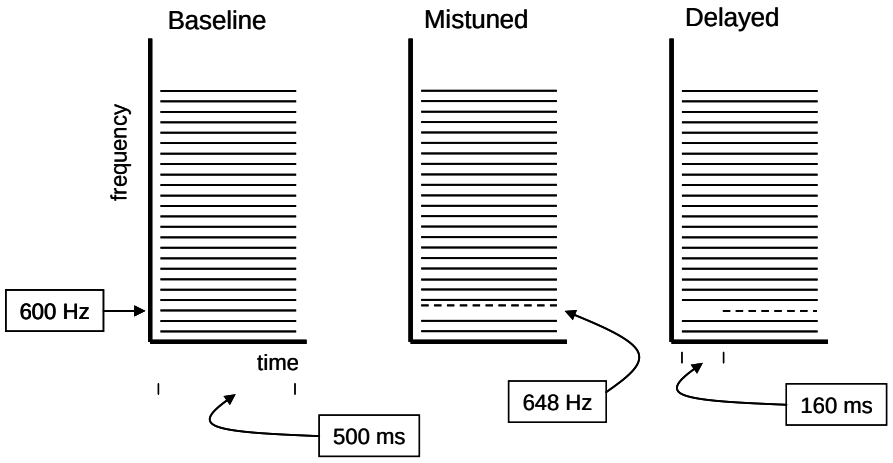

Figure 1: Schematic spectrograms of Baseline, Mistuned and Delayed stimuli.

In order to familiarise participants with the experiment, a training session preceded the MEG-recordings. Stimuli were presented through headphones (Dynamic Stereo Headphones MDR-CD470, Sony), while participants performed the localisation task (see below). For the MEG recordings stimuli were presented through tube-phones (Etymotic Research, ER30). For each ear, a 6.1-m length of 4-mm diameter tube terminated in a right-angle adapter (ER30-9) ending with a disposable foam eartip (ER13-14 for adults, ER13-14B for children). A digital filter was designed to compensate for the low-pass frequency response of this sound delivery system, resulting in a frequency response that was flat to within -4–0 dB from 100–4800 Hz when measured with a B&K Ear Simulator (Type 4157) with 0.5-inch microphone (Type 4134), Microphone Preamplifier (ZC 0032) and B&K Hand-held Analyzer (Type 2250). Using the same measurement system, the presentation level of *Baseline* stimuli was set to 70 dB (A).

Task

On each trial of the *Attend* condition, participants reported whether they heard the third partial on the left or right using two keys on a keyboard. Feedback on the accuracy of the response was given by displaying a coloured rectangle on the side where the key had been pressed. It was green if the response was correct or red if the response was incorrect. The inter-stimulus interval (ISI) was the response latency plus 500 ms.

On each trial of the *Non-attend* condition, a faint grey circle subtending a visual angle of 1° appeared for 50 ms at a randomly-chosen location within 8° right or left of the fixation cross. Participants were instructed to report whether the circle appeared left or right of the

fixation cross by making a corresponding key press. Feedback was provided in the same format as used in the *Attend* condition. The ISI was the response latency plus a random value in the range from 500 to 1500 ms. At the same time, the acoustic stimuli were presented asynchronously with an ISI that was chosen randomly from the range 500 to 1500 ms. Participants were instructed to concentrate on the visual stimuli and to ignore acoustic stimuli.

MEG recordings

Per attention condition (*Attend, Non-attend*) one hundred samples of each stimulus type (*Baseline, Mistuned, Delayed*) were presented randomly. Recording was done with a 148-channel magnetometer (MAGNESTM 2500 WH, 4D Neuroimaging, San Diego, USA). Acoustic stimuli were delivered with an analog-to-digital converter (Motu 2408) and amplifier (Servo 200, Samson) through tube-phones (Etymotic Research, ER30). The foam tips were introduced carefully into the ear canals of participants and were additionally secured with medical tape. Participants lay on a height-adjusTable bed in a magnetically shielded room (Vakuumschmelze Hanau). Eye movements (EOG) were recorded from four electrodes attached to the left and right temple and above and below the right eye, as well as cardiac activity (ECG) via two electrodes, one on each forearm. A Neurofax amplifier (EEG-11006, Nihon Kohden) served for the recording of EOG and ECG. A video camera installed inside the chamber allowed monitoring the subject's behaviour and compliance throughout the experiment. Visual stimuli and feedback were projected (D-ILA Projector, JVC™, DLA-G11E) onto a screen placed ~40cm away from the participants' eyes through a mirror system. Data from each channel were band-pass filtered between 0.1 and 200 Hz and sampled at a rate of 678.17 Hz. Recording was continuous.

MEG data analysis

Data were corrected for ocular and cardiac artefacts using BESA™ 5.3 and were averaged separately for *Baseline, Delayed,* and *Mistuned* stimuli in *Attend* and *Non-attend* conditions using a pre-stimulus baseline of 100 ms and a post-stimulus time-window of 800 ms. Epochs were excluded if the signal amplitude in any channel exceeded 3.5 pT or if the signal amplitude differed by more than 2.5 pT between adjacent sampling points. The minimum number of averaged epochs was 68 for children and 65 for adults. The average number of averaged epochs was 94 for children and 93 for adults.

The sensor groups of interest were two groups of 7 sensors located bilaterally over the temporal lobes, which showed the biggest N1m amplitudes in adults and accordingly the biggest N2m amplitudes in children in the averaged cortical response elicited by *Baseline* stimuli (Fig. 2). This approach was chosen, as it was objective and ensured that sensor groups reflected auditory processing (the auditory N2 component is very prominent in children) (Johnstone, Barry, Anderson, & Coyle, 1996; Sussman, Steinschneider, Gumenyuk, Grushko, & Lawson, 2008). The left channel group covered the in-going magnetic field (positive values) and the right channel group the out-going magnetic field (negative values) of the underlying auditory sources. Therefore, the sign of the signals over the left hemisphere was reversed to facilitate comparisons with right-hemisphere signals.

Mean amplitudes in time-windows of interest were computed for each sensor group, participant, stimulus, and condition. Time-windows were chosen to embrace prominent deflections from zero in either direction in the difference waveforms (*Mistuned minus Baseline* and *Delayed minus Baseline*). Mean amplitudes were analysed statistically with analyses of variance (ANOVAs). Condition *(Baseline, Mistuned, Delayed),* Hemisphere (*Left, Right*), and Attention (*Attend, Non-attend*) were within-group variables. Group *(children, adults)* was used as between group factor [5]. Additionally, paired t-tests and effect sizes (r) were computed for all relevant contrasts using the following formulae:

$$\text{Cohen's } d = M_1 - M_2 \,/\, \sigma_{pooled}$$

$$\text{where } \sigma_{pooled} = \sqrt{[(\sigma_1^2 + \sigma_2^2) / 2]}$$

$$r_{Y\lambda} = d / \sqrt{(d^2 + 4)}$$

The differences of the mean amplitudes (*Mistuned – Baseline, Delayed – Baseline*) characterising the components (ESPm, ORNm, LSPm) were analysed statistically with ANOVAs. Hemisphere (*Left, Right*) and Attention (*Attend, Non-attend*) were within-group variables. Group *(children, adults)* was used as between group factor.

To specify individual peak latencies, individual difference waveforms (*Mistuned–Baseline* and *Delayed–Baseline*) for both conditions (*Attend, Non-attend*) and both hemispheres (*Left, Right*) were band-pass filtered from 1 to 25 Hz for detecting ESPm and ORNm, and

[5] Except for ESPm and LSPm in *Delayed* condition, because for adults no ESPm and for children no LSPm was identified. Consequently, comparisons of ESPm and LSPm (e.g. differences of the mean amplitudes and peak latencies) between adults and children in *Delayed* condition could not be calculated.

from 1 to 8 Hz for detecting LSPm. Peak latencies were compared between conditions with ANOVAs for each component (ESPm, ORNm, LSPm). Hemisphere (*Left*, *Right*) and Attention (*Attend*, *Non-attend*) were within-group variables. Group *(children, adults)* was used as between group factor.

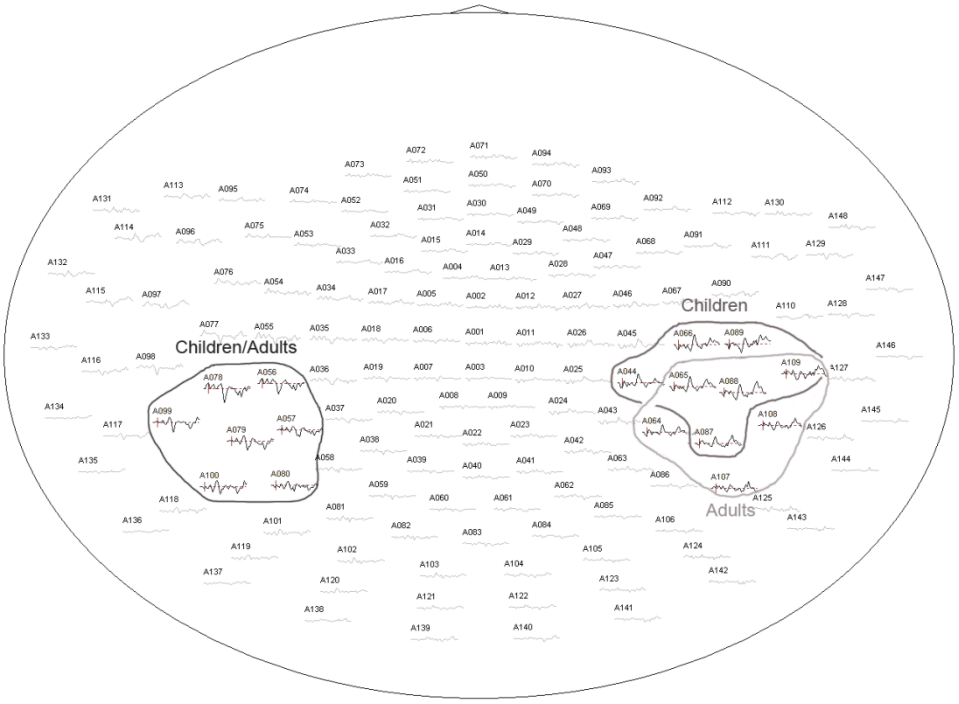

Figure 2: Difference waveforms (Delayed - Baseline) for each sensor averaged over 15 children in the Attend condition. Channel Groups used for data analysis are framed (7 channels for each hemisphere). Left hemisphere (same channels for adults and children): A056, A057, A078, A079, A080, A099, and A100. Right hemisphere: A044, A065, A066, A087, A088, A089, A109 for children, and A64, A65, A87, A88, A107, A108, A109 for adults.

Behavioural data

The percentages of correct responses to *Baseline*, *Mistuned*, and *Delayed* stimuli were compared in a multivariate analysis of variance (MANOVA). Condition (*Baseline*, *Mistuned*, *Delayed*) was within-group variable. Group *(children, adults)* was used as between group factor. Further, Spearman´s rank correlations were calculated between behavioural results in MEG and amplitudes/latencies of MEG components (ESPm, ORNm, LSPm). P-values were corrected for multiple tests using the Bonferroni-Holm method. Hearing thresholds for the frequencies 1 kHz, 2 kHz and 4 kHz were compared between groups with t-tests. *Post-hoc* analyses were performed with Tukey's HSD test. All analyses were conducted with Statistica (StatSoft, Inc., Version 6, 2003).

2.2.3 Results (Study 2)

Hearing thresholds

All participants had pure-tone audiometric thresholds in the normal range (≤20 dB HL) for both ears and all tested frequencies (200 Hz, 595 Hz, 1 kHz, 2 kHz and 4 kHz for children, 500 Hz, 1 kHz, 2 kHz and 4 kHz for adults). All mean hearing levels were lower than 6 dB HL in both ears. T-tests showed that children had a lower threshold for detecting the 4 kHz frequency in the left ear ($t(28)=3.04$, $p=.005$). No further significant differences between adults and children were found.

Behavioural results

Adults and children performed well above chance localising the target stimulus partial in the *Delayed* and *Mistuned* condition, indicating that they segregated it successfully from the harmonic background (main effect Condition ($F(2,27)=54.38$, $p<.0001$; *Delayed*>*Baseline* $p=.0001$, *Mistuned*>*Baseline* $p=.0001$, Fig. 3). Overall, children performed worse than adults ($F(1,28)=28.25$, $p<.0001$). No other significant main effects or interactions were found.

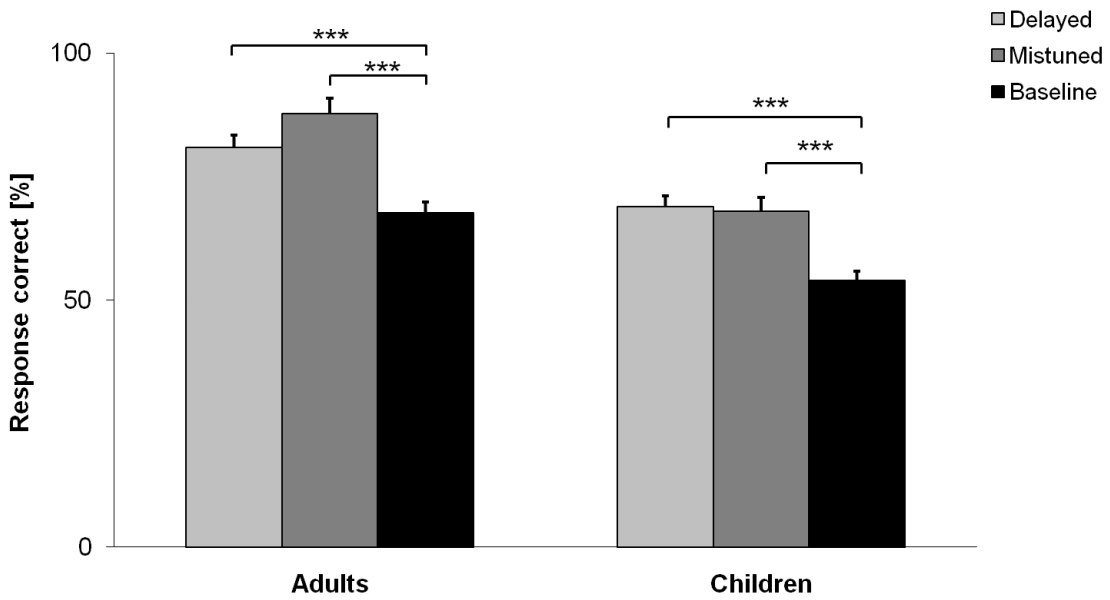

Figure 3: Localisation performance of adults and children during the MEG scan for Baseline, Mistuned, and Delayed stimuli. Error bars indicate standard errors.

Cortical response elicited by Baseline stimuli

The adult grand average AEFs elicited by *Baseline* stimuli comprised a P1m with a peak latency of ~ 45 ms after stimulus onset, an N1m (~ 85 ms) and a P2m (~150 ms), followed by a sustained field (~300 – 400 ms). In children the N1m (~110 ms) with a peak amplitude of ca. -85 fT was less prominent in contrast to a clearly evoked N2m (~225 ms) with a peak amplitude of ca. -180 fT (see Fig. 4 for right hemisphere, *Non-Attend* condition). As in adults, AEFs in children comprised a P1m (~65 ms), but lacked a P2m.

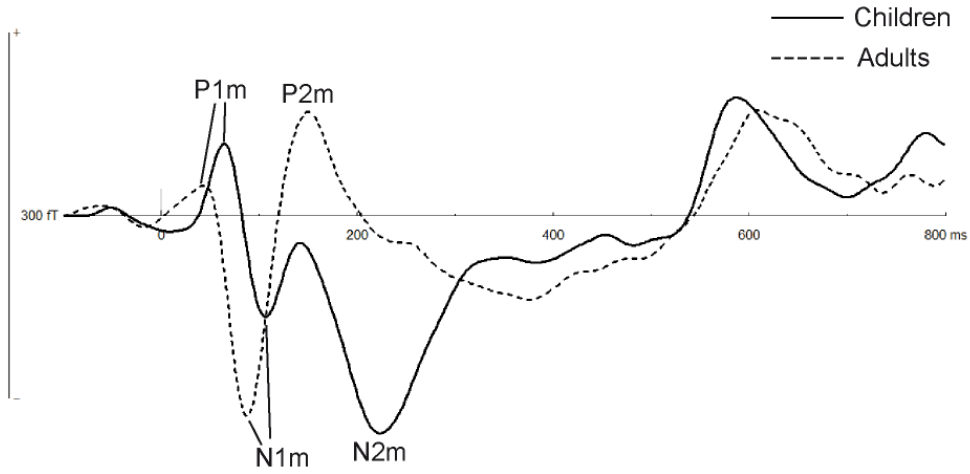

Figure 4: Waveform averaged over adults (15) and children (15) in the right hemisphere (7 channels), evoked by Baseline stimuli in the Non-Attend condition; band-pass filtered from 1 to 25 Hz.

Mistuning the third harmonic

The difference waveforms *Mistuned* minus *Baseline* for children and adults (see Fig. 5 and Fig. 6) were characterised by distinct components peaking at ~87 ms (ESPm) and ~153 ms (ORNm) in children and at ~73 ms (ESPm) and ~132 ms (ORNm) in adults (see Table 1). This likely corresponds to P75m and ORNm already described by Lipp and colleagues (2010) and to the early positivity (~80 ms), and the object-related negativity (ORN, ~160 ms) originally identified by Alain et. al (2007). Additionally, a broader component was identified that peaked at ~320 ms for children and ~257 ms for adults (Fig. 6 and Table 1). A similar component was found in adults, which was labelled P400 (Alain et al., 2001), P230 (Alain & McDonald, 2007) or P230m (Lipp et al., 2010). Here is it referred to as the late segregation positivity (LSPm).

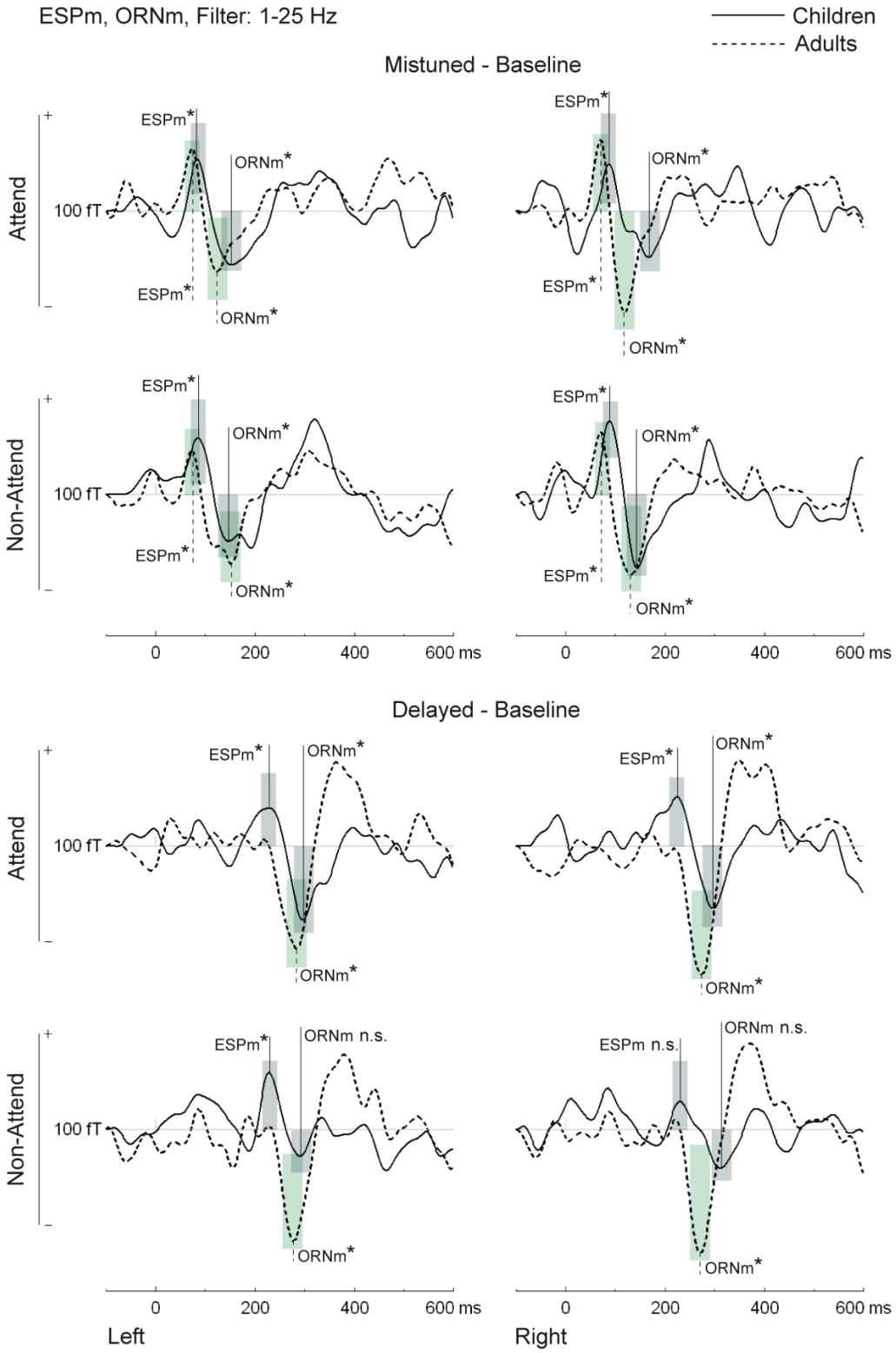

Figure 5: Difference waveforms for components ESPm and ORNm in attended and non-attended conditions for left and right hemispheric channel groups averaged over all participants (15) of each group (children/adults). Difference waveforms (Mistuned – Baseline) are displayed in the upper four graphs; difference waveforms (Delayed – Baseline) are displayed in the lower four graphs (here: no ESPm for adults are marked). The polarity of the signal is reversed in the left hemisphere. Coloured boxes emblematize time-windows used for statistical analyses. Signals were band-pass filtered from 1 to 25 Hz.

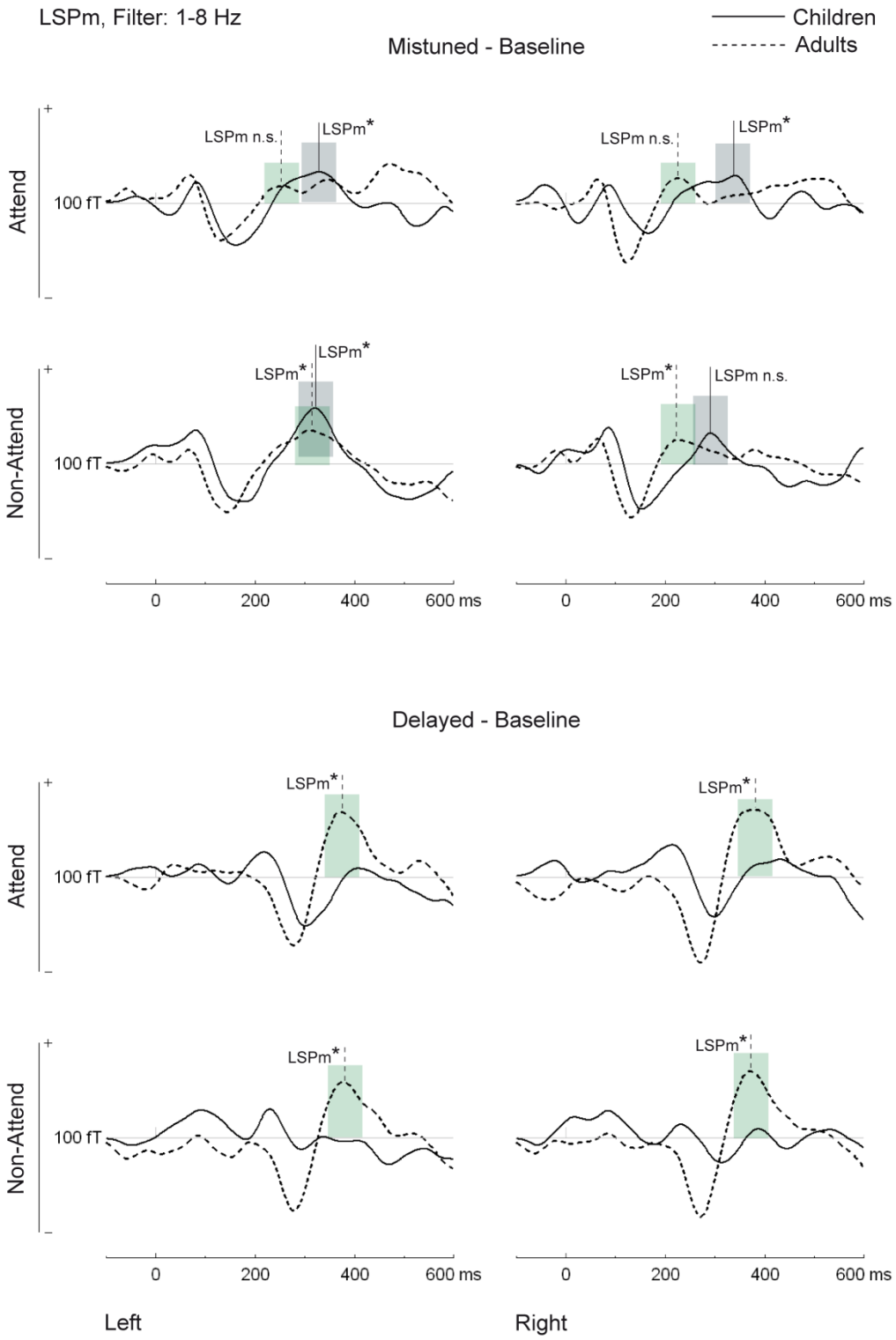

Figure 6: Difference waveforms (Mistuned – Baseline) for component LSPm in attended and non-attended conditions for left and right hemispheric channel groups averaged over all participants (15) of each group (children/adults). Difference waveforms (Mistuned – Baseline) are displayed in the upper four graphs; difference waveforms (Delayed – Baseline) are displayed in the lower four graphs (here: no LSPm for children are marked). The polarity of the signal is reversed in the left hemisphere. Coloured boxes emblematize time-windows used for statistical analyses. Signals were band-pass filtered from 1 to 8 Hz.

Table 1: Peak latencies for components ESPm, ORNm and LSPm in attended and non-attended conditions for left- and right-hemispheric channel groups averaged over all children and adults. Time-ranges used for calculating mean amplitudes are displayed.

Peak Latencies and Time-Ranges				**Children** (n=15)			**Adults** (n=15)		
Component	**Comparison**	**Attention**	**Hemis-phere**	**Peak Latency (ms)**	**Time-Range (ms)**	**Mean Peak Latency (ms)**	**Peak Latency (ms)**	**Time-Range (ms)**	**Mean Peak Latency (ms)**
ESPm	*Mistuned* vs. *Baseline*	Attend	Left	84	69-99	87	74	59-89	73
			Right	89	74-104		71	56-86	
		Non-attend	Left	86	71-101		74	59-89	
			Right	89	74-104		71	56-86	
	Delayed vs. *Baseline*	Attend	Left	230	215-245	229	----	----	----
			Right	226	211-241		----	----	
		Non-attend	Left	230	215-245		----	----	
			Right	231	216-246		----	----	
ORNm	*Mistuned* vs. *Baseline*	Attend	Left	153	133-173	153	124	104-144	132
			Right	168	148-188		119	99-139	
		Non-attend	Left	147	127-167		153	133-173	
			Right	145	125-165		132	112-152	
	Delayed vs. *Baseline*	Attend	Left	298	278-318	300	285	265-305	277
			Right	298	278-318		274	254-294	
		Non-attend	Left	292	272-312		279	259-299	
			Right	311	291-331		271	251-291	
LSPm	*Mistuned* vs. *Baseline*	Attend	Left	328	293-363	320	257	222-292	257
			Right	339	304-374		229	194-264	
		Non-attend	Left	321	286-356		313	278-348	
			Right	291	256-326		229	194-264	
	Delayed vs. *Baseline*	Attend	Left	----	----	----	375	340-410	378
			Right	----	----		383	348-418	
		Non-attend	Left	----	----		379	344-414	
			Right	----	----		373	338-408	

Mean amplitudes in the two temporal channel groups were computed for attended and non-attended *Baseline* and *Mistuned* conditions in specific time-ranges. ESPm time-ranges covered ±15 ms around the ESPm peak latency. Mean amplitudes in the ORNm time-ranges covered ±20 ms around the ORNm peak latency, and for LSPm time-ranges covered ±35 ms around the LSPm peak latency (Table 1). Mean amplitudes were retrieved per participant (Table 2).

Main effects for condition (Table 3) show, that the ESPm was significant, i.e. *Baseline* stimuli led to higher amplitudes than *Mistuned* stimuli for both hemispheres and both attention conditions (*Attend* and *Non-Attend*). Likewise, the ORNm was significant, i.e. amplitudes in *Mistuned* conditions were higher than in *Baseline* conditions. Furthermore, the LSPm was significant, i.e. amplitudes in *Baseline* conditions were higher than in *Mistuned* conditions. The interaction Condition *(Baseline, Mistuned, Delayed)* * Group *(children, adults)* was not significant.

Table 2: Mean amplitudes and standard errors of the difference waveforms (Mistuned–Baseline and Delayed–Baseline) characterizing components ESPm, ORNm and LSPm displayed in attended and non-attended conditions for left- and right-hemispheric channel groups averaged over all children and adults.

Mean Amplitudes of Components			**Children** (n=15)			**Adults** (n=15)	
Component	**Comparison**	**Attention**	**Hemisphere**	**Mean Amplitude (fT)**	**Standard Error (fT)**	**Mean Amplitude (fT)**	**Standard Error (fT)**
ESPm	*Mistuned–Baseline*	Attend	Left	22.51	9.20	27.46	4.28
			Right	20.08	6.49	29.24	5.40
		Non-attend	Left	26.87	6.90	18.33	5.80
			Right	33.77	8.00	26.26	4.72
	Delayed–Baseline	Attend	Left	18.93	6.88	----	----
			Right	23.48	8.44	----	----
		Non-attend	Left	25.61	7.72	----	----
			Right	12.76	5.97	----	----
ORNm	*Mistuned–Baseline*	Attend	Left	-26.74	6.55	-27.32	5.49
			Right	-20.94	6.83	-44.62	5.12
		Non-attend	Left	-22.23	7.18	-31.03	5.68
			Right	-32.42	10.05	-38.42	6.64
	Delayed–Baseline	Attend	Left	-33.41	5.7	-48.14	5.20
			Right	-28.68	8.43	-60.91	7.28
		Non-attend	Left	-11.21	8.83	-51.49	7.54
			Right	-17.83	8.49	-56.64	5.12
LSPm	*Mistuned–Baseline*	Attend	Left	14.79	5.22	7.00	5.79
			Right	11.11	5.11	10.91	5.23
		Non-attend	Left	24.74	5.56	16.20	4.50
			Right	12.10	7.30	10.65	4.61
	Delayed–Baseline	Attend	Left	----	----	30.74	7.16
			Right	----	----	34.92	5.76
		Non-attend	Left	----	----	26.06	5.02
			Right	----	----	30.92	4.86

Table 3: Main Effect of Sound Condition. F- and p-values of ANOVAs with Group (children, adults) as between group factor. Exception: ANOVAs for ESPm and LSPm in Delayed condition were calculated respectively for one group.

Component	Main Effect Sound Condition	F(1,28)/ F(1,14)*	p
ESPm	*Mistuned* vs. *Baseline*	96.90	<.001
	Delayed vs. *Baseline (children only)*	29.01*	<.001
ORNm	*Mistuned* vs. *Baseline*	92.79	<.001
	Delayed vs. *Baseline*	125.97	<.001
LSPm	*Mistuned* vs. *Baseline*	36.82	<.001
	Delayed vs. *Baseline (adults only)*	42.79*	<.001

Detailed results for each group are summarized in Table 4. Whereas in children all components (ESPm, ORNm, and LSPm) were significant with only one exception (LSPm, *Non-attend*, right hemisphere, p=0.1), late positivities (LSPm) were non-significant in adults irrespective of hemisphere, when they attended the sounds (LSPm, *Attend*, right hemisphere, p=.07, left hemisphere, p=0.2).

Delaying the third harmonic

Differences in morphology between *Delayed minus Baseline* waveforms and *Mistuned minus Baseline* waveforms were found for both adults and children (see Fig. 5 and Fig. 6). However, two noticeable differences between children and adults were observable. First, adults elicited no ESPm in the *Delayed* condition and second, children showed no clear LSPm.

The ESPm-peak in children and ORNm-peak in both groups occurred later than in the *Mistuned* condition (ESPm ~229 ms for children; ORNm ~300 ms for children and ~277 ms for adults, see Table 1), which was expected, because the onset of the third harmonic was delayed and sound segregation thus occurred later. Corrected for the 160 ms delay of the third harmonic, the ESPm- and ORNm-peak would occur approximately in the same

time-range or even a little earlier than in the *Mistuned* condition (corrected for delay: ESPm ~69 ms for children; ORNm ~140 ms for children and ~117 ms for adults).

Table 4: Paired t-test results of the comparison between conditions (Baseline, Mistuned, and Delayed) for each component (ESPm, ORNm, LSPm) and both groups (children, adults).

Comparison between Conditions				**Children** (n=15)			**Adults** (n=15)		
Component	**Comparison**	**Attention**	**Hemis-phere**	**t(14)**	**Effect Size r**	**p**	**t(14)**	**Effect Size r**	**p**
ESPm	*Mistuned* vs. *Baseline*	Attend	Left	2.4	0.19	0.03	6.4	0.18	<.001
			Right	3.1	0.11	0.008	5.4	0.23	<.001
		Non-attend	Left	3.9	0.26	0.002	3.2	0.13	0.007
			Right	4.2	0.23	<.001	5.6	0.20	<.001
	Delayed vs. *Baseline*	Attend	Left	2.7	0.15	0.02	----	----	----
			Right	2.8	0.18	0.01	----	----	----
		Non-attend	Left	3.3	0.16	0.005	----	----	----
			Right	2.1	0.11	**0.05 n.s.**	----	----	----
ORNm	*Mistuned* vs. *Baseline*	Attend	Left	-4.1	-0.19	0.001	-5.0	-0.18	<.001
			Right	-3.1	-0.20	0.008	-8.6	-0.28	<.001
		Non-attend	Left	-3.1	-0.20	0.008	-5.5	-0.20	<.001
			Right	-3.2	-0.21	0.006	-5.8	-0.27	<.001
	Delayed vs. *Baseline*	Attend	Left	-5.9	-0.39	<.001	-9.3	-0.59	<.001
			Right	-3.4	-0.21	0.004	-8.4	-0.64	<.001
		Non-attend	Left	-1.3	-0.12	**0.2 n.s.**	-6.8	-0.64	<.001
			Right	-2.1	-0.14	**0.05 n.s.**	-11.1	-0.79	<.001
LSPm	*Mistuned* vs. *Baseline*	Attend	Left	2.8	0.18	0.01	1.5	0.15	**0.2 n.s.**
			Right	2.2	0.08	<.05	2.1	0.16	**0.07 n.s.**
		Non-attend	Left	4.4	0.27	<.001	3.6	0.33	0.003
			Right	1.7	0.10	**0.1 n.s.**	2.3	0.19	0.04
	Delayed vs. *Baseline*	Attend	Left	----	----	----	4.3	0.35	<.001
			Right	----	----	----	6.1	0.47	<.001
		Non-attend	Left	----	----	----	5.2	0.44	<.001
			Right	----	----	----	6.4	0.46	<.001

In the ESPm, ORNm, and LSPm time-ranges, mean amplitudes in the two temporal channel groups were computed for attended and non-attended *Baseline* and *Delayed* conditions as described above. Mean amplitudes were retrieved per participant. As summarised in Table 3, the ESPm was significant in children, i.e. *Baseline* stimuli led to higher amplitudes than *Delayed* stimuli for both attention conditions (*Attend* and *Non-Attend*) and hemispheres. Likewise, the ORNm was significant, i.e. amplitudes in *Delayed* conditions were higher than in *Baseline* conditions irrespective of group, attention condition, and hemisphere. This

suggests that both groups generated the ORNm characterizing concurrent sound segregation elicited by onset asynchrony. Finally, the LSPm was significant in adults, i.e. amplitudes in *Baseline* conditions were higher than in *Delayed* conditions irrespective of attention condition and hemisphere.

Table 4 summarises the significance of the components ESPm, ORNm, and LSPm per group. Whereas the ESPm was significant in children when they attended to the sounds and only missed significance in the right hemisphere when they did not attend to the sounds (left hemisphere p=.005, right hemisphere p=.05), it was completely absent in adults. In contrast, the LSPm was absent in children, but was significant in adults in both hemispheres, regardless if attention was paid to the sounds or not. Finally, ORNm was significant in both children and adults when they paid attention to the sounds. In adults, this was also true for the *Non-attend* condition in both hemispheres, whereas the component did not reach significance in the left hemisphere (p=0.2) and only missed significance in the right hemisphere (p=.05) in children.

Comparison of component amplitude and peak latency between children and adults

Overall, children had smaller ORNm amplitudes than adults in the *Delayed* condition ($F(1,28)=21.05$, $p<.0001$). Comparison of component peak latencies for *Delayed* stimuli, indicated longer ORNm peak latencies for children than adults ($F(1,27)=11.18$, $p=.002$). Further, a significant main effect was found for ORNm peak latencies in the *Mistuned* condition ($F(1,28)=18.66$, $p=.0002$), indicating that ORNm peak latencies were longer for children than adults. Post-hoc testing of the significant interaction Attention (*Attend, Non-attend*)*Group *(children, adults)* ($F(1,28)=8.93$, $p=.006$) in the *Mistuned* condition showed that children had longer ORNm peak latencies than adults only when attention was paid to the sounds (p=.0005), whereas children did not differ in ORNm peak latencies from adults when they did not attend to the sounds (p=0.5). For *Mistuned* stimuli, children also had longer ESPm ($F(1,24)=16.15$, $p=.0005$) and LSPm ($F(1,27)=26.21$, $p=.0001$) peak latencies than adults. Eventually, a main effect for Hemisphere (*Left, Right*) ($F(1,27)=11.95$, $p=.002$) was found for the LSPm, indicating that peak latencies in the left hemisphere were longer than in the right hemisphere.

Correlations between localisation performance and MEG components (ESPm, ORNm, LSPm)

Across both groups of participants no significant correlations were found between localisation performance (*Mistuned, Delayed* and *Baseline* condition) and MEG component amplitude. However, significant correlations between localisation performance in the *Baseline* and *Mistuned* condition and MEG component latency were found, when stimuli were mistuned and participants attend them (see Table 5). Results indicate, that the shorter the peak latencies of ESPm, ORNm, and LSPm were, especially in the right hemisphere, the better participants segregated the target harmonic in *Baseline*, and respectively *Mistuned* stimuli.

Table 5: Significant results of Spearman´s rank correlations between localisation performance (Mistuned and Baseline condition) and MEG component latency (ESPm, ORNm, LSPm).

Behaviour	**&**	**Peak Lateny (*Attend, Mistuned*)**	**n**	**Spearman R**	**p**	**α-level**
Correct Mistuned	&	ORNm right	30	-0.58	0.0007	0.0062
		ESPm right	26	-0.57	0.002	
		LSPm left	30	-0.52	0.003	
		LSPm right	29	-0.52	0.004	
Correct Baseline	&	ORNm right	30	-0.54	0.002	0.0042
		ESPm right	26	-0.56	0.003	
		LSPm left	30	-0.52	0.004	
		LSPm right	29	-0.53	0.003	

n=number of valid cases, α-level after Bonferroni-Holm correction

2.2.4 Discussion (Study 2)

Behavioural differences in concurrent sound segregation between adults and children

Both, adults and children were more likely to correctly report on which side they had heard the third partial when it was mistuned or delayed compared to being in tune and synchronous with all other harmonics of the complex sound. This implies that in those conditions, the harmonic was perceived as a separate auditory object. Accordingly, as in adults, inharmonicity and onset asynchrony evoked sound segregation already in children. However, overall, children performed worse than adults. This is in line with results obtained by Alain et al. (2003) who reported that children were less sensitive in detecting

mistuning than adults. The present data suggest that processing mechanisms for concurrent sound segregation are still underdeveloped in children ranging from 9 to 13 years. The demonstrated behavioural changes in concurrent sound perception from childhood to adulthood were paralleled by changes in neuromagnetic activity, as discussed below.

Cortical response to a complex harmonic in adults and children

To begin with, AEFs elicited by complex sounds showed morphological differences between the groups, demonstrating an immature auditory processing in children compared to adults. In line with a previous MEG study of Alain, Quan, McDonald, & Van Roon (2009) complex sounds consistently evoked a typical triphasic P1m-N1m-P2m complex in adults (for a review see Wunderlich & Cone-Wesson, 2006). In children, the present results show differences in the AEF morphology compared to adults: the N2m component was stronger, whereas N1m and P2m elicitation was weaker in children. These findings are consistent with studies of age-related changes in auditory evoked potentials and magnetic fields, which showed a dominant N2 component (the electrical counterpart of the N2m wave) in 8 – 11-year-old children (Sussman et al., 2008), respectively N250m (synonymic to N2m) in 6 – 14-year-old children (Takeshita et al., 2002), and a linear decrease of N2 amplitude and an increase of P2 amplitude with increasing age from 8 to 17 years resulting in an adult morphology of the N1-P2 complex at about 14 – 17 years (Johnstone et al., 1996). In both, adults and children, auditory processing was most prominent at bilateral temporal channels reflecting generators in auditory cortices along the Sylvian fissure. This is in line with a study of Albrecht, Suchodoletz, & Uwer (2000) who found that both the child and adult AEP could be represented similarly by bilateral dipole sources located in the temporal lobe (superior temporal lobe and lateral temporal area). Consistently, neural generators of N100m, N250m (Takeshita et al., 2002) and concurrent acoustic objects were localised along the superior temporal plane in auditory cortices (i.e., primary auditory cortex) (Alain & McDonald, 2007; Dyson & Alain, 2004).

Group differences in AEF components

The most important differences between adults and children concern sound segregation based on *onset asynchrony*. First, in the *Delayed* condition children elicited an early segregation positivity (ESPm), whereas in adults no ESPm was elicited following the delayed harmonic. However, adults and children clearly showed an ESPm in the difference

waveform Mistuned – Baseline, as recently reported for adults (Lipp et al., 2010). In adults, the ESPm was interpreted to represent an automatic registration of the *mistuned* partial, as its amplitude is not affected by attention. Further support comes from a study of Alain and McDonald (2007), who reported an early positivity (P80) which has been evoked by a mistuned partial of a complex sound while participants ignored the stimuli and interpreted it as reflecting an early registration of inharmonicity in primary auditory cortex. However, the present results suggest that the ESPm in children seems to reflect an early, bottom-up driven, automatic process of concurrent sound segregation based on inharmonicity *and* onset asynchrony, as it was elicited in both hemispheres by a delayed and mistuned partial and was present irrespective of attention paid to the sounds or not. Therefore, in children the ESPm seems to be a more general marker for concurrent sound segregation and could imply an immature specificity for frequency periodicity compared to adults.

Second, in contrast to adults, children had no LSPm in the *Delayed* condition, indicating immature sound segregation processing of a temporal cue in children. Only a mistuned harmonic revealed a roughly adult-like late deflection in children. Therefore, the LSPm in children may reflect the activity of neurons specifically sensitive to frequency periodicity rather than represent concurrent sound segregation in general.

In both, adults and children, an ORNm was evoked by mistuning *and* delaying the third partial of a harmonic complex, supporting the interpretation that ORNm is a general marker of sound segregation in the brain signal (Lipp et al., 2010, for a review see Alain, 2007). Further, in adults the ORNm was significant in all conditions irrespective of attention and only missed statistical significance in children in the *Delayed* condition when they ignored the sounds. This is consistent with the proposal that concurrent sound segregation may occur independently of listeners` attention. However, in children the automation of concurrent sound segregation seems to be partly immature, especially when onset asynchrony serves as cue for concurrent sound segregation.

Different maturational patterns of concurrent sound segregation?

The findings indicate that concurrent sound segregation undergoes developmental changes. These changes seem to follow different maturational patterns depending on the type of cue (onset asynchrony, inharmonicity). Whereas, in children AEF morphology is adult-like for mistuning, it differs for delaying by (i) ESPm, which is *not* apparent in adults for *Delayed*

stimuli, (ii) ORNm, which is smaller in children and solely statistically significant when sounds are attended, and (iii) a missing LSPm in children compared to adults. Therefore, the present AEF morphology in children indicates an immature automatic segregation mechanism concerning onset asynchrony. In the immature auditory system incomplete myelination and synaptogenesis, affecting the latency and synchrony of the neural signal, was assumed to lead to slower cortical recovery and probably lower cortical excitability in children (Gilley, Sharma, Dorman, & Martin, 2005; Surwillo, 1981). Less synchrony, lower cortical excitability and/or less recruited neurons in children may have contributed to the present finding that children had smaller ORNm amplitudes than adults in the *Delayed* condition. This finding contradicts the results of Alain and colleagues (2003) who found larger ORN amplitudes following a mistuned harmonic in children than adults. However, a comparison has to be drawn with caution, because of noticeable differences between the present and Alain´s study concerning the kind of stimuli used (onset delay vs. 16% mistuning), stimulus duration, inter-stimulus interval, task and measurement method.

The suggestion that the detection of temporal and spectral changes in the acoustic signal may mature at different rates is supported by previous research. For example, the length of the temporal window of integration, defining the time window in which discrete sound elements are processed as a single event, was found to be shorter in adults (<200 ms) than 9–11-year-old children (<300 ms) and 5–8-year-old children (<350 ms) (Wang et al., 2005). This shows that temporal resolution is still immature in 9–11-year-old children, whereas, by age 9 frequency discrimination abilities are adult-like (Thompson, Cranford, & Hoyer, 1999). However, one has to keep in mind that maturation of complex sound patterns follows a longer developmental course than maturation of simple feature discrimination. For example, Sussman & Steinschneider (2009) demonstrated that 9 – 12-year-old children required a larger frequency separation than adults to perceive two streams. This is in line with the present finding that *also* auditory segregation of a *mistuned* harmonic is still not fully developed, regarding longer AEF latencies and poorer localization performance in children than adults. Accordingly, the suggestion of a deferred maturation of concurrent sound segregation based on onset asynchrony compared to sound segregation based on inharmonicity has to be put into perspective. As there is no difference in the localization performance between *delayed* and *mistuned* stimuli, children may use the information of the partly matured primitive grouping mechanism, reflected by the ESPm and ORNm, in a similar efficiency for both cues, yet less efficient than adults.

Peak latencies

When participants paid attention to the sounds, behavioural segregation performance across adults and children was the better, the shorter the peak latencies of ESPm, ORNm, and LSPm in the *Mistuned* and *Baseline* condition were. ESPm, ORNm, and LSPm in the *Mistuned* and ORNm in the *Delayed* condition occurred earlier in adults than in children and adults performed better than children localising the third harmonic. Therefore, the longer peak latencies of AEF components might be related with immature bottom-up processing in children. There is general agreement that latencies are longer in children than in adults at least for P1 and N1 (Albrecht et al., 2000; Johnstone et al., 1996; Kummer et al., 2007; Ponton, Eggermont, Kwong, & Don, 2000; Takeshita et al., 2002; Tonnquist-Uhlen, Borg, & Spens, 1995; Wang et al., 2005). Johnstone and colleagues (1996) investigated age-related changes of auditory event-related potentials in an auditory oddball task in participants ranging from 8 to 17 years. For target tones, they found a decrease in a behavioural reaction time measure, as well as a linear decrease in latency of the N1, N2 and P3 components with increasing age. Similarly, Wang and colleagues (2005) found that the latency of the mismatch negativity (MMN) decreased as age increased in participants ranging from 5 to 39 years. The latency of N170 (a component in the difference curve of monosyllabic speech and intensity-equivalent noise stimuli) indicating an electrophysiological correlate of speech perception, also decreased with increasing age, demonstrated for children from 9 to 14 years (Kummer et al., 2007). The maturation of the auditory system as reflected in the age-related latency decrease is assumed to indicate an increased transmission velocity due to increases in myelinisation, axonal diameter and synaptic efficiency (Tonnquist-Uhlen, 1996). Since longer latencies for AEFs, especially ORNm, were revealed in the *Mistuned* and *Delayed* condition, this suggests that in children bottom-up sound segregation is immature regardless of cue (temporal or spectral).

Influence of attention on sound segregation

Although markers for pre-attentive concurrent sound segregation, namely the ORN (Alain et al., 2001; Alain & Izenberg, 2003; Alain et al., 2002), P75m and ORNm (Lipp et al., 2010) have been identified and further evidence was found that attention is not always required for the formation of auditory streams (Sussman et al., 2007), it was suggested that there may be circumstances in which attention is needed to strengthen segregation or to resolve an ambiguous auditory scene. Attention may be needed to fine-tune segmentation

of a complex auditory scene (Cusack et al., 2004). Focused auditory attention was found to sharpen frequency tuning in human auditory cortex and to improve auditory performance in noisy environments by enhancing the processing of task-relevant stimuli and/or suppressing task-irrelevant information (Okamoto, Stracke, Wolters, Schmael & Pantev, 2007). For adults, the cues to evoke sound segregation in the present study may have been relatively strong (8% mistuning, 160 ms onset delay). Thus, attention may not have been needed to elicit AEFs (ESPm and ORNm), which therefore probably reflect primitive grouping mechanisms. In other situations, in which the cues for concurrent sound segregation are less strong or are ambiguous (e.g., in a noisy environment), attention may be beneficial to strengthen segregation. It is important to accentuate that under certain circumstances top-down processes may facilitate the process indexed by the ORN, e.g. when the experimental design encourages listeners to focus their attention on a particular frequency region. Accordingly, Alain et al. (2001) reported that the ORN in adults was found in active and passive listening conditions, but its amplitude was larger during active than passive listening (see experiments 1 and 3 from Alain et al., 2001). The present results indicate that in children automatic (bottom-up) sound segregation processes have not fully matured. The cues of the present study may fall into an ambiguous domain in which children need to focus attention to the sounds in order to support bottom-up segregation mechanisms. A possible "gain" for the ORNm due to attention could explain the present finding in children that *only* when the sounds were attended, a statistically significant ORNm followed the delayed harmonic. Focused attention probably helps children to partly compensate for their underdeveloped automatic organization of auditory processing. Further support for this hypothesis is provided by the finding of Sussman & Steinschneider (2009) who suggest a crucial role of attention for stream segregation in children. Larger frequency contrasts seemed to be required in *passive* than *active* listening conditions for children but not adults to elicit a MMN, the physiological index to indicate two auditory streams.

2.3 Concurrent sound segregation in children with and without attention deficit hyperactivity disorder (ADHD) (Study 3)

Abstract

Children with attention deficit hyperactivity disorder (ADHD) and children with auditory processing disorders constitute heterogeneous groups presenting strikingly similar symptoms. It is unclear whether children with ADHD have a genuine perceptual deficit or merely an attention impairment leading to perceptual deficits. This issue was addressed by investigating auditory segregation skills in children with and without ADHD. Auditory evoked fields (AEFs) were measured using magnetoencephalography (MEG). Children were presented with harmonic complex tones with (i) all harmonics in-tune, (ii) the third harmonic mistuned by 8% of its original value, or (iii) the onset of the third harmonic delayed by 160 ms compared to the other harmonics. Active and passive listening conditions were compared to evaluate the contribution of attention on sound segregation. Both cues - inharmonicity and onset asynchrony - elicited sound segregation in both groups of children. However ADHD children performed worse than control children. We found an early segregation positivity (ESPm), probably reflecting a stimulus-driven, automatic process of concurrent sound segregation based on inharmonicity and onset asynchrony. ESPm in ADHD children seems less consistent and more variable when the third partial was mistuned or delayed compared to the healthy control group. Results also suggest that, especially when inharmonicity serves as cue for concurrent sound segregation, the ESPm may indicate a genuine auditory processing deficit in children with ADHD. The present results support the idea, that at least a subgroup of children with ADHD has some impairment in pre-attentive processing of sounds.

2.3.1 Introduction (Study 3)

Attention deficit hyperactivity disorder (ADHD) is an early-onset neurobehavioral disorder estimated to affect about 5%−10% of children with an estimated heritability of 0.76 (Biederman & Faraone, 2005; Scahill & Schwab-Stone, 2000). ADHD is characterized by behavioural symptoms of inattention, hyperactivity and impulsivity. It is associated with an increased incidence of other psychiatric problems, such as anxiety, oppositional defiant, learning, conduct and mood disorders (Biederman, 2005; Pliszka, 2000). Following the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV, APA, 1994) ADHD can

be divided into three subtypes: the predominantly inattentive subtype (ADHD-I), the predominantly hyperactive/impulsive subtype (ADHD-H) and the combined subtype (ADHD-C). Originally, all three subtypes were thought to be different manifestations of the same disorder.

The current view is that subtypes of ADHD or alternative subgroups (e.g. characterized by deficient inhibitory control or delay aversion) might not have the same etiology (Nigg, 2005; Sonuga-Barke, 2005). In his model of executive functioning deficits in ADHD, Barkley (2006; Barkley, DuPaul, & McMurray, 1990) discusses the subtypes ADHD-H and ADHD-C, which may be developmental stages of the same disorder and concludes that ADHD-I may be a distinct disorder rather than a subtype of a common attention deficit. Moreover, in addition to the different ADHD subtypes likely having different etiologies, symptoms within one subtype are heterogeneous. Therefore, the view has evolved that ADHD is a complex, multifactoral disorder (Nigg, 2005; Sonuga-Barke, 2005).

Emerging evidence documents the strong influence of fronto-subcortical circuits, which mediate executive abilities such as working memory, attention regulation, behavioral inhibition, planning, and set-shifting. Structural and functional imaging studies consistently suggest ADHD is associated with dysfunctions in fronto-subcortical pathways and imbalances in dopaminergic and noradrenergic systems, probably resulting in core deficits like failure of inhibitory control, dysregulation of brain systems mediating reward and response cost, and deficits in arousal, activation, and effortful control (Biederman, 2005; Biederman & Faraone, 2005; di Michele, Prichep, John, & Chabot, 2005). However, although executive functioning deficits and deficits in reward signaling are a much-replicated finding in many experiments, a meta-analysis by Willcutt et al. (2005) revealed that executive functioning deficits are neither necessary nor sufficient to cause all cases of ADHD. For instance, Nigg et. al (2005) have demonstrated that fewer than 50% of children with ADHD-C exhibit impairment on any specific executive function tasks. Furthermore, deficits in reward signaling alone, indicated by measures of delay aversion, showed modest validity to discriminate ADHD children from control children (Solanto et al., 2001). Simultaneous analyses of both domains – executive functioning and delay aversion – in a preschool study (Sonuga-Barke, Dalen, & Remington, 2003) revealed that 29% of ADHD-range children displayed both delay aversion and executive dysfunction,

27% delay aversion only, 15% executive dysfunction only, and 29% neither problem (Nigg et al., 2005). Thus, although this biologically based model of a fronto-subcortical dysfunction appears compelling, it is clear that there must other etiological factors that cause ADHD symptoms in at least a subgroup of children.

ADHD, with or without hyperactivity, still remains one of the most controversial issues in child psychiatry, especially in the endeavour to clarify the relationship between this disorder and (central) auditory processing disorder ((C)APD) (Bamiou, Musiek, & Luxon, 2001; Cacace & McFarland, 2005). A heterogeneous population of children may be subsumed under the denominations of ADHD-I, ADHD-C and (C)APD (Barkley, 2006; Cacace & McFarland, 2005), as behavioural deficits associated with ADHD-like inability to sustain focused attention (Barkley, Grodzinsky, & DuPaul, 1992), deficits in behavioral inhibition and reduced working memory (Barkley, 1997) could potentially affect performance on perceptual tasks. However, it is also possible that listening difficulties lead to ADHD symptoms, as children might be less able to focus their attention on one particular sound source when other sounds are present. Behaviourally, it is impossible to determine, which deficit (auditory processing or attention deficit) precedes the other. Thus, debate continues over the hypothesis that children with ADHD have a genuine auditory processing deficit.

Recent findings indicate that ADHD is a significant factor in children´s psychoacoustic performance. For example, the presence of ADHD resulted in a general reduction of performance in the perception of auditory temporal and non-temporal cues (Breier, Fletcher, Foorman, Klaas, & Gray, 2003). Children having attention deficit disorder (ADD) showed poorer speech discrimination abilities than control children when background noise was introduced (Geffner, Lucker, & Koch, 1996), and children with attention deficits asked to judge the loudness of speech as comfortable or tolerable, preferred significantly lower levels than control children, implying a dysfunction in suprathreshold loudness perception (Lucker, Geffner, & Koch, 1996). Ptok and colleagues (2006) analysed questionnaires for ADHD and (C)APD symtomatology of 312 children and showed that children with malfunctioning attention had significantly greater deficits in understanding speech in background noise, sound localization, and auditory memory than children with no attentional impairments.

The overlapping symptomatology with (C)APD was particularly marked for the ADHD-I subtype. Recent research suggests that these two disorders are often co-morbid, but can occur independently (Keller & Tillery, 2002; Tillery, Katz, & Keller, 2000). They have in common academic difficulties, distraction, poor listening skills, asking for things to be repeated, auditory divided attention deficit, difficulty hearing in background/ambient noise (Chermak, Tucker, & Seikel, 2002), and inattention (Chermak, Hall, & Musiek, 1999; Chermak, Somers, & Seikel, 1998). Despite the overlap between disorders, Chermak et al. (2002) emphasize that in ADHD-I the deficits may reflect a cognitive disorder involving executive dysfunction (i.e., dysfunction in general control processes that regulate behavior). In contrast, in (C)APD these difficulties are due to a perceptual disorder which causes deficits in processing information through the auditory modality. Similarly, Sutcliffe et al. (2006) concluded that poor performance of children with ADHD on tasks that involve discrimination of brief auditory stimuli may be due to problems in temporal synchronization of attention, rather than genuine inability to hear differences.

Children with attention disorders, as well as children with auditory processing disorders, constitute a etiologically heterogeneous group presenting strikingly similar symptoms. Therefore, the challenge for accurate differential diagnosis is to measure modality-specific perceptual dysfunctions (Cacace & McFarland, 1998). Identifying a biological measure that could aid in this distinction would help to refine diagnostic criteria and may provide more specific diagnostic tests for ADHD and (C)APD.

With this motivation, the current study compared auditory segregation skills in children with and without ADHD, controlling for attention. In everyday circumstances, e.g. in the classroom, children are faced with the perceptual challenge of identifying and parsing the acoustic elements that belong to a particular sound source (e.g., teacher) from other co-occuring auditory sources (e.g., other pupils talking), often when the background sources are relatively intense and when the environment is reverberant. The ability to assign incoming acoustic elements to perceptual auditory objects corresponding to different physical sound sources is known as auditory scene analysis (ASA) (Bregman, 1990). In his model of ASA, Bregman (1990) proposes two classes of mechanisms that help to decide which components belong to a particular sound source. An initial pre-attentive process partitions the acoustic wave into distinct groups of sounds according to Gestalt principles, such as grouping by physical similarity, temporal proximity, and good continuity. Sounds

are more likely to be assigned to separate sources if they differ widely in frequency, intensity, onset time, and spatial location. This stage is thought to be largely independent of listeners` attention being driven by the incoming acoustic data (bottom-up). The outcome of these "primitive grouping mechanisms" may then be subjected to a more detailed analysis by controlled processes (top-down). Whereas primitive grouping mechanisms group sounds on the basis of similarity in physical stimulus properties such as frequency, intensity, phase, harmonicity, and temporal synchrony, "schema-governed mechanisms" are thought to be learnt, and to use prior knowledge to group auditory objects in the acoustic data, which is particularly useful in listening situations with unfavorable signal-to-noise levels. It should be considered that stimulus-driven sensory mechanisms can be modified under certain circumstances by attentive processes (Carlyon, Cusack, Foxton, & Robertson, 2001; Sussman, Winkler, Huotilainen, Ritter, & Naatanen, 2002), suggesting that top-down and bottom-up effects of sound organization may interact at an early stage of auditory processing.

It is reasonable to expect that difficulties in processing one or more of the acoustic grouping cues (e.g., onset synchrony and harmonicity) could impair a listener`s ability to segregate speech from the background sound, and lead to reductions in the intelligibility of speech, particularly when listening in a noisy environment like a classroom. Investigating the ability of children with ADHD to segregate auditory objects based on such cues as inharmonicity and onset asynchrony should lead to a better understanding of the neural underpinning of auditory processing dysfunctions in these children.

Alain and colleagues investigated sound segregation in a series of experiments measuring the electroencephalogram (EEG) of adults (Alain et al., 2001, 2002; Alain & Izenberg, 2003) and children (Alain, Theunissen, Chevalier, Batty, & Taylor, 2003), while they were presented with auditory stimuli. The stimuli either promoted sound segregation leading to the perception of two sounds rather than one (Segregation condition), or did not promote sound segregation leading to the perception of one sound (No Segregation condition). The presented sounds comprised multiple harmonics of 200 Hz, one of which could be mistuned so that it was no longer an integer multiple of the fundamental. In the Segregation condition the third harmonic was mistuned causing the third harmonic to stand out perceptually from the sound complex formed by the other harmonics. The authors identified a negative deflection in the EEG difference waveform (Segregation – No

Segregation) approximately 180 ms after stimulus onset, which they referred to as the object-related negativity (ORN) and another later, positive deflection around 400 ms, labelled P400. The ORN was present when participants actively listened to the sounds and also when their attention was distracted by reading a book (Alain et al., 2001) or watching a silent movie (Alain et al., 2002). The authors concluded that the ORN reflects automatic, bottom-up sound segregation and that its generation is minimally affected by attentional load (Alain & Izenberg, 2003). In contrast, the P400 was only present in the active listening condition. The authors therefore concluded that the P400 reflects awareness-related top-down processes of sound segregation. In children the EEG difference wave showed a different pattern: mistuning elicited a clear ORN, but there was no significant effect of mistuning on ERPs 300-400 ms post-stimulus (Alain et al., 2003). The authors concluded that perceptual segregation of concurrent sounds by children may be adult-like, but children are less efficient than adults in processing the information following the detection of mistuning.

Lipp and colleagues (2010) measured auditory evoked fields (AEFs) using magnetoencephalography. Participants were presented with complex sounds in witch: (i) all harmonics were in-tune (= *Baseline*), (ii) the third harmonic was mistuned by 8% of its original value, or (iii) the onset of the third harmonic was delayed by 160ms compared to the other harmonics. During MEG recording, participants listened to the sounds and performed an auditory localisation task involving the third harmonic, whereas in another session they ignored the sounds and performed a visual localisation task. Active and passive listening was chosen to evaluate the contribution of attention on sound segregation. In the difference waveforms (*Mistuned – Baseline*) an early deflection at ~75ms (P75m) after sound onset was found, probably reflecting an automatic registration of the mistuned harmonic. Subsequent deflections — the object-related negativity (ORNm) and a later displacement (P230m) — seem to be more general markers of concurrent sound segregation, as they were elicited by both mistuning and delaying the third harmonic. Results indicate that the ORNm reflects relatively automatic, bottom-up sound segregation processes, whereas the P230m may be more sensitive to attention, especially with inharmonicity as the cue for concurrent sound segregation. Similar displacements reported by Alain and McDonald (2007) have been evoked by a mistuned third partial of a complex sound while participants were not required to pay attention to the stimuli. The early positivity was not correlated with perceptual judgement and was interpreted as reflecting an

early registration of inharmonicity in primary auditory cortex. The authors concluded that all three components likely reflect automatic sound segregation that may occur independently of listeners´ attention. However, attention was not manipulated, making it difficult to determine whether the early P80, ORN, and P230 are modulated by attention.

The present study investigated neural processes involved in concurrent sound segregation based on primitive grouping mechanisms in children with and without ADHD. Two distinct cues — inharmonicity and onset asynchrony — were used to invoke sound segregation (Darwin & Carlyon, 1995; Lipp et al., 2010). To ensure that participants based their judgment on sound segregation rather than other cues, we used a sound localisation task that could only be performed correctly if participants truly heard out the segregated partial, which was presented either left or right from straight ahead (*Attend* condition). In order to investigate the susceptibility of the *early segregation positivity* (ESPm/P75m), the ORNm and the *late segregation positivity* (LSPm/P230m) to attentional focus, an additional task was included that distracted the participants' attention away from the sounds, but kept the over-all attentional load comparable (*Non-Attend* condition).

The study aimed to investigate differences in the process of sound segregation between children with and without ADHD, and if potentially differences are affected by attention. We hypothesized that, if children with ADHD have no genuine auditory processing deficit concerning concurrent sound segregation, they should not differ from control children in AEF components, which potentially reflect an automatic, bottom-up concurrent sound segregation process.

2.3.2 Methods (Study 3)

Participants

Fifteen control children (mean age 11.0 years, ±0.3, range 9-13 years, 5 female, 13 right-handed, 2 bimanual) and fifteen children who were diagnosed with ADHD (mean age 10.3 years, ±0.4, range 9-13 years, 6 female, 14 right-handed, 1 left-handed) took part in the experiment. There was no significant group difference in age ($F(1,28)=2.6$, $p=.12$). Eleven children were diagnosed with ADHD combined type (ADHD-C, DSM-IV code 314.01), four with ADHD predominantly inattentive type (ADHD-I, DSM-IV code 314.00). Diagnoses were made by the head psychiatrist and his team of psychologists at two paediatric out-patient centres in Konstanz and Friedrichshafen. Diagnoses were based on

standardised parent and teacher questionnaires, diagnostic interviews and psychometric tests. Six ADHD children were currently under medical treatment with methylphenidate (one child took Equasym® 20 mg; three took Medikinet® 20 mg, 25 mg, 40 mg; two took Concerta® 27 mg, 36 mg). They were asked to omit medication for at least 48 h before testing. 60% of ADHD children had co-morbid disorders (e.g., specific reading disorder and developmental disorder of motor function). All parents gave their written informed consent to participate. Ethical approval of the study was granted by the Ethical Review Board of the University of Konstanz.

MEG recordings

In session one, hearing thresholds were determined and children were familiarised with the experiment. In session two, MEG data were collected in two conditions: a) children listened to sounds and performed an auditory localisation task (*Attend* condition), and b) children were played the same sounds while performing a visual localisation task (*Non-attend* condition). *Attend* and *Non-Attend* conditions were presented in a counterbalanced order.

Hearing thresholds, attention test, and questionnaires

Hearing thresholds were determined for five frequencies (200 Hz, 595 Hz, 1 kHz, 2 kHz and 4 kHz) with clinical audiometer (AC40). One subtest of the computerised test battery *Test of Attention for Children* (KITAP), testing sustained attention, was administered (Zimmermann, Gondan, & Fimm, 2002). Handedness was assessed with the Edinburgh handedness inventory (Oldfield, 1971). In order to screen for any maladaptive behavioral or emotional problems, all children were evaluated with the *Child Behavior Checklist* (CBCL) (Doepfner et al., 1998). ADHD symptomatology was evaluated with the *Diagnostic System of Psychiatric Disorders in Children and Adolescents* (DISYPS) (Doepfner, Goertz-Dorten, & Lehmkuhl, 2008), using the questionnaire for ADHD (FFB-ADHD). To screen for abnormalities concerning hearing all children were evaluated with the *Questionnaire on Auditory Development* (AUDIVA, 2004; Rosenkötter, 1998).

Acoustic Stimuli

Stimuli were created at the Department of Psychology of the University of York, UK (Summerfield, Paul, Kitterick, Briley, & Bailey, 2007). *Baseline* stimuli were complex tones consisting of 24 sinusoidal partials (f_0 = 200 Hz), which were synthesised digitally

(44100 samples/s, 16-bit amplitude quantization) with equal amplitudes. Two variants of the *Baseline* stimulus were formed by perturbing the third partial at 600 Hz. In *Mistuned* stimuli, the frequency of the third partial was raised by 8% to 648 Hz. In *Delayed* stimuli, the third partial started 160 ms after the other partials, but ended simultaneously (Fig. 1). One hundred samples of each stimulus condition (*Baseline, Mistuned, Delayed*) were created; stimuli lasted 500 ms.

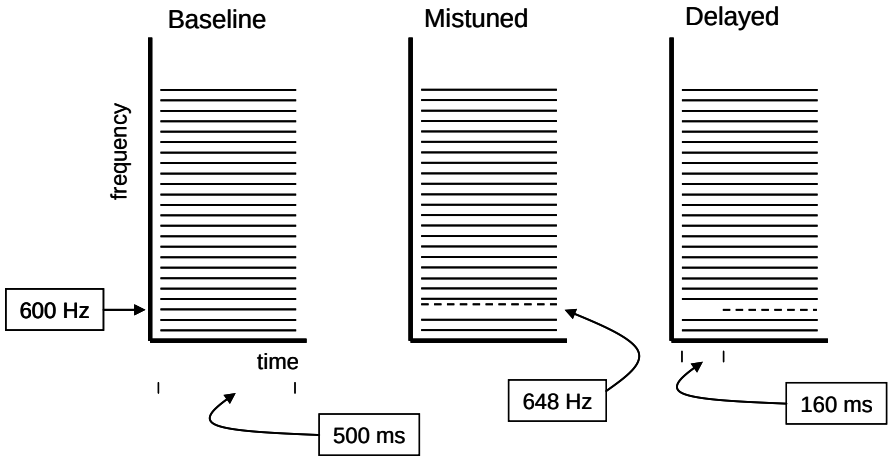

Figure 1: Schematic spectrograms of Baseline, Mistuned and Delayed stimuli.

Stimuli were created so that the third partial was always perceived 45° left or right of straight ahead, with equal probability. For specifics of stimulus generation, please see (Lipp et al., 2010).

In order to familiarise the children with the experiment, a training session preceded the MEG-recordings. Stimuli were presented through headphones (Dynamic Stereo Headphones MDR-CD470, Sony), while children performed the localisation task (see below). For the MEG recordings stimuli were presented through tube-phones (Etymotic Research, ER30). For each ear, a 6.1-m length of 4-mm diameter tube terminated in a right-angle adapter (ER30-9) ending with a disposable foam eartip (ER13-14B). A digital filter was designed to compensate for the low-pass frequency response of this sound delivery system, resulting in a frequency response that was flat to within -4–0 dB from 100–4800 Hz when measured with a B&K Ear Simulator (Type 4157) with 0.5-inch microphone (Type 4134), Microphone Preamplifier (ZC 0032) and B&K Hand-held Analyzer (Type 2250).

Using the same measurement system, the presentation level of *Baseline* stimuli was set to 70 dB (A).

Task

On each trial of the Attend condition, children reported whether they heard the third partial on the left or right using two keys on a keyboard. Feedback on the accuracy of the response was given by displaying a coloured rectangle on the side where the key had been pressed. It was green if the response was correct or red if the response was incorrect. The inter-stimulus interval (ISI) was the response latency plus 500 ms.

On each trial of the *Non-attend* condition, a faint grey circle subtending a visual angle of 1° appeared for 50 ms at a randomly-chosen location within 8° right or left of the fixation cross. Children were instructed to report whether the circle appeared left or right of the fixation cross by making a corresponding key press. Feedback was provided in the same format as used in the *Attend* condition. The ISI was the response latency plus a random value in the range from 500 to 1500 ms. At the same time, the acoustic stimuli were presented asynchronously with an ISI that was chosen randomly from the range 500 to 1500 ms. Children were instructed to concentrate on the visual stimuli and to ignore acoustic stimuli.

MEG recordings

Per attention condition (*Attend, Non-attend*) one hundred samples of each stimulus type (*Baseline, Mistuned, Delayed*) were presented randomly. Recording was done with a 148-channel magnetometer (MAGNESTM 2500 WH, 4D Neuroimaging, San Diego, USA). Acoustic stimuli were delivered with an analog-to-digital converter (Motu 2408) and amplifier (Servo 200, Samson) through tube-phones (Etymotic Research, ER30). The foam tips were introduced carefully into the ear canals of children and were additionally secured with medical tape. Children lay on a height-adjustable bed in a magnetically shielded room (Vakuumschmelze Hanau). Eye movements (EOG) were recorded from four electrodes attached to the left and right temple and above and below the right eye, as well as cardiac activity (ECG) via two electrodes, one on each forearm. A Neurofax amplifier (EEG-11006, Nihon Kohden) served for the recording of EOG and ECG. A video camera installed inside the chamber allowed monitoring the subject's behaviour and compliance throughout the experiment. Visual stimuli and feedback were projected (D-ILA Projector,

JVC™, DLA-G11E) onto a screen placed ~40cm away from the participants' eyes through a mirror system. Data from each channel were band-pass filtered between 0.1 and 200 Hz and sampled at a rate of 678.17 Hz. Recording was continuous.

MEG data analysis

Data were corrected for ocular and cardiac artefacts using BESA™ 5.3 and were averaged separately for *Baseline, Delayed,* and *Mistuned* stimuli in *Attend* and *Non-attend* conditions using a pre-stimulus baseline of 100 ms and a post-stimulus time-window of 800 ms. Epochs were excluded if the signal amplitude in any channel exceeded 3.5 pT or if the signal amplitude differed by more than 2.5 pT between adjacent sampling points.

The sensor groups of interest were two groups of 7 sensors located bilaterally over the temporal lobes, which showed the biggest N2m amplitudes in the averaged cortical response elicited by *Baseline* stimuli (Fig. 2). This approach was chosen, as it was objective and ensured that sensor groups reflected auditory processing (the auditory N2 component is very prominent in children) (Johnstone, Barry, Anderson, & Coyle, 1996; Sussman, Steinschneider, Gumenyuk, Grushko, & Lawson, 2008).

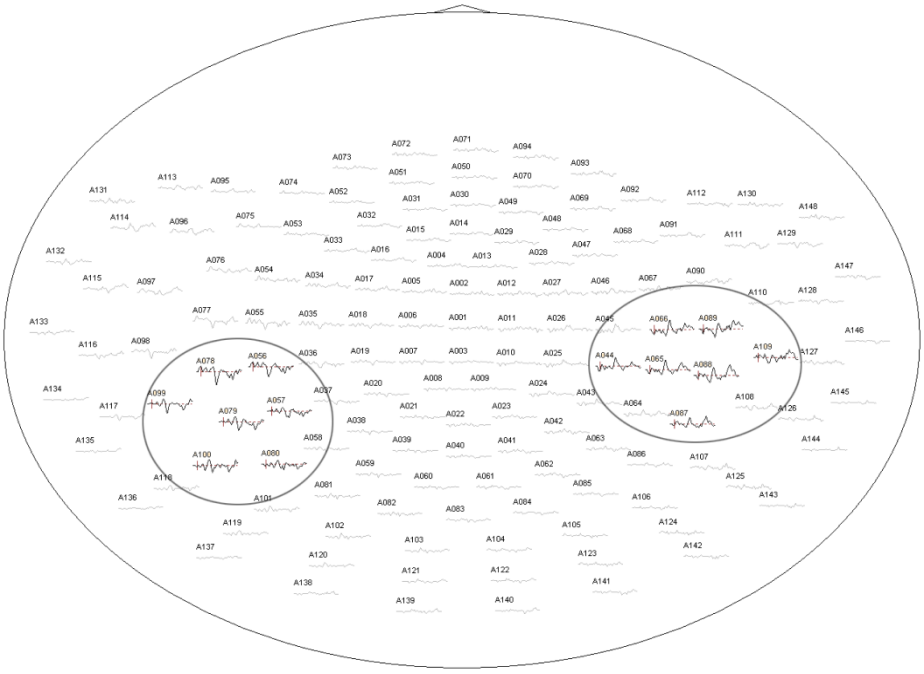

Figure 2: Difference waveforms (Delayed - Baseline) for each sensor averaged over 15 control children in the Attend condition. Channel Groups used for data analysis are coloured black (7 channels for each hemisphere). Left hemisphere: A056, A057, A078, A079, A080, A099, and A100. Right hemisphere: A044, A065, A066, A087, A088, A089, and A109.

In both groups of children, the same 7 sensors showed the biggest N2 amplitudes in each hemisphere. The left channel group covered the in-going magnetic field (positive values) and the right channel group the out-going magnetic field (negative values) of the underlying auditory sources. Therefore, the sign of the signals over the left hemisphere was reversed to facilitate comparisons with right-hemisphere signals.

Mean amplitudes in time-windows of interest were computed for each sensor group, participant, stimulus, and condition. Time-windows were chosen to embrace prominent deflections from zero in either direction in the difference waveforms (*Mistuned minus Baseline* and *Delayed minus Baseline*). Paired t-tests and effect sizes (r) were computed for all relevant contrasts using the following formulae:

$$\text{Cohen's } d = M_1 - M_2 \,/\, \sigma_{pooled}$$

$$\text{where } \sigma_{pooled} = \sqrt{[(\sigma_1^2 + \sigma_2^2) / 2]}$$

$$r_{Y\lambda} = d \,/\, \sqrt{(d^2 + 4)}$$

The differences of the mean amplitudes (*Mistuned – Baseline*, *Delayed – Baseline*) characterising the components (ESPm, ORNm, LSPm) were analysed statistically with ANOVAs. Hemisphere[6] (*Left, Right*) and Attention (*Attend, Non-attend*) were within-group variables. Group *(children with ADHD, control children)* was used as between group factor.

To specify individual peak latencies, individual difference waveforms (*Mistuned–Baseline* and *Delayed–Baseline*) for both conditions (*Attend, Non-attend*) and both hemispheres (*Left, Right*) were band-pass filtered from 1 to 25 Hz for detecting ESPm and ORNm, and from 1 to 8 Hz for detecting LSPm. Peak latencies were compared between conditions with ANOVAs for each component (ESPm, ORNm, LSPm). Hemisphere (*Left, Right*) and Attention (*Attend, Non-attend*) were within-group variables. Group *(children with ADHD, control children)* was used as between group factor.

[6] Except for ESPm in *Mistuned* condition. Left and right hemispheric values (e.g., differences of the mean amplitudes or peak latencies) were averaged together for ANOVAs, because for children with ADHD no ESPm on the left side in the *Attend* session could be identified.

Behavioural data

The percentages of correct responses to *Baseline, Mistuned,* and *Delayed* stimuli were compared in a multivariate analysis of variance (MANOVA). Condition (*Baseline, Mistuned, Delayed*) was within-group variable. Group *(children with ADHD, control children)* was used as between group factor. Further, Spearman´s rank correlations were calculated between behavioural results in MEG and attention test scores (KITAP), as well as amplitudes/latencies of MEG components (ESPm, ORNm, LSPm). P-values were corrected for multiple tests using the Bonferroni-Holm method. Hearing thresholds and scores of questionnaires (DISYPS, AUDIVA, CBCL) were compared between groups with t-tests. Attention test scores (KITAP) were compared between groups with Mann-Whitney U-tests, as scores were not normally distributed. *Post-hoc* analyses were performed with Tukey's HSD test. All analyses were conducted with Statistica (StatSoft, Inc., Version 6, 2003).

2.3.3 Results (Study 3)

Hearing thresholds

All participants had normal hearing levels (≤20 dB HL) for both ears and all tested frequencies (200 Hz, 595 Hz, 1 kHz, 2 kHz and 4 kHz). Mean hearing levels were lower than 8 dB HL in both ears. T-tests showed no statistically significant differences between children with ADHD and control children.

Questionnaire scores and attention test

Results of the *Child Behavior Checklist* (CBCL) showed that children with ADHD had higher scores than control children on all 1st order syndrome scales except *somatic complaints.* ADHD children also reached higher scores on all 2nd order syndrome scales (Table 1).

Table 1: Group comparison of CBCL syndrome scales and number of children with clinically relevant T-scores for syndromes.

CBCL	**Control children** (n=15)			**ADHD** (n=15)			**t**	**p**
	Mean (T-score)	SD	No. of children (clinically relevant)	Mean (T-score)	SD	No. of children (clinically relevant)		
Syndrome Scales (1st order)								
Social withdrawal	52.0	3.0	0	58.0	6.1	0	-3.2	**.003**
Somatic complaints	56.4	7.9	**2**	57.4	8.0	**1**	-0.3	.73
Anxious/ depressed	53.7	5.6	0	65.4	11.9	**3**	-3.4	**.002**
Social problems	52.5	4.8	0	62.7	10.1	**1**	-3.5	**.001**
Schizoid/ compulsive	50.9	2.5	0	56.7	6.4	**1**	-3.3	**.003**
Attention problems	50.4	0.8	0	69.3	8.9	**7**	-8.2	**<.001**
Antisocial behavior	52.5	3.4	0	60.7	7.8	**1**	-3.8	**<.001**
Aggressive behavior	52.0	2.9	0	70.9	12.0	**6**	-5.9	**<.001**
Syndrome Scales (2nd order)								
Internalizing	51.7	7.1	0	61.8	9.2	**6**	-3.4	**.002**
Externalising	47.8	7.8	0	61.8	9.5	**11**	-5.8	**<.0001**
Total problems	48.7	7.4	0	67.1	8.2	**11**	-6.4	**<.0001**

SD=standard deviation, 1st order Syndrome Scales: clinically relevant T-scores > 70, 2nd order Syndrome Scales: clinically relevant T-scores > 63, Internalizing problems=social withdrawal, somatic problems, and anxious/depressed, Externalizing problems=social problems, schizoid/compulsive, attentional problems, antisocial behavior, and aggressive behavior

Results of the questionnaire on auditory development (AUDIVA) showed higher scores for ADHD children compared to control children in questions concerning hearing ($t(28)=-3.4$, $p=.002$) and concentration ($t(28)=-10.1$, $p<.0001$), indicating more aversion against noise, more problems in conversations, and more problems to concentrate.

Table 2 shows the results of the questionnaire on ADHD symptoms (DISYPS, FFB-ADHD). ADHD children had higher age-normalized stanine scores (M=5, SD=2) on all scales. One control child scored above the clinically relevant cut-off for impulsivity (SN=8).

Finally, children with ADHD made more errors than controls in the subtest of KITAP for sustained attention ($p=.003$; $Z=-3.0$).

Table 2: Results of ADHD questionnaire (FBB-ADHD) from Diagnostic System of Psychiatric Disorders in Children and Adolescents (DISYPS). Stanine scores for attention deficit, hyperactivity, impulsivity, and ADHD symptomatology in total for ADHD children and control children. Number of ADHD children and control children with clinically relevant cut-off scores for each syndrome scale.

ADHD symptoms	**Control children** (n=15)			**ADHD** (n=15)			**t**	**p**
	Mean (SN-score)	SD	No. of children (clinically relevant)	Mean (SN-score)	SD	No. of children (clinically relevant)		
Syndrome Scales								
Attention deficit	3.7	1.8	0	7.7	1.1	**8**	-7.4	**<.0001**
Hyperactivity	4.9	1.1	0	7.8	1.0	**8**	-7.5	**<.0001**
Imulsivity	4.9	1.2	**1**	8.2	1.1	**13**	-7.7	**<.0001**
$Total_{ADHD}$	4.0	1.6	0	8.0	0.8	**10**	-8.4	**<.0001**

SN=stanine, SD=standard deviation, clinically relevant SN scores ≥8

Behavioural results

Both control children and children with ADHD performed well above chance localising the target partial in the Delayed and Mistuned condition, indicating that they segregated it successfully from the harmonic background (main effect Condition ($F(2,27)=45.25$, $p<.0001$; *Delayed>Baseline* p=.0001, *Mistuned>Baseline* p=.0001, Fig. 3). Overall, ADHD children performed worse than control children ($F(1,28)=4.78$, p=.037). No other significant main effects or interactions were found.

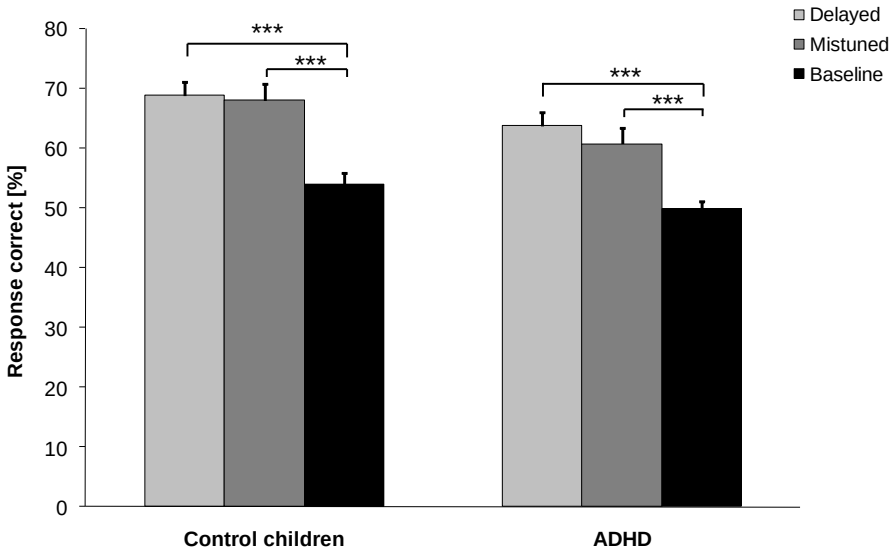

Figure 3: Localisation performance of ADHD children and control children during the MEG scan for Baseline, Mistuned, and Delayed stimuli. Error bars indicate standard errors.

Mistuning the third harmonic

Figure 4 and 5 show the difference waveforms *Mistuned* minus *Baseline* for ADHD children and control children. The difference waveforms were characterised by distinct components peaking at ~87 ms (ESPm) and ~153 ms (ORNm) (also see Table 3). This likely corresponds to the early positivity (~80 ms), and the object-related negativity (ORN, ~160 ms) originally identified by Alain and colleagues (2007) and to P75m and ORNm (~132 ms) found in MEG (Lipp et al., 2010). The ORNm was followed by a broader component that peaked at ~320 ms for controls and ~356 ms for ADHD children (Fig. 4 and Table 3). A similar component was also found in adults by Alain and colleagues, which was labelled P400 (Alain et al., 2001), P230 (Alain & McDonald, 2007) or P230m (Lipp et al., 2010). Here is it referred to as the late segregation positivity (LSPm).

Table 3: Peak latencies for components ESPm, ORNm and LSPm in attended and non-attended conditions for left- and right-hemispheric channel groups averaged over all ADHD children and control children. Time-ranges used for calculating mean amplitudes are displayed.

Peak Latencies and Time-Ranges				**Control children** (n=15)			**ADHD** (n=15)		
Component	**Comparison**	**Attention**	**Hemisphere**	**Peak Latency (ms)**	**Time-Range (ms)**	**Mean Peak Latency (ms)**	**Peak Latency (ms)**	**Time-Range (ms)**	**Mean Peak Latency (ms)**
ESPm	*Mistuned* vs. *Baseline*	Attend	Left	84	69-99	87	----	----	88
			Right	89	74-104		90	75-105	
		Non-attend	Left	86	71-101		88	73-103	
			Right	89	74-104		87	72-102	
	Delayed vs. *Baseline*	Attend	Left	230	215-245	229	207	192-222	229
			Right	226	211-241		215	200-230	
		Non-attend	Left	230	215-245		249	234-264	
			Right	231	216-246		245	230-260	
ORNm	*Mistuned* vs. *Baseline*	Attend	Left	153	133-173	153	146	126-166	153
			Right	168	148-188		153	133-173	
		Non-attend	Left	147	127-167		164	144-184	
			Right	145	125-165		147	127-167	
	Delayed vs. *Baseline*	Attend	Left	298	278-318	300	314	294-334	319
			Right	298	278-318		333	313-353	
		Non-attend	Left	292	272-312		326	306-346	
			Right	311	291-331		304	284-324	
LSPm	*Mistuned* vs. *Baseline*	Attend	Left	328	293-363	320	350	315-385	356
			Right	339	304-374		353	318-388	
		Non-attend	Left	321	286-356		361	326-396	
			Right	291	256-326		359	324-394	

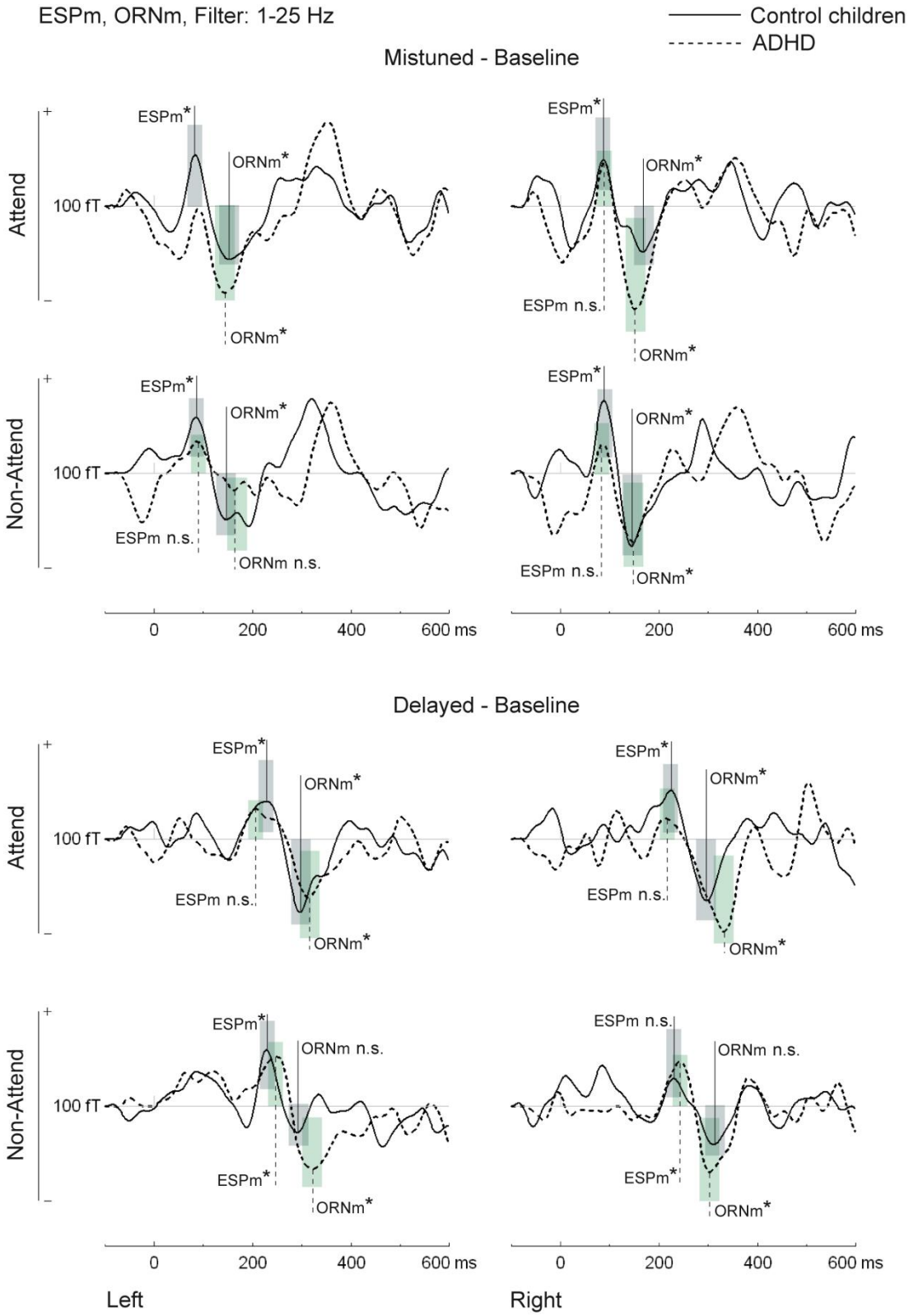

Figure 4: Difference waveforms for components ESPm and ORNm in attended and non-attended conditions for left and right hemispheric channel groups averaged over all participants (15) of each group of children. Difference waveforms (Mistuned – Baseline) are displayed in the upper four graphs; difference waveforms (Delayed – Baseline) are displayed in the lower four graphs. The polarity of the signal is reversed in the left hemisphere. Coloured boxes emblematize time-windows used for statistical analyses. Signals were band-pass filtered from 1 to 25 Hz.

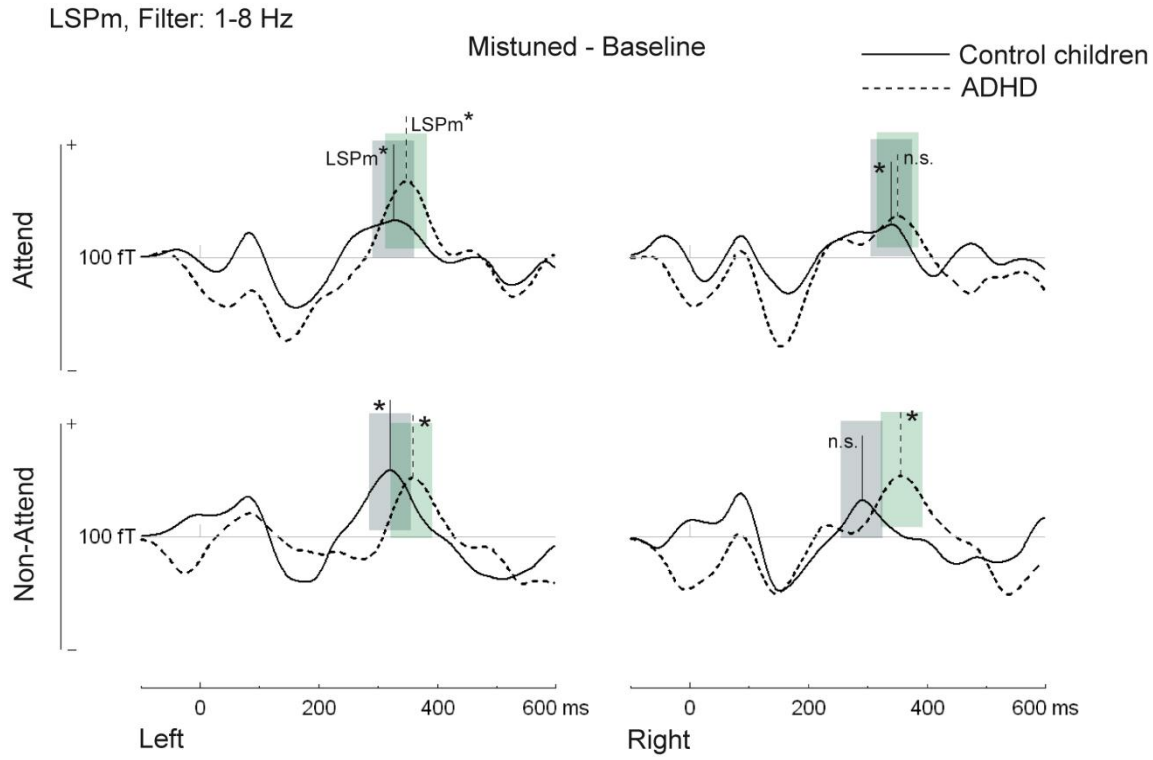

Figure 5: Difference waveforms (Mistuned – Baseline) for component LSPm in attended and non-attended conditions for left and right hemispheric channel groups averaged over all participants (15) of each group of children. The polarity of the signal is reversed in the left hemisphere. Coloured boxes emblematize time-windows used for statistical analyses. Signals were band-pass filtered from 1 to 8 Hz.

Mean amplitudes in the two temporal channel groups were computed for attended and non-attended *Baseline* and *Mistuned* conditions in specific time-ranges. ESPm time-ranges covered ±15 ms around the ESPm peak latency. Mean amplitudes in the ORNm time-ranges covered ±20 ms around the ORNm peak latency, and for LSPm time-ranges covered ±35 ms around the LSPm peak latency (Table 3). Mean amplitudes were retrieved per participant (Table 4). In ADHD children ESPm could not be detected in the left hemisphere when children attended to the sounds in *Mistuned* condition.

Detailed results for each group are summarized in Table 5. The most important difference between ADHD children and control children concerned the early segregation positivity. Whereas in control children all components (ESPm, ORNm, and LSPm) were significant with only one exception (LSPm, *Non-attend*, right hemisphere, p=0.1), all early positivities (ESPm) were non-significant in ADHD children irrespective of attention and hemisphere. Further, the ORNm failed to reach significance in the left hemisphere (p=0.6)

when ADHD children did not attend to the sounds; the LSPm was not significant in the right hemisphere (p=.06) when they did attend to the sounds.

Table 4: Mean amplitudes and standard errors of the difference waveforms (Mistuned–Baseline and Delayed–Baseline) characterizing components ESPm, ORNm and LSPm displayed in attended and non-attended conditions for left- and right-hemispheric channel groups averaged over all ADHD children and control children.

Mean Amplitudes of Components				**Control children** (n=15)		**ADHD** (n=15)	
Component	**Comparison**	**Attention**	**Hemisphere**	**Mean Amplitude (fT)**	**Standard Error (fT)**	**Mean Amplitude (fT)**	**Standard Error (fT)**
ESPm	*Mistuned–Baseline*	Attend	Left	22.51	9.20	----	----
			Right	20.08	6.49	17.23	9.45
		Non-attend	Left	26.87	6.90	15.05	8.01
			Right	33.77	8.00	11.87	9.49
	Delayed–Baseline	Attend	Left	18.93	6.88	14.82	8.99
			Right	23.48	8.44	10.14	11.50
		Non-attend	Left	25.61	7.72	24.84	6.58
			Right	12.76	5.97	21.23	8.95
ORNm	*Mistuned–Baseline*	Attend	Left	-26.74	6.55	-43.79	11.16
			Right	-20.94	6.83	-50.62	13.60
		Non-attend	Left	-22.23	7.18	-6.69	13.06
			Right	-32.42	10.05	-32.68	9.74
	Delayed–Baseline	Attend	Left	-33.41	5.7	-28.04	4.78
			Right	-28.68	8.43	-44.64	7.69
		Non-attend	Left	-11.21	8.83	-31.55	7.58
			Right	-17.83	8.49	-30.85	13.18
LSPm	*Mistuned–Baseline*	Attend	Left	14.79	5.22	29.03	7.46
			Right	11.11	5.11	15.86	7.86
		Non-attend	Left	24.74	5.56	22.53	5.68
			Right	12.10	7.30	24.59	7.23

Table 5: Paired t-test results of the comparison between conditions (Baseline, Mistuned, and Delayed) for each component (ESPm, ORNm, LSPm) and both groups (ADHD children, control children).

Comparison between Conditions				**Control children** (n=15)			**ADHD** (n=15)		
Component	**Comparison**	**Attention**	**Hemisphere**	**t(14)**	**Effect Size r**	**p**	**t(14)**	**Effect Size r**	**p**
ESPm	*Mistuned* vs. *Baseline*	Attend	Left	2.4	0.19	0.03	----	----	----
			Right	3.1	0.11	0.008	1.8	0.09	**0.09 n.s.**
		Non-attend	Left	3.9	0.26	0.002	1.9	0.10	**0.08 n.s.**
			Right	4.2	0.23	<.001	1.2	0.06	**0.2 n.s.**
	Delayed vs. *Baseline*	Attend	Left	2.7	0.15	0.02	1.6	0.10	**0.1 n.s.**
			Right	2.8	0.18	0.01	0.9	0.05	**0.4 n.s.**
		Non-attend	Left	3.3	0.16	0.005	3.8	0.22	0.002
			Right	2.1	0.11	**0.05 n.s.**	2.4	0.12	0.03
ORNm	*Mistuned* vs. *Baseline*	Attend	Left	-4.1	-0.19	0.001	-3.9	-0.21	0.002
			Right	-3.1	-0.20	0.008	-3.7	-0.18	0.002
		Non-attend	Left	-3.1	-0.20	0.008	-0.5	-0.04	**0.6 n.s.**
			Right	-3.2	-0.21	0.006	-3.4	-0.11	0.004
	Delayed vs. *Baseline*	Attend	Left	-5.9	-0.39	<.001	-5.9	-0.25	<.001
			Right	-3.4	-0.21	0.004	-5.8	-0.39	<.001
		Non-attend	Left	-1.3	-0.12	**0.2 n.s.**	-4.2	-0.33	<.001
			Right	-2.1	-0.14	**0.05 n.s.**	-2.3	-0.23	0.03
LSPm	*Mistuned* vs. *Baseline*	Attend	Left	2.8	0.18	0.01	3.9	0.24	0.002
			Right	2.2	0.08	<.05	2.0	0.16	**0.06 n.s.**
		Non-attend	Left	4.4	0.27	<.001	4.0	0.33	0.001
			Right	1.7	0.10	**0.1 n.s.**	3.4	0.28	0.004

Delaying the third harmonic

The grand average difference waveforms *Delayed* minus *Baseline* are also shown in Fig. 4. The ESPm- and ORNm-peak occurred later than in the *Mistuned* condition (ESPm ~229 ms for both groups; ORNm ~300 ms for control children and ~319 ms for ADHD children, see Table 3).

This was anticipated, due to the fact that the onset of the third harmonic was delayed and sound segregation thus occurred later. Corrected for the 160 ms delay of the third harmonic, the ESPm- and ORNm-peak would occur approximately in the same time-range or even a little earlier than in the *Mistuned* condition (corrected for delay: ESPm ~69 ms for both groups; ORNm ~140 ms for control children and ~159 ms for ADHD children). The

ORNm in the *Delayed* condition was also found in adults, peaking at ~277 ms (corrected for delay: ORNm ~117 ms) (Lipp et al., 2010).

A noticeable difference between the morphology of *Delayed* difference waveforms and *Mistuned* difference waveforms was the absence of the late positivity in the *Delayed* condition in both groups of children.

In the ESPm and ORNm time-ranges, mean amplitudes in the two temporal channel groups were computed for attended and non-attended *Baseline* and *Delayed* conditions as described above. Mean amplitudes were retrieved per participant. Table 5 summarises the significance of the components ESPm and ORNm per group. The ESPm was significant in control children when they attended to the sounds and only missed significance in the right hemisphere when they did not attend to the sounds (left hemisphere p=.005, right hemisphere p=.05). In contrast, the ESPm completely lacked significance in ADHD children when they attended to the sounds. It was significant, however, in the non-attended condition. Again, the main difference between ADHD children and control children seemed to lie in the early segregation positivity. The ORNm was significant in both ADHD children and control children when they paid attention to the sounds. In ADHD children, this was also true for the *Non-attend* condition in both hemispheres, whereas the component did not reach significance in the left hemisphere (p=0.2) and only missed significance in the right hemisphere (p=.05) in control children.

Comparison of component amplitude and peak latency between ADHD children and control children

Comparison of component amplitudes (ESPm, ORNm, LSPm) between ADHD children and control children showed no significant main effects. The interaction Hemisphere (*Left, Right*)*Attention (*Attend, Non-attend*) was significant (F(1,28)=4.34, p=.05). ORNm amplitudes for non-attended *Mistuned* conditions in the left hemisphere were smaller than in the right hemisphere (p=.03) and smaller than for *Attend* conditions in the left and right hemisphere (p<.01, respectively). There were no further significant interactions.

Comparison of component peak latencies for *Delayed* stimuli, indicated longer ORNm peak latencies for ADHD children than control children (F(1,26)=11.35, p=.002). Further, a significant main effect was found for Attention (F(1,26)=7.91, p=.009), indicating that ESPm peak latencies were longer in *Non-Attend* conditions than in *Attend*

conditions. Post-hoc testing of the significant interaction Attention (*Attend, Non-attend*)*Group *(children with ADHD, control children)* (F(1,26)=7.16, p=.01) showed that the main effect for Attention could be attributed soley to ADHD children, who had longer ESPm peak latencies in *Non-Attend* conditions than in *Attend* conditions (p=.003). For *Mistuned* stimuli, a main effect for Group *(children with ADHD, control children)* (F(1,27)=10.19, p=.004) was found for the LSPm, indicating that ADHD children had longer peak latencies than control children. Post-hoc testing of the significant interaction Attention (*Attend, Non-attend*)*Group *(children with ADHD, control children)* (F(1,27)=6.01, p=.01) showed that for the *Non-Attend* condition ADHD children had longer peak latencies than control children (p=.01), while for the *Attend* condition ADHD children and control children did not differ.

Correlations between localisation performance, attention test results and MEG components (ESPm, ORNm, LSPm)

Across both groups of children no significant correlations were found between localisation performance (*Mistuned, Delayed* and *Baseline* condition) and MEG component amplitude or latency. A significant correlation was found between localisation performance in the *Baseline* condition and correct reactions and omissions in the sustained attention test (R=0.51, p=.004 and R=-0.51, p=.004, respectively), indicating that the better the children performed on the attention test (high number of correct trials, few omissions), the better they segregated the target harmonic in the *Baseline* stimuli.

2.3.4 Discussion (Study 3)

Interpretation of behavioral results in consideration of neurological markers of primitive grouping mechanisms

Present results indicate that control children and children with ADHD segregated the target partial successfully from the harmonic background. They performed well above chance localizing it, whether it was delayed or mistuned. This suggests that in both conditions, the target partial was perceived as a separate auditory object. However, overall, ADHD children performed worse than control children. This could be due to a perceptual deficit of auditory processing, impaired sustained attention, or a combination of both. It is impossible to distinguish the causes for the impaired performance of ADHD children by looking only at the behavioral results. However, the difference in AEF components between ADHD

children and control children suggests that the poorer sound segregation performance in children with ADHD cannot be attributed solely to inattention: The neurological markers of Bregman`s primitive grouping mechanisms are thought to be largely independent of listener`s attention (Bregman, 1990). The automatic stage of auditory scene analysis (ASA) that parses the acoustic signal into separate auditory objects is proposed to be indexed by the ORNm and early positivity (P80/P75m, here related to as the ESPm) found in MEG studies (Alain & McDonald, 2007; Lipp et al., 2010). Consequently, aberrations in early AEFs (ESPm, ORNm) are thought to reflect impairments of automatic, bottom-up concurrent sound segregation – the primitive grouping mechanism of Bregman's model of ASA, whereas aberrations in subsequent processing stages reflect impairments of top-down mechanisms. Therefore, the present finding of abnormal elicitation of early AEFs in ADHD children is consistent with the assumption that primitive grouping mechanisms of ASA are specifically impaired in children with ADHD.

Group differences in AEF components

The most important difference between ADHD children and control children concerned the early segregation positivity (ESPm). In the *Mistuned* condition ESPm components in all experimental conditions were significant in control children, but non-significant in ADHD children irrespective of attention and hemisphere. The early deflection around 75 ms in the difference waveform Mistuned – Baseline, recently found in adults (Lipp et al., 2010), was interpreted to represent an automatic registration of the mistuned partial, because its amplitude was not affected by attention. Similarly, Alain and McDonald (2007) reported an early positivity (P80) which has been evoked by a mistuned partial of a complex sound while participants were not required to pay attention to the stimuli and interpreted it as reflecting an early registration of inharmonicity in primary auditory cortex. These interpretations can be applied in part to the data from children. The ESPm in control children seemed to reflect an early, bottom-up driven, automatic process of concurrent sound segregation based on inharmonicity *and* onset asynchrony, as it was elicited in both hemispheres by a delayed or mistuned partial, and was present irrespective of attention paid to the sounds or not. In the present study ADHD and control children did not differ significantly in ESPm amplitudes, suggesting that this early marker of concurrent sound segregation exists in both groups. However, based on the lack of statistical significance, the ESPm in ADHD children seemed less consistent and more variable when the third partial

was mistuned. This could reflect a very early impaired automatic processing of concurrent sound segregation based on inharmonicity in children diagnosed with ADHD.

One has to keep in mind that later components, like the ORNm and probably also partly the LSPm, also suggested to reflect bottom-up mechanisms of concurrent sound segregation based on inharmonicity in children, are similarly elicited in both groups regarding their robustness, as indexed by statistical significance, and amplitudes, which did not differ significantly between the groups. This could lead to the conclusion that to differentiate between ADHD children and control children one should focus on the very beginning of the sound segregation process, which seems deficient in ADHD children as manifest by the impaired elicitation of the ESPm when a mistuned harmonic serves as cue for concurrent sound segregation.

In some conditions, children with ADHD had longer peak latencies than control children (ORNm, *Delayed* condition; LSPm, *Non-attend* condition), never conversely. This indicates that concurrent sound segregation is slower in ADHD children. One possible interpretation is that auditory grouping mechanisms require more time in ADHD children, potentially contributing to auditory processing deficits.

Concerning concurrent sound segregation based on *onset asynchrony*, the main difference between ADHD children and control children seemed again to lie in the early segregation positivity. The ESPm completely lacked significance in ADHD children when they attended to the sounds, in contrast to control children. It was significant, however, in the non-attended condition. These results may reflect some impairment in the early segregation process of a delayed partial in children with ADHD, especially, when attention is required. But as an ESPm was elicited in ADHD children when their attention was guided to visual stimuli, it is unlikely that a potential impairment in the early segregation process of a delayed partial is as robustly reflected by a missing ESPm as in the *Mistuned* condition. It is possible that when performing the visual task ADHD children attended to the sounds, even though they were instructed to ignore them, and in contrast became less attentive to the sounds when they were performing the localisation task. This pattern of behaviour could lead to the results outlined above, given that concurrent sound segregation based on onset asynchrony is immature and requires attention to elicit an ESPm in children. However, as *Mistuned, Delayed,* and *Baseline* sounds were presented together in a random order, the absence of similar effects for the *Mistuned* condition is hard to explain.

In summary, at least a subgroup of ADHD children seems to have a genuine and specific impairment in automatic stages of concurrent sound segregation (especially when based on inharmonicity), rather than a general auditory processing deficit.

The chicken or the egg dilemma – or what may be the cause of an auditory processing deficit in ADHD children?

Sussman and Steinschneider (2009) investigated the role of attention in the perception of sound streams by children and young adults with passively and actively obtained electrophysiological indices (mismatch negativity (MMN) and P3b). For children and adults there was good correspondence between the perception of streams and *active* electrophysiological indices. However, in contrast to adults, perception of stream segregation in children did not match with *passive* electrophysiological measures. Thus, MMN should have been elicited at a frequency separation of about 11 semitones, a clearly segregated organization according to behaviour. However, in the passive condition, when children were instructed to ignore the sounds, MMN did not appear until a difference of 31 semitones. Larger contrasts seemed to be required in passive than active listening conditions for children but not adults. The authors` interpretation was that in children of this age (9-12 years), basic auditory processing mechanisms for sound pattern organization are immature and that attention plays a crucial role during normal development in shaping the automatic processes and strengthening the neural networks used for auditory scene analysis. They suggested that for children with disturbed auditory processing, attentional impairments may hamper the ability to develop and refine automatic sound organization processes and ultimately impact on perception. In the context of the present study, the results of Sussman and Steinschneider (2009) are consistent with the idea that the deficient ESPm in children with ADHD may reflect an under-development of automatic auditory processing mechanisms as a consequence of attention deficits during auditory maturation. However, it is also reasonable to expect, that effective sound segregation is required to focus attention on a certain auditory object. The present study delivers convincing support for the assumption that attention deficits in ADHD children may be a consequence of a genuine auditory processing deficit concerning concurrent sound segregation based on inharmonicity.

3 General discussion

The research objective of the present thesis was to extent existing knowledge about the mechanisms of auditory scene analysis (ASA, Bregman, 1990) in three groups of participants: Adults and children with and without attention deficit hyperactivity disorder (ADHD). The present thesis addressed neurological markers of concurrent sound segregation and the bottom-up and top-down processing stages of ASA. Additionally, it aimed to investigate the maturational changes underlying concurrent sound segregation and tried to examine if children with ADHD suffer from a genuine auditory processing deficit. Concurrent sound segregation abilities were analysed by measuring auditory evoked fields (AEFs) using magnetoencephalography (MEG). An experimental design was applied, which allowed to diversify the contribution of attention on sound segregation by an active and passive listening condition and the usage of two cues – inharmonicity and onset asynchrony within a harmonic complex tone – to invoke sound segregation (Darwin & Carlyon, 1995).

The comparison of adults and school-aged children (9-13 years of age) suggests that the mechanisms for concurrent sound segregation operate similarly in both groups. In adults, as well as in children, both cues - inharmonicity and onset asynchrony - evoked sound segregation. However, it could be shown that children performed worse than adults. Children seem to be less efficient than adults in processing the information of mistuning and delay.

3.1 AEFs associated with concurrent sound segregation

Three MEG components associated with concurrent sound segregation were identified in adults and children: an early component (*early segregation positivity*, ESPm) and later deflections, namely the *object related negativity* (ORNm), originally defined by Alain and colleagues (Alain, Arnott, & Picton, 2001; Alain, Schuler, & McDonald, 2002), and a *late segregation positivity* (LSPm). These AEFs seem to reflect relatively automatic, bottom-up sound segregation processes, congruent to Bregman´s primitive grouping mechanism. However, top-down and bottom-up effects of sound organization may interact at an early stage of auditory processing, as it has been previously demonstrated that stimulus-driven sensory mechanisms can be modified under certain circumstances by attentive processes (Carlyon, Cusack, Foxton, & Robertson, 2001; Sussman, Winkler, Huotilainen, Ritter, &

Naatanen, 2002). In adults, there may be a weak effect of attention on the LSPm partly indexing top-down influences, when a mistuned harmonic serves as a cue for concurrent sound segregation. For children, the cues of the present study may fall into an ambiguous domain in which focussed attention to the sounds may support the bottom-up segregation mechanisms. This could explain the present finding in children that *only* when the sounds were attended, a statistically significant ORNm followed the delayed harmonic.

3.2 Maturation of concurrent sound segregation

Further MEG results also demonstrate that automatic (bottom-up) sound segregation processes are not fully developed in children. Especially, sound segregation based on onset asynchrony differs distinctly between adults and children. This is shown in marked differences in the morphology of the difference waveform *Delayed – Baseline* in children compared to adults, whereas the morphology of the difference waveform *Mistuned – Baseline* in children bears great similarity to that of adults. The ESPm was found to be a more general marker for a bottom-up sound segregation process in children, as it was elicited by a delayed *and* mistuned partial. However, in adults the ESPm probably reflects an automatic registration of the *mistuned* harmonic. The ORNm was adult-like in children and was interpreted as a general marker of sound segregation in both groups. Contrary to adults, children lacked a later displacement (LSPm), when onset asynchrony served as cue for concurrent sound segregation. The mentioned differences between AEF morphology in adults and children, as well as longer AEF peak latencies in children compared to adults emphasized the hypothesis that bottom-up mechanisms of concurrent sound segregation are still immature in children of the present age group.

However, concurrent sound segregation in children could be scrutinized in future studies. The present sample of children ranges from 9 to 13 years. As the auditory system still matures within this time-window, as shown for various AEP components (e.g. Kummer et al., 2007; Pang & Taylor, 2000; Ponton, Eggermont, Kwong, & Don, 2000), it is likely that sound segregation skills undergo developmental changes within this span of life, too. Further investigation of groups with smaller age ranges could provide a more detailed view of the maturation of concurrent sound segregation.

It has been cautiously suggested that the detection of temporal changes in the acoustic signal may mature deferred to the detection of spectral changes. An alternative explanation

for the more obvious difference between the children´s and adults´ AEF morphology in the *Delayed* compared to the *Mistuned* condition might be that the applied 160 ms onset delay was too short, whereas 8% mistuning was big enough to maximally elicit sound segregation in children of the present age group with regard to their stage of development. Applying a larger onset delay could have diminished the differences between the children´s and adults´ AEF morphology in the *Delayed* condition. This idea is supported by a study of Alain et al. (2003) who found no significant difference between ERPs elicited by sounds with a mistuned harmonic of 8% or 16% in children (8–12 years), whereas successive auditory stimuli were integrated into a unitary percept with an onset-to-onset pace of 250 ms and smaller in 9–11-year-old children (Wang, Datta, & Sussman, 2005). Nevertheless, given the fact that children showed equal performance of localizing the delayed, respectively the mistuned harmonic, it is likely that the chosen cues (8% mistuning, 160 ms delay) elicited AEFs, which similarly well represent the developmental stage of concurrent sound segregation in children. However, it would be of great interest to manipulate the amount of mistuning and delay for children in future research studies in order to learn more about the thresholds of mistuning and delay, which are necessary to elicit AEFs of concurrent sound segregation automatically in children of a particular age group.

3.3 Concurrent sound segregation in children with ADHD

A further aspect of the present thesis concerned concurrent sound segregation in children with attention problems and potential aberrations of their auditory perceptual abilities. Although both cues - inharmonicity and onset asynchrony - elicited sound segregation in both groups of children, ADHD children performed worse than control children. Is this due to a perceptual deficit of auditory processing, impaired sustained attention, or a combination of both?

While the implication of attention in forming and localizing an auditory object cannot be ruled out entirely, many results are consistent with Bregman´s model of ASA in which an initial pre-attentive mechanism partitions the auditory input into distinct groups of sounds. Although in school-aged children auditory segregation does not operate as efficiently as in adults, the primitive grouping mechanism has been shown to be roughly adult-like for concurrent sound segregation and auditory stream segregation (Alain et al., 2003; Sussman, Ceponiene, Shestakova, Naatanen, & Winkler, 2001). The pre-attentive

analysis of sounds may help the attention processes by facilitating the allocation of the attention focus to a particular subset of sounds. When the outcome of the pre-attentive system reveals more than one sound source, then attention can be efficiently allocated to only one of these sources, allowing the listener to automatically exclude those elements that do not belong to the attended object (for review see Alain & Arnott, 2000). In other words, when the automatic sound segregation process is impaired, it is difficult to pay attention to a specific auditory object. Consequently, this could lead to symptoms of inattention in children with ADHD.

The present results support the idea that, at least for a subgroup of children with ADHD, there is some impairment in pre-attentive processing of sounds. This assumption is supported by the unstable expression of the early segregation positivity in children with ADHD when compared to a healthy control group. Especially when inharmonicity serves as cue for concurrent sound segregation, the ESPm may indicate a genuine auditory processing deficit.

ADHD is believed to constitute a spectrum of disorders that subsume different subtypes of a larger population of children with attention and hearing problems. Still today, the diagnosis (APD vs. ADHD) may depend upon whether an audiologist or a psychologist/pediatrician evaluates the child (Cacace & McFarland, 2005; Moss & Sheiffele, 1994), as overlap exists in the conceptions of audiologists about (C)APD and the conceptions of pediatricians about ADHD. Cacace and McFarland (2005) critically discussed the position that (C)APD could be a cause of ADHD and ask for a demonstration "that there are cases diagnosed as ADHD that can be characterized [...] by an auditory modality-specific deficit that is causally related to attention." In other words, is there deficient auditory processing (bottom-up) in children with ADHD that impairs attention? The present thesis gives rise to the assumption that there are such children, and that there are potential neurological markers which could help to diagnose them. A major challenge will be to refine diagnostic criteria in order to precisely identify those children with auditory processing deficits.

Not all ADHD children respond to stimulant medication (Greenhill, Findling, & Swanson, 2002; Solanto, 1998), e.g. treatment with methylphenidate, which increases the levels of dopamine in the brain through reuptake inhibition of the dopamine transporter (di Michele, Prichep, John, & Chabot, 2005). This supports the hypothesis that in some ADHD

children the primary cause for attention problems might not be an imbalance in dopaminergic and noradrenergic systems. As the outcome of the present thesis indicates that children, diagnosed having ADHD, might suffer from auditory processing deficits, a further challenge will be to assure that they receive adequate treatment (e.g. auditory training), as there is clear evidence that auditory training can be an effective intervention for a variety of auditory-based disorders (Moore, Halliday, & Amitay, 2009).

3.4 Concluding remark and perspective

To conclude, the findings of the present thesis do not only add some important extensions to the knowledge about auditory scene analysis using the example of concurrent sound segregation but also contribute to the controversial issue about auditory processing deficits in children with ADHD. Future research should further elucidate the complex mechanisms of auditory scene analysis in order to upgrade diagnostic criteria and methods of treatment of people with auditory disorders.

4 References

Alain, C. (2007). Breaking the wave: effects of attention and learning on concurrent sound perception. *Hear Res*, 229(1-2), 225-236.

Alain, C., & Arnott, S. R. (2000). Selectively attending to auditory objects. *Front Biosci, 5,* D202-212.

Alain, C., Arnott, S. R., & Picton, T. W. (2001). Bottom-up and top-down influences on auditory scene analysis: evidence from event-related brain potentials. *J Exp Psychol Hum Percept Perform,* 27(5), 1072-1089.

Alain, C., & Bernstein, L. J. (2008). From sounds to meaning: the role of attention during auditory scene analysis. *Current Opinion in Otolaryngology & Head and Neck Surgery(16),* 485-489.

Alain, C., & Izenberg, A. (2003). Effects of attentional load on auditory scene analysis. J *Cogn Neurosci, 15*(7), 1063-1073.

Alain, C., & McDonald, K. L. (2007). Age-related differences in neuromagnetic brain activity underlying concurrent sound perception. *J Neurosci, 27*(6), 1308-1314.

Alain, C., Quan, J., McDonald, K., & Van Roon, P. (2009). Noise-induced increase in human auditory evoked neuromagnetic fields. *Eur J Neurosci, 30*(1), 132-142.

Alain, C., Reinke, K., He, Y., Wang, C., & Lobaugh, N. (2005). Hearing two things at once: neurophysiological indices of speech segregation and identification. *J Cogn Neurosci, 17*(5), 811-818.

Alain, C., Schuler, B. M., & McDonald, K. L. (2002). Neural activity associated with distinguishing concurrent auditory objects. *J Acoust Soc Am, 111*(2), 990-995.

Alain, C., Theunissen, E. L., Chevalier, H., Batty, M., & Taylor, M. J. (2003). Developmental changes in distinguishing concurrent auditory objects. *Brain Res Cogn Brain Res, 16*(2), 210-218.

Albayrak, O., Friedel, S., Schimmelmann, B. G., Hinney, A., & Hebebrand, J. (2008). Genetic aspects in attention-deficit/hyperactivity disorder. *J Neural Transm, 115*(2), 305-315.

Albrecht, R., Suchodoletz, W., & Uwer, R. (2000). The development of auditory evoked dipole source activity from childhood to adulthood. *Clin Neurophysiol, 111*(12), 2268-2276.

American Psychiatric Association (1994). Diagnostic and statistical manual mental disorders DSM-VI. (4th ed.) , American Psychiatric Association, Washington, DC.

Arnsten, A. F., & Li, B. M. (2005). Neurobiology of executive functions: catecholamine influences on prefrontal cortical functions. *Biol Psychiatry, 57*(11), 1377-1384.

AUDIVA. (2004). Elternfragebogen zur Anamnese und Verlaufskontrolle - AUDIVA, Hören und Bewegen. Kandern: AUDIVA

Bamiou, D. E., Musiek, F. E., & Luxon, L. M. (2001). Aetiology and clinical presentations of auditory processing disorders--a review. *Arch Dis Child, 85*(5), 361-365.

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD. *Psychol Bull, 121*(1), 65-94.

Barkley, R. A. (2006). Attention-deficit/ Hyperactivity Disorder: A handbook for diagnosis and treatment (Vol. 3rd edition). New York: Guilford Press.

Barkley, R. A., DuPaul, G. J., & McMurray, M. B. (1990). Comprehensive evaluation of attention deficit disorder with and without hyperactivity as defined by research criteria. *J Consult Clin Psychol, 58*(6), 775-789.

Barkley, R. A., Grodzinsky, G., & DuPaul, G. J. (1992). Frontal lobe functions in attention deficit disorder with and without hyperactivity: a review and research report. *J Abnorm Child Psychol, 20*(2), 163-188.

Barry, R. J., Clarke, A. R., & Johnstone, S. J. (2003). A review of electrophysiology in attention-deficit/hyperactivity disorder: I. Qualitative and quantitative electroencephalography. *Clin Neurophysiol, 114*(2), 171-183.

Bey, C., & McAdams, S. (2002). Schema-based processing in auditory scene analysis. *Percept Psychophys, 64*(5), 844-854.

Bidet-Caulet, A., Fischer, C., Bauchet, F., Aguera, P. E., & Bertrand, O. (2008). Neural substrate of concurrent sound perception: direct electrophysiological recordings from human auditory cortex. *Front Hum Neurosci, 1*, 1-11.

Biederman, J. (2005). Attention-deficit/hyperactivity disorder: a selective overview. *Biol Psychiatry, 57*(11), 1215-1220.

Biederman, J., & Faraone, S. V. (2005). Attention-deficit hyperactivity disorder. *Lancet, 366*(9481), 237-248.

Bregman, A. S. (1990). Auditory Scene Analysis. The Perceptual Organisation of Sounds Camebridge: MIT Press

Bregman, A. S., & Pinker, S. (1978). Auditory streaming and the building of timbre. *Can J Psychol, 32*(1), 19-31.

Breier, J. I., Fletcher, J. M., Foorman, B. R., Klaas, P., & Gray, L. C. (2003). Auditory temporal processing in children with specific reading disability with and without attention deficit/hyperactivity disorder. *J Speech Lang Hear Res, 46*(1), 31-42.

Brickenkamp, R. (2002). Test d2: Aufmerksamkeits-Beslastungs-Test Göttingen: Hogrefe.

Brunner, M. (2007). ["Let the cobbler stick to his last"]. *HNO, 55*(4), 241-244; discussion 244.

Cacace, A. T., & McFarland, D. J. (1998). Central auditory processing disorder in school-aged children: a critical review. *J Speech Lang Hear Res, 41*(2), 355-373.

Cacace, A. T., & McFarland, D. J. (2005). Delineating Auditory Processing Disorder (APD) and Attention Deficit Hyperactivity Disorder (ADHD): A Conceptual, Theoretical, and Practical Framework. In T. K. Parthasarathy (Ed.), An Introduction to Auditory Processing Disorders in Children (pp. 39-61): Lawrence Erlbaum Associates Inc.

Carlyon, R. P. (2004). How the brain separates sounds. *Trends Cogn Sci, 8*(10), 465-471.

Carlyon, R. P., Cusack, R., Foxton, J. M., & Robertson, I. H. (2001). Effects of attention and unilateral neglect on auditory stream segregation. *J Exp Psychol Hum Percept Perform, 27*(1), 115-127.

Chalikia, M. H., & Bregman, A. S. (1989). The perceptual segregation of simultaneous auditory signals: pulse train segregation and vowel segregation. *Percept Psychophys, 46*(5), 487-496.

Chamberlain, S. R., Robbins, T. W., & Sahakian, B. J. (2007). The neurobiology of attention-deficit/hyperactivity disorder. *Biol Psychiatry, 61*(12), 1317-1319.

Chermak, G. D., Hall, J. W., 3rd, & Musiek, F. E. (1999). Differential diagnosis and management of central auditory processing disorder and attention deficit hyperactivity disorder. *J Am Acad Audiol, 10*(6), 289-303.

Chermak, G. D., Somers, E. K., & Seikel, J. A. (1998). Behavioral signs of central auditory processing disorder and attention deficit hyperactivity disorder. *J Am Acad Audiol, 9*(1), 78-84; quiz 85.

Chermak, G. D., Tucker, E., & Seikel, J. A. (2002). Behavioral characteristics of auditory processing disorder and attention-deficit hyperactivity disorder: predominantly inattentive type. *J Am Acad Audiol, 13*(6), 332-338.

Clarke, A. R., Barry, R. J., McCarthy, R., & Selikowitz, M. (1998). EEG analysis in Attention-Deficit/Hyperactivity Disorder: a comparative study of two subtypes. *Psychiatry Res, 81*(1), 19-29.

Clarke, A. R., Barry, R. J., McCarthy, R., Selikowitz, M., Clarke, D. C., & Croft, R. J. (2003). Effects of stimulant medications on children with attention-deficit/hyperactivity disorder and excessive beta activity in their EEG. *Clin Neurophysiol, 114*(9), 1729-1737.

Culling, J. F., & Darwin, C. J. (1993). Perceptual separation of simultaneous vowels: within and across-formant grouping by F0. *J Acoust Soc Am, 93*(6), 3454-3467.

Cusack, R., Deeks, J., Aikman, G., & Carlyon, R. P. (2004). Effects of location, frequency region, and time course of selective attention on auditory scene analysis. *J Exp Psychol Hum Percept Perform, 30*(4), 643-656.

Dannenbring, G. L., & Bregman, A. S. (1978). Streaming vs. fusion of sinusoidal components of complex tones. *Percept Psychophys, 24*(4), 369-376.

Darwin, C. J., & Carlyon, R. P. (1995). Auditory Grouping. In B. C. Moore (Ed.), Hearing. London: Academic Press Limited.

Demany, L. (1982). Auditory Stream Segregation in Infancy. *Infant Behav Dev*(5), 261-276.

di Michele, F., Prichep, L., John, E. R., & Chabot, R. J. (2005). The neurophysiology of attention-deficit/hyperactivity disorder. *Int J Psychophysiol, 58*(1), 81-93.

Doepfner, J., Goertz-Dorten, A., & Lehmkuhl, G. (2008). DISYPS-II. Diagnostik-System fuer psychische Stoerungen nach ICD-10 und DSM-IV fuer Kinder und Jugendliche - II. Bern: Verlag Hans Huber, Hogrefe AG.

Doepfner, J., Plueck, J., Boelte, K., Lenz, P., Melchers, P., & Heim, K. (1998). Elternfragebogen über das Verhalten von Kindern und Jugendlichen. Deutsche Bearbeitung der Child Behavior Checklist (CBCL/4-18). Einführung und Anleitung zur Handauswertung (2. Aufl. mit deutschen Normen), Arbeitsgruppe Kinder-, Jugend-, und Familiendiagnostik (KJFD). Köln.

Dünker, H., Lienert, G. A., Lukesch, H., & Mayrhofer, S. (2001). KLT-R, Konzentrations-Leistungs-Test - Revidierte Fassung. Göttingen: Hogrefe.

Dyson, B. J., & Alain, C. (2004). Representation of concurrent acoustic objects in primary auditory cortex. *J Acoust Soc* Am, 115(1), 280-288.

Dyson, B. J., Alain, C., & He, Y. (2005). Effects of visual attentional load on low-level auditory scene analysis. *Cogn Affect Behav Neurosci, 5*(3), 319-338.

Fallon, M., Trehub, S. E., & Schneider, B. A. (2000). Children's perception of speech in multitalker babble. *J Acoust Soc Am, 108*(6), 3023-3029.

Faraone, S. V., Biederman, J., Weiffenbach, B., Keith, T., Chu, M. P., Weaver, A., et al. (1999). Dopamine D4 gene 7-repeat allele and attention deficit hyperactivity disorder. *Am J Psychiatry, 156*(5), 768-770.

Fitch, R. H., Miller, S., & Tallal, P. (1997). Neurobiology of speech perception. *Annu Rev Neurosci,* 20, 331-353.

Geffner, D., Lucker, J. R., & Koch, W. (1996). Evaluation of auditory discrimination in children with ADD and without ADD. *Child Psychiatry Hum Dev, 26*(3), 169-179.

Gerlach, M., Deckert, J., Rothenberger, A., & Warnke, A. (2008). Pathogenesis and pathophysiology of attention-deficit/hyperactivity disorder: from childhood to adulthood. *J Neural Transm, 115*(2), 151-153.

Gilley, P. M., Sharma, A., Dorman, M., & Martin, K. (2005). Developmental changes in refractoriness of the cortical auditory evoked potential. *Clin Neurophysiol, 116*(3), 648-657.

Greenhill, L. L., Findling, R. L., & Swanson, J. M. (2002). A double-blind, placebo-controlled study of modified-release methylphenidate in children with attention-deficit/hyperactivity disorder. *Pediatrics, 109*(3), E39.

Grossberg, S., Govindarajan, K. K., Wyse, L. L., & Cohen, M. A. (2004). ARTSTREAM: a neural network model of auditory scene analysis and source segregation. *Neural Netw, 17*(4), 511-536.

Hartmann, W. M., McAdams, S., & Smith, B. K. (1990). Hearing a mistuned harmonic in an otherwise periodic complex tone. *J Acoust Soc Am, 88*(4), 1712-1724.

Haykin, S., & Chen, Z. (2005). The cocktail party problem. *Neural Comput, 17*(9), 1875-1902.

Hill, N. I., & Darwin, C. J. (1996). Lateralization of a perturbed harmonic: effects of onset asynchrony and mistuning. *J Acoust Soc Am, 100*(4 Pt 1), 2352-2364.

Hulse, S. H., MacDougall-Shackleton, S. A., & Wisniewski, A. B. (1997). Auditory scene analysis by songbirds: stream segregation of birdsong by European starlings (Sturnus vulgaris). *J Comp Psychol, 111*(1), 3-13.

Izumi, A. (2002). Auditory stream segregation in Japanese monkeys. *Cognition, 82*(3), B113-122.

Johnson, C. E. (2000). Children's phoneme identification in reverberation and noise. *J Speech Lang Hear Res, 43*(1), 144-157.

Johnstone, S. J., Barry, R. J., Anderson, J. W., & Coyle, S. F. (1996). Age-related changes in child and adolescent event-related potential component morphology, amplitude and latency to standard and target stimuli in an auditory oddball task. *Int J Psychophysiol, 24*(3), 223-238.

Kaas, J. H., & Hackett, T. A. (2000). Subdivisions of auditory cortex and processing streams in primates. *Proc Natl Acad Sci U S A, 97*(22), 11793-11799.

Keller, W. D., & Tillery, K. L. (2002). Reliable Differential Diagnosis and Effective Management of Auditory Processing and Attention Deficit Hyperactivity Disorders. *Seminars of Hearing, 23*(4), 337-347.

Kemner, C., Jonkman, L. M., Kenemans, J. L., Bocker, K. B., Verbaten, M. N., & Van Engeland, H. (2004). Sources of auditory selective attention and the effects of methylphenidate in children with attention-deficit/hyperactivity disorder. *Biol Psychiatry, 55*(7), 776-778.

Klump, G. M., van Hemmen, J. L., & Sejnowski, T. J. (2006). How does the hearing system perform auditory scene analysis? 23 problems in systems neuroscience, 303-321. New York, NY US: Oxford University Press.

Krause, K. H., Dresel, S. H., Krause, J., la Fougere, C., & Ackenheil, M. (2003). The dopamine transporter and neuroimaging in attention deficit hyperactivity disorder. *Neurosci Biobehav Rev, 27*(7), 605-613.

Kummer, P., Burger, M., Schuster, M., Rosanowski, F., Eysholdt, U., & Hoppe, U. (2007). Cortical auditory evoked potentials to acoustic changes in speech stimuli in children. *Folia Phoniatr Logop, 59*(5), 273-280.

Lahat, E., Avital, E., Barr, J., Berkovitch, M., Arlazoroff, A., & Aladjem, M. (1995). BAEP studies in children with attention deficit disorder. *Dev Med Child Neurol, 37*(2), 119-123.

LaHoste, G. J., Swanson, J. M., Wigal, S. B., Glabe, C., Wigal, T., King, N., et al. (1996). Dopamine D4 receptor gene polymorphism is associated with attention deficit hyperactivity disorder. *Mol Psychiatry, 1*(2), 121-124.

Liotti, M., Pliszka, S. R., Perez, R., Kothmann, D., & Woldorff, M. G. (2005). Abnormal brain activity related to performance monitoring and error detection in children with ADHD. *Cortex, 41*(3), 377-388.

Lipp, R., Kitterick, P., Summerfield, Q., Bailey, P. J., & Paul-Jordanov, I. (2010). Concurrent sound segregation based on inharmonicity and onset asynchrony. *Neuropsychologia, 48*(5), 1417-1425.

Lucker, J. R., Geffner, D., & Koch, W. (1996). Perception of loudness in children with ADD and without ADD. *Child Psychiatry Hum Dev, 26*(3), 181-190.

Macken, W. J., Tremblay, S., Houghton, R. J., Nicholls, A. P., & Jones, D. M. (2003). Does auditory streaming require attention? Evidence from attentional selectivity in short-term memory. *J Exp Psychol Hum Percept Perform, 29*(1), 43-51.

McAdams, S., & Bertoncini, J. (1997). Organization and discrimination of repeating sound sequences by newborn infants*. J Acoust Soc Am, 102*(5 Pt 1), 2945-2953.

McDonald, K. L., & Alain, C. (2005). Contribution of harmonicity and location to auditory object formation in free field: evidence from event-related brain potentials. *J Acoust Soc Am, 118*(3 Pt 1), 1593-1604.

Moore, B. C., Glasberg, B. R., & Peters, R. W. (1986). Thresholds for hearing mistuned partials as separate tones in harmonic complexes. *J Acoust Soc Am, 80*(2), 479-483.

Moore, D. R., Halliday, L. F., & Amitay, S. (2009). Use of auditory learning to manage listening problems in children. *Philos Trans R Soc Lond B Biol Sci, 364*(1515), 409-420.

Morrongiello, B. A., & Trehub, S. E. (1987). Age-related changes in auditory temporal perception*. J Exp Child Psychol, 44*(3), 413-426.

Moss, W. L., & Sheiffele, W. A. (1994). Can we differentially diagnose an attention deficit disorder without hyperactivity from a central auditory processing problem? *Child Psychiatry Hum Dev, 25*(2), 85-96.

Multiaxial classification of child and adolescent psychiatric disorders: The ICD-10 classification of mental and behavioural disorders in children and adolescents (1996). New York, NY US: Cambridge University Press.

N'Diaye, K., Ragot, R., Garnero, L., & Pouthas, V. (2004). What is common to brain activity evoked by the perception of visual and auditory filled durations? A study with MEG and EEG co-recordings*. Brain Res Cogn Brain Res, 21*(2), 250-268.

Nichols, S. L., & Waschbusch, D. A. (2004). A review of the validity of laboratory cognitive tasks used to assess symptoms of ADHD. *Child Psychiatry Hum Dev, 34*(4), 297-315.

Nickisch, A., Gross, M., Schonweiler, R., Uttenweiler, V., am Zehnhoff-Dinnesen, A., Berger, R., et al. (2007). [Auditory processing disorders : Consensus statement by the German Society for Phoniatry and Paedaudiology.]. *HNO, 55*(1), 61-72.

Nigg, J. T. (2005). Neuropsychologic theory and findings in attention-deficit/hyperactivity disorder: the state of the field and salient challenges for the coming decade. *Biol Psychiatry, 57*(11), 1424-1435.

Nigg, J. T., Willcutt, E. G., Doyle, A. E., & Sonuga-Barke, E. J. (2005). Causal heterogeneity in attention-deficit/hyperactivity disorder: do we need neuropsychologically impaired subtypes? *Biol Psychiatry, 57*(11), 1224-1230.

Okamoto, H., Stracke, H., Wolters, C. H., Schmael, F., & Pantev, C. (2007). Attention improves population-level frequency tuning in human auditory cortex. *J Neurosci, 27*(39), 10383-10390.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia, 9*(1), 97-113.

Pang, E. W., & Taylor, M. J. (2000). Tracking the development of the N1 from age 3 to adulthood: an examination of speech and non-speech stimuli. *Clin Neurophysiol, 111*(3), 388-397.

Pliszka, S. R. (2000). Patterns of psychiatric comorbidity with attention-deficit/hyperactivity disorder*. Child Adolesc Psychiatr Clin N Am, 9*(3), 525-540, vii.

Pliszka, S. R. (2005). The neuropsychopharmacology of attention-deficit/hyperactivity disorder. *Biol Psychiatry, 57*(11), 1385-1390.

Pliszka, S. R., Liotti, M., & Woldorff, M. G. (2000). Inhibitory control in children with attention-deficit/hyperactivity disorder: event-related potentials identify the processing component and timing of an impaired right-frontal response-inhibition mechanism*. Biol Psychiatry, 48*(3), 238-246.

Polley, D. B., Steinberg, E. E., & Merzenich, M. M. (2006). Perceptual learning directs auditory cortical map reorganization through top-down influences. *J Neurosci, 26*(18), 4970-4982.

Ponton, C. W., Eggermont, J. J., Kwong, B., & Don, M. (2000). Maturation of human central auditory system activity: evidence from multi-channel evoked potentials. *Clin Neurophysiol, 111*(2), 220-236.

Ptok, M., Blachnik, P., & Schonweiler, R. (2004). [Late auditory potentials (NC-ERP) in children with symptoms of auditory processing and perception disorder. With and without attention deficit disorder]. *HNO, 52*(1), 67-75.

Ptok, M., Buller, N., Schwemmle, C., Bergmann, C., & Luerssen, K. (2006). [Auditory processing disorder versus attention deficit/hyperactivity disorder. A dysfunction complex or different entities?]. *HNO, 54*(5), 405-408, 410-404.

Purvis, K. L., & Tannock, R. (2000). Phonological processing, not inhibitory control, differentiates ADHD and reading disability. *J Am Acad Child Adolesc Psychiatry, 39*(4), 485-494.

Rosenkötter, H. (1998). Praktische Diagnostik der auditiven Wahrnehmung. Auditive Wahrnehmung und Hörtraining, 2. Arbeitstagung in Ludwigsburg: AUDIVA Verlag.

Sams, M., Paavilainen, P., Alho, K., & Naatanen, R. (1985). Auditory frequency discrimination and event-related potentials. *Electroencephalogr Clin Neurophysiol, 62*(6), 437-448.

Scahill, L., & Schwab-Stone, M. (2000). Epidemiology of ADHD in school-age children. *Child Adolesc Psychiatr Clin N Am, 9*(3), 541-555, vii.

Snyder, J. S., & Alain, C. (2007). Toward a neurophysiological theory of auditory stream segregation. *Psychol Bull, 133*(5), 780-799.

Solanto, M. V. (1998). Neuropsychopharmacological mechanisms of stimulant drug action in attention-deficit hyperactivity disorder: a review and integration. *Behav Brain Res, 94*(1), 127-152.

Solanto, M. V., Abikoff, H., Sonuga-Barke, E., Schachar, R., Logan, G. D., Wigal, T., et al. (2001). The ecological validity of delay aversion and response inhibition as measures of impulsivity in AD/HD: a supplement to the NIMH multimodal treatment study of AD/HD. *J Abnorm Child Psychol, 29*(3), 215-228.

Sonuga-Barke, E. J. (2003). The dual pathway model of AD/HD: an elaboration of neuro-developmental characteristics. *Neurosci Biobehav Rev, 27*(7), 593-604.

Sonuga-Barke, E. J. (2005). Causal models of attention-deficit/hyperactivity disorder: from common simple deficits to multiple developmental pathways. *Biol Psychiatry, 57(*11), 1231-1238.

Sonuga-Barke, E. J., Dalen, L., & Remington, B. (2003). Do executive deficits and delay aversion make independent contributions to preschool attention-deficit/hyperactivity disorder symptoms? *J Am Acad Child Adolesc Psychiatry, 42*(11), 1335-1342.

Summerfield, Q., Paul, I., Kitterick, P., Briley, P., & Bailey, P. (2007). Neuro-magnetic indices of auditory-perceptualsegregation of mistuned and delayed tones by adults and children. 30th Mid-Winter Meeting of the Association For Research In Otolaryngology, Denver, Colorado, USA, p. 147.

Surwillo, W. W. (1981). Recovery of the cortical evoked potential from auditory stimulation in children and adults. *Dev Psychobiol, 14*(1), 1-12.

Sussman, E., Ceponiene, R., Shestakova, A., Naatanen, R., & Winkler, I. (2001). Auditory stream segregation processes operate similarly in school-aged children and adults. *Hear Res, 153*(1-2), 108-114.

Sussman, E., & Steinschneider, M. (2009). Attention effects on auditory scene analysis in children. *Neuropsychologia, 47*(3), 771-785.

Sussman, E., Steinschneider, M., Gumenyuk, V., Grushko, J., & Lawson, K. (2008). The maturation of human evoked brain potentials to sounds presented at different stimulus rates. *Hear Res, 236*(1-2), 61-79.

Sussman, E., Winkler, I., Huotilainen, M., Ritter, W., & Naatanen, R. (2002). Top-down effects can modify the initially stimulus-driven auditory organization. *Brain Res Cogn Brain Res, 13*(3), 393-405.

Sussman, E. S., Horvath, J., Winkler, I., & Orr, M. (2007). The role of attention in the formation of auditory streams. *Percept Psychophys, 69*(1), 136-152.

Sutcliffe, P. A., Bishop, D. V., Houghton, S., & Taylor, M. (2006). Effect of attentional state on frequency discrimination: a comparison of children with ADHD on and off medication. *J Speech Lang Hear Res, 49*(5), 1072-1084.

Takeshita, K., Nagamine, T., Thuy, D. H., Satow, T., Matsuhashi, M., Yamamoto, J., et al. (2002). Maturational change of parallel auditory processing in school-aged children revealed by simultaneous recording of magnetic and electric cortical responses. *Clin Neurophysiol, 113*(9), 1470-1484.

Tannock, R. (1998). Attention deficit hyperactivity disorder: advances in cognitive, neurobiological, and genetic research. *J Child Psychol Psychiatry, 39*(1), 65-99.

Thompson, N. C., Cranford, J. L., & Hoyer, E. (1999). Brief-tone frequency discrimination by children. *J Speech Lang Hear Res, 42*(5), 1061-1068.

Thompson, R. F., & Spencer, W. A. (1966). Habituation: a model phenomenon for the study of neuronal substrates of behavior. *Psychol Rev, 73*(1), 16-43.

Tillery, K. L., Katz, J., & Keller, W. D. (2000). Effects of methylphenidate (Ritalin) on auditory performance in children with attention and auditory processing disorders. *J Speech Lang Hear Res, 43*(4), 893-901.

Tonnquist-Uhlen, I. (1996). Topography of auditory evoked cortical potentials in children with severe language impairment. *Scand Audiol Suppl, 44*, 1-40.

Tonnquist-Uhlen, I., Borg, E., & Spens, K. E. (1995). Topography of auditory evoked long-latency potentials in normal children, with particular reference to the N1 component. *Electroencephalogr Clin Neurophysiol, 95*(1), 34-41.

Trehub, S. E., Schneider, B. A., & Henderson, J. L. (1995). Gap detection in infants, children, and adults. *J Acoust Soc Am, 98*(5 Pt 1), 2532-2541.

Wang, W., Datta, H., & Sussman, E. (2005). The development of the length of the temporal window of integration for rapidly presented auditory information as indexed by MMN. *Clin Neurophysiol, 116*(7), 1695-1706.

Werner, L. A., Marean, G. C., Halpin, C. F., Spetner, N. B., & Gillenwater, J. M. (1992). Infant auditory temporal acuity: gap detection. *Child Dev, 63*(2), 260-272.

Wienbruch, C., Paul, I., Bauer, S., & Kivelitz, H. (2005). The influence of methylphenidate on the power spectrum of ADHD children - an MEG study. *BMC Psychiatry*, 5, 29.

Willcutt, E. G., Doyle, A. E., Nigg, J. T., Faraone, S. V., & Pennington, B. F. (2005). Validity of the executive function theory of attention-deficit/hyperactivity disorder: a meta-analytic review. *Biol Psychiatry, 57*(11), 1336-1346.

Winkler, I., Kushnerenko, E., Horvath, J., Ceponiene, R., Fellman, V., Huotilainen, M., et al. (2003). Newborn infants can organize the auditory world. *Proc Natl Acad Sci U S A, 100*(20), 11812-11815.

Wood, N. L., & Cowan, N. (1995). The cocktail party phenomenon revisited: attention and memory in the classic selective listening procedure of Cherry (1953). *J Exp Psychol Gen, 124*(3), 243-262.

Wunderlich, J. L., & Cone-Wesson, B. K. (2006). Maturation of CAEP in infants and children: a review. *Hear Res, 212*(1-2), 212-223.

Zimmermann, P., Gondan, M., & Fimm, B. (2002). Testbatterie zur Aufmerksamkeits-prüfung für Kinder. Herzogenrath: Psytest, Fimm, V., Psychologische Testsysteme.

5 Indices

5.1 Figure index

General introduction

Figure 1	*Schematic spectrogram of galloping rhythm sequence used to study streaming. [Adapted from Carylon (2004)].*	*4*

Study 1

Figure 1 (Study 1)	*Schematic spectrograms of Baseline, Mistuned and Delayed stimuli.*	*27*
Figure 2 (Study 1)	*Difference waveforms (Delayed – Baseline) for each sensor averaged over 16 participants. Channel groups used for data analysis are circled (6 channels for each hemisphere).*	*31*
Figure 3 (Study 1)	*Localisation performance during the MEG scan and during headphone presentation. Error bars indicate standard errors.*	*33*
Figure 4 (Study 1)	*Waveform averaged over all participants (16) in the right-hemisphere (6 channels), evoked by Baseline, Mistuned and Delayed stimuli in the Attend condition; band-pass filtered from 1 Hz to 25 Hz.*	*33*
Figure 5 (Study 1)	*Difference waveforms for components P75m, ORNm and P230m in attended and non-attended conditions for left- and right-hemispheric channel groups averaged over all participants. The polarity of the signal is reversed in the left hemisphere. Blue boxes outline time-windows used for statistical analyses. P75m and ORNm are displayed in the upper four graphs (signals were band-pass filtered from 1 Hz to 25 Hz). P230m waveforms are displayed in the lower four graphs (band-pass filter between 1Hz and 8 Hz).*	*35*

Study 2

Figure 1 (Study 2)	*Schematic spectrograms of Baseline, Mistuned and Delayed stimuli.*	*50*
Figure 2 (Study 2)	*Difference waveforms (Delayed - Baseline) for each sensor averaged over 15 children in the Attend condition. Channel Groups used for data analysis are framed (7 channels for each hemisphere). Left hemisphere (same channels for adults and children): A056, A057, A078, A079, A080, A099, and A100. Right hemisphere: A044, A065, A066, A087, A088, A089, A109 for children, and A64, A65, A87, A88, A107, A108, A109 for adults.*	*53*
Figure 3 (Study 2)	*Localisation performance of adults and children during the MEG scan for Baseline, Mistuned, and Delayed stimuli. Error bars indicate standard errors.*	*54*

Figure 4 (Study 2) *Waveform averaged over adults (15) and children (15) in the right hemisphere (7 channels), evoked by Baseline stimuli in the Non-Attend condition; band-pass filtered from 1 to 25 Hz.* *55*

Figure 5 (Study 2) *Difference waveforms for components ESPm and ORNm in attended and non-attended conditions for left and right hemispheric channel groups averaged over all participants (15) of each group (children/adults). Difference waveforms (Mistuned – Baseline) are displayed in the upper four graphs; difference waveforms (Delayed – Baseline) are displayed in the lower four graphs (here: no ESPm for adults are marked). The polarity of the signal is reversed in the left hemisphere. Coloured boxes emblematize time-windows used for statistical analyses. Signals were band-pass filtered from 1 to 25 Hz.* *56*

Figure 6 (Study 2) *Difference waveforms (Mistuned – Baseline) for component LSPm in attended and non-attended conditions for left and right hemispheric channel groups averaged over all participants (15) of each group (children/adults). Difference waveforms (Mistuned – Baseline) are displayed in the upper four graphs; difference waveforms (Delayed – Baseline) are displayed in the lower four graphs (here: no LSPm for children are marked). The polarity of the signal is reversed in the left hemisphere. Coloured boxes emblematize time-windows used for statistical analyses. Signals were band-pass filtered from 1 to 8 Hz.* *57*

Study 3

Figure 1 (Study 3) *Schematic spectrograms of Baseline, Mistuned and Delayed stimuli.* *77*

Figure 2 (Study 3) *Difference waveforms (Delayed - Baseline) for each sensor averaged over 15 control children in the Attend condition. Channel Groups used for data analysis are coloured black (7 channels for each hemisphere). Left hemisphere: A056, A057, A078, A079, A080, A099, and A100. Right hemisphere: A044, A065, A066, A087, A088, A089, and A109.* *79*

Figure 3 (Study 3) *Localisation performance of ADHD children and control children during the MEG scan for Baseline, Mistuned, and Delayed stimuli. Error bars indicate standard errors.* *83*

Figure 4 (Study 3) *Difference waveforms for components ESPm and ORNm in attended and non-attended conditions for left and right hemispheric channel groups averaged over all participants (15) of each group of children. Difference waveforms (Mistuned – Baseline) are displayed in the upper four graphs; difference waveforms (Delayed – Baseline) are displayed in the lower four graphs. The polarity of the signal is reversed in the left hemisphere. Coloured boxes emblematize time-windows used for statistical analyses. Signals were band-pass filtered from 1 to 25 Hz.* *85*

Figure 5 (Study 3) *Difference waveforms (Mistuned – Baseline) for component LSPm in attended and non-attended conditions for left and right hemispheric channel groups averaged over all participants (15) of each group of children. The polarity of the signal is reversed in the left hemisphere. Coloured boxes emblematize time-windows used for statistical analyses. Signals were band-pass filtered from 1 to 8 Hz.* *86*

5.2 Table index

Study 1

Table 1 (Study 1) — *Attention test scores of the long (KLT-R) and short term (D2) attention test.* — *32*

Table 2 (Study 1) — *Peak latencies for components P75m, ORNm and P230m in attended and non-attended conditions for left- and right-hemispheric channel groups averaged over all participants. Time-ranges used for calculating mean amplitudes and further statistical analyses are displayed.* — *34*

Table 3 (Study 1) — *Main Effect of Sound Condition. F- and p-values of ANOVAs.* — *36*

Table 4 (Study 1) — *Paired t-test results of the comparison between conditions (Baseline, Mistuned and Delayed) for each component (P75m, ORNm, P230m) displayed along with mean amplitudes and standard errors of the difference waveforms (Mistuned–Baseline and Delayed–Baseline).* — *37*

Study 2

Table 1 (Study 2) — *Peak latencies for components ESPm, ORNm and LSPm in attended and non-attended conditions for left- and right-hemispheric channel groups averaged over all children and adults. Time-ranges used for calculating mean amplitudes are displayed.* — *58*

Table 2 (Study 2) — *Mean amplitudes and standard errors of the difference waveforms (Mistuned–Baseline and Delayed–Baseline) characterizing components ESPm, ORNm and LSPm displayed in attended and non-attended conditions for left- and right-hemispheric channel groups averaged over all children and adults.* — *59*

Table 3 (Study 2) — *Main Effect of Sound Condition. F- and p-values of ANOVAs with Group (children, adults) as between group factor. Exception: ANOVAs for ESPm and LSPm in Delayed condition were calculated respectively for one group.* — *60*

Table 4 (Study 2) — *Paired t-test results of the comparison between conditions (Baseline, Mistuned, and Delayed) for each component (ESPm, ORNm, LSPm) and both groups (children, adults).* — *61*

Table 5 (Study 2) — *Significant results of Spearman´s rank correlations between localisation performance (Mistuned and Baseline condition) and MEG component latency (ESPm, ORNm, LSPm).* — *63*

Study 3

Table 1 (Study 3) — *Group comparison of CBCL syndrome scales and number of children with clinically relevant T-scores for syndromes.* — *82*

Table 2 (Study 3) — *Results of ADHD questionnaire (FBB-ADHD) from Diagnostic System of Psychiatric Disorders in Children and Adolescents (DISYPS). Stanine scores for attention deficit, hyperactivity, impulsivity, and ADHD symptomatology in total for ADHD children and control children. Number of ADHD children and control children with clinically relevant cut-off scores for each syndrome scale.* — *83*

Table 3 (Study 1) *Peak latencies for components ESPm, ORNm and LSPm in attended and non-attended conditions for left- and right-hemispheric channel groups averaged over all ADHD children and control children. Time-ranges used for calculating mean amplitudes are displayed.* *84*

Table 4 (Study 3) *Mean amplitudes and standard errors of the difference waveforms (Mistuned–Baseline and Delayed–Baseline) characterizing components ESPm, ORNm and LSPm displayed in attended and non-attended conditions for left- and right-hemispheric channel groups averaged over all ADHD children and control children.* *87*

Table 5 (Study 3) *Paired t-test results of the comparison between conditions (Baseline, Mistuned, and Delayed) for each component (ESPm, ORNm, LSPm) and both groups (ADHD children, control children).* *88*