Newborn infants detect cues of concurrent sound segregation

Alexandra Bendixen[1,2], Gábor P. Háden[3,4], Renáta Németh[3,5], Dávid Farkas[3,6], Miklós Török[7], and István Winkler[3,8]

[1]Auditory Psychophysiology Lab, Department of Psychology, Cluster of Excellence “Hearing4all”, European Medical School, Carl von Ossietzky University of Oldenburg, Oldenburg, Germany

[2]Institute of Psychology, University of Leipzig, Leipzig, Germany

[3]Institute of Cognitive Neuroscience and Psychology, Research Centre for Natural Sciences, Hungarian Academy of Sciences, Budapest, Hungary

[4]Institute for Logic, Language and Computation, University of Amsterdam, Amsterdam, The Netherlands

[5]Department of Cognitive Science, Central European University, Budapest, Hungary

[6]Department of Cognitive Science, Faculty of Natural Sciences, Budapest University of Technology and Economics, Budapest, Hungary

[7]Department of Obstetrics-Gynaecology and Perinatal Intensive Care Unit, Military Hospital, Budapest, Hungary

[8]Department of Cognitive and Neuropsychology, Institute of Psychology, University of Szeged, Szeged, Hungary

Address correspondence to:

Alexandra Bendixen

Auditory Psychophysiology Lab

Department of Psychology

Cluster of Excellence “Hearing4all”

European Medical School

Carl von Ossietzky University of Oldenburg

Küpkersweg 74

D-26129 Oldenburg, Germany

Tel: +49-(0)441-798-4844

Fax: +49-(0)441-798-5522

E-mail: alexandra.bendixen@uni-oldenburg.de

Keywords

object-related negativity (ORN), mistuning, onset asynchrony, electroencephalogram (EEG), neonates, sound segregation, instantaneous cues, vertical grouping, auditory scene analysis

Abstract

Separating concurrent sounds is fundamental for veridical perception of one's auditory surroundings. Sound components that are harmonically related and start at the same time are usually grouped into a common perceptual object, whereas components that are not in harmonic relation or have different onset times are more likely to be perceived in terms of separate objects. Here we tested whether neonates are able to pick up the cues supporting this sound organization principle. We presented newborn infants with series of complex tones with their harmonics in tune (creating the percept of a unitary sound object) and with manipulated variants, which gave the impression of two concurrently active sound sources. The manipulated variant had either one partial mistuned (single-cue condition), or the onset of this mistuned partial was also delayed (double-cue condition). Tuned and manipulated sounds were presented in random order with equal probabilities. Recording the neonates' electroencephalographic (EEG) responses allowed us to evaluate their processing of the sounds. Results show that in both conditions, mistuned sounds elicited a negative displacement of the event-related potential (ERP) relative to tuned sounds from 360 to 400 ms after sound onset. The mistuning-related ERP component resembles the object-related negativity (ORN) component in adults, which is associated with concurrent sound segregation. Delayed onset additionally led to a negative displacement from 160 to 200 ms, probably more related to the physical parameters of the sounds than to their perceptual segregation. The elicitation of an ORN-like response in newborn infants suggests that neonates possess the basic capabilities for segregating concurrent sounds by detecting inharmonic relations between the co-occurring sounds.

Introduction

Our auditory environment abounds in simultaneously active sound sources whose signals overlap and must be disentangled to retrieve meaningful information [1]. The decomposition process, termed *auditory scene analysis* [2], has been conceptualized in terms of two types of operations [3,4]: *sequential* analysis (helping the listener to identify whether consecutive sounds belong together or not), and *simultaneous* analysis (helping to decide whether concurrent sounds belong together or not). Because these processes operate smoothly in most individuals, we are hardly aware of the challenges they pose, and we usually take for granted that other individuals hear the “auditory world” in a similar way as we do. The present study investigates whether this assumption is justified for newborn infants: Are the cues needed for disentangling concurrent sounds processed at birth?

Cues for concurrent sound segregation refer to acoustic properties that are immediately (i.e., on a millisecond time scale) indicative of the presence of more than one sound source. Examples are concurrent signals coming from different locations [5], being in inharmonic relation with each other [6,7], or starting at slightly different times (*onset asynchrony* [8]).

Sound segregation by inharmonicity can be simulated by a complex tone (a fundamental with multiple upper harmonic partials, all in integer relation with the fundamental frequency) of which one partial is slightly mistuned upwards or downwards [9]. Perceptual reports suggest that the inharmonic partial pops out with pure-tone quality from the complex sound. This has been exploited for behavioral investigations of detecting mistuning and using it for concurrent sound segregation [6,7,9]. A recent study has adapted this behavioral approach for testing six-month-old infants, and has shown that infants at this age possess the capacity to detect mistuning [10].

Although this suggests an early developmental time-course for the processing of concurrent cues, one cannot infer that the capacity is innate given the widespread changes that auditory processing undergoes between birth and six months of age (e.g., [11–14]). Fortunately, concurrent sound segregation can be measured not only behaviorally (which would be difficult in neonates) but also by electrophysiological methods. In adults, concurrent segregation is accompanied by the elicitation of the object-related negativity (ORN), an early frontocentrally negative component of the event-related potential (ERP) [15].

ORN amplitude reliably reflects the extent of perceptual segregation in adults [15–17]. ORN and its magnetic equivalent (ORNm) are not only sensitive to mistuning [15,16,18–23] but also to other cues of concurrent segregation such as onset asynchrony [24,25], differences or discontinuities in location [5,26], dichotic pitch [17,27–31], and simulated echo [32,33].

The present study tests whether newborn infants can detect cues of concurrent sound segregation as reflected in ERP responses. Two different stimulus sets were used to probe concurrent sound segregation: stimuli based on mistuning-only that had been used for testing six-month olds in a previous study [10], and a cue combination of mistuning and onset asynchrony previously shown to elicit a robust ORN response in adults [34].

Materials and Methods

Participants. EEG was recorded from 41 healthy, full-term newborn infants (22 male, 19 female) during day 1 to 6 postpartum (mean 2.5 days, standard deviation [SD] 0.98 days). All infants had regular gestational age (39.1 ± 1.01 weeks), birth weight (3331.5 ± 347.1 g), and Apgar scores (consistently 9/10). Data from an additional 14 newborns were recorded, but discarded due to excessive electrical artifacts (see below). Informed consent was obtained from one or both parents. The experiment was carried out in a dedicated experimental room. The mother of the infant could opt to be present during the recording. The study was conducted in full accordance with the Declaration of Helsinki and all applicable national laws. It was approved by the Ethics Committee of Science and Research of the Medical Research Council of Hungary (ETT-TUKEB).

Apparatus and Stimuli. Complex tones were presented binaurally by ER-1 headphones (Etymotic Research Inc., Elk Grove Village, IL, USA) connected via sound tubes to selfadhesive ear-couplers (Natus Medical Inc., San Carlos, CA, USA) placed over the newborns' ears. Sounds were presented at an intensity level of 68 decibel sound pressure level (dB SPL) in the double-cue condition, and at 62 dB SPL in the single-cue condition. Sound delivery was controlled using the E-Prime stimulus presentation software (Psychology Software Tools, Inc., Pittsburgh, PA). Two different stimulus sets were presented in two stimulus blocks, whose order was counterbalanced across participants. Each block consisted of 400 complex tones (50% tuned, 50% manipulated) delivered in random order.

The stimulus set for the double-cue condition was taken from a previous adult study [34]. The *tuned* sound was a complex tone with a base frequency of 240 Hz, comprising the first five harmonics (i.e., 240 Hz, 480 Hz, 720 Hz, 960 Hz, 1200 Hz). All harmonics started in sine phase and were of the same amplitude. Sound duration was 250 ms, including raised-cosine onset and offset ramps of 5 ms duration each. The *manipulated* sound was identical to the tuned sound except for the 2^{nd} partial, which was both mistuned and delayed relative to the rest of the complex. Mistuning was created by shifting the 2^{nd} partial's frequency upwards by 8% to 518.4 Hz. Delay was created by starting the 2^{nd} partial 100 ms later than the other harmonics (but ending it at the same time). Sounds in this condition were delivered with a stimulus onset asynchrony (SOA) of 1100 ms. Overall duration of the stimulus block was 7.33 minutes.

The stimulus set for the single-cue condition was taken from a previous study in six-month-old infants [10]. The *tuned* sound was a complex tone with a base frequency of 240 Hz, comprising the first six harmonics (i.e., 240 Hz, 480 Hz, 720 Hz, 960 Hz, 1200 Hz, 1440 Hz). Harmonics were combined in random phase, and their intensity level decreased towards higher harmonics with a 6dB roll off per octave. Sound duration was 500 ms, including onset and offset ramps of 50 ms duration each. The *manipulated* sound was identical to the tuned sound except for the 3^{rd} partial, which was mistuned relative to the rest of the complex. Mistuning was created by shifting the 3^{rd} partial's frequency upwards by 8% to 777.6 Hz. Sounds in this condition were delivered with an SOA of 1500 ms. Overall duration of the stimulus block was 10 minutes.

Each infant underwent additional testing with different stimuli in the same recording session. These additional tests were always administered before the current ones and will be reported elsewhere. They included 13 minutes of homogenous tone sequences with occasional deviations of different types (as described in [35]) followed by 5 minutes of EEG recording without any auditory stimulus presentation.

EEG recording and analysis. EEG was continuously recorded with Ag/AgCl electrodes attached to 15 locations (Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, P3, Pz, P4) according to the International 10-20 system [36]. Recordings were made against a nose reference using a direct-coupled amplifier (V-Amp, Brain Products, Munich, Germany) at 24-bit resolution with a sampling rate of 1000 Hz. An additional electrode was attached near to the outer canthus of the left eye. This electrode was referenced offline against Fp1 to create a bipolar electrooculogram (EOG) channel for monitoring eye-movement-related artifacts.

Note that the broad spatial electrode coverage was chosen for the additional testing described above (cf. [35]), whereas a reduced set of frontocentral electrodes (F3, Fz, F4, C3, Cz, C4) was of interest for the present analysis. The spatial coverage of these channels is consistent with the frontocentral distribution of the ORN component in adults and children [20] as well as with typical auditory ERP responses in newborn infants [13]. The six frontocentral electrodes as well as the bipolarized EOG channel were thus selected for further analysis. Discarding the remaining channels (Fp2, F7, F8, T3, T4, P3, Pz, P4) from the analysis led to a substantial improvement in data quality, as each channel was found to contain a certain amount of unique artifacts (resulting, e.g., from changes in impedance during the recording) in at least one of the infants.

EEG data were analyzed with EEGlab [37]. Continuous data were filtered from 1-16 Hz using a finite-impulse response (FIR) band-pass filter (Kaiser-windowed, Kaiser $\beta = 5.65$, filter length 1813 points). Epochs were extracted from -200 ms to 500 ms relative to the onset of each sound. The 200 ms pre-stimulus interval was used for baseline correction. Epochs with an amplitude change exceeding 100 μV on any channel were rejected. Individual data were used for further analysis if there were at least 100 artifact-free epochs (50%) for each of the four stimulus types (double- vs. single-cue x tuned/manipulated sound). This criterion was fulfilled in 41 out of 55 datasets (75% of the newborn infants).

All artifact-free epochs were averaged separately for each condition and sound type per participant. Difference waves were formed by subtracting the ERPs elicited by tuned sounds from the ERPs elicited by their manipulated counterparts separately for the two conditions.

Based on visual inspection, two relevant component intervals were defined as 40-ms wide windows around the corresponding grand-average peak. The earlier window ranged from 160-200 ms and was termed *eN* (for *early negativity*), while the later window ranged from 360-400 ms and was termed *ORN* to indicate possible correspondence of the obtained effect to that in adults [34]. In the *eN* and ORN intervals, mean amplitudes elicited by the four stimulus types at the six electrodes (F3, Fz, F4, C3, Cz, C4) were measured from each babies' average ERPs.

Statistical analysis of mean ERP amplitudes was performed separately for the two component intervals, incorporating the 2x3 electrode grid. This resulted in a repeated-measures ANOVA with 4 factors: Condition (2 levels: double- vs. single-cue as specified above) x Sound type (2 levels: tuned vs. manipulated) x Anteriority (2 levels: frontal vs. central) x Laterality (3 levels: left vs. middle vs. right). Greenhouse-Geisser correction was performed when the assumption of sphericity was violated, and the ε correction factors are reported in those cases.

Results

Figures 1, 2 and 3 show that an early (*eN* latency range) negative deflection was elicited by mistuned-plus-delayed sounds (double-cue condition), but not by mistuned-only sounds (single-cue condition), while a late (ORN latency range) negative deflection was elicited by mistuned sounds in both conditions. These observations were statistically verified in repeated-measures ANOVAs. Full ANOVA results are given in Tables 1 and 2; here only significant effects are reported and only those that involve Sound type or Condition are followed upon.

--- insert Figures 1, 2, 3 at about here ---

During the early (*eN*) window, a significant 4-way interaction of Condition x Sound type x Anteriority x Laterality was observed ($p < .05$). The interaction was decomposed by separate 3-factorial ANOVAs (Sound type x Anteriority x Laterality) for the two conditions. No significant effects or interactions were observed for the single-cue condition (all p values > .240). For the double-cue condition, there was a significant 3-way interaction of Sound type x Anteriority x Laterality ($p < .01$). This was further decomposed by separate 2-factorial ANOVAs (Sound type x Laterality) for each level of Anteriority. There was no effect involving Sound type at central electrodes (both p values > .107). At frontal electrodes, there was a significant main effect of Sound type ($p < .05$) and a significant interaction between the effects of Sound type and Laterality ($p < .05$). This interaction was due to significant effects of Sound type at F3 and Fz (both p values < .05) but not at F4 ($p = .555$). The effects of Sound type resulted from more negative amplitudes for mistuned-plus-delayed sounds than for tuned sounds. Thus in summary, a negative deflection in an early window (160-200 ms) was elicited by mistuned-plus-delayed sounds in the double-cue condition, but not by mistuned-only sounds in the single-cue condition (see Table 1 for full ANOVA results).

--- insert Table 1 at about here ---

During the late (ORN) window, a significant main effect of Sound type was observed ($p < .01$) that was not different between conditions (all p values for interactions involving Sound type > .204). The only significant finding involving the Condition factor was an interaction of the effects of Condition and Anteriority ($p < .01$), caused by more negative amplitudes at central channels in the double-cue than in the single-cue condition, independent of Sound type (see Table 2 for full ANOVA results). To prospectively test whether the main effect of Sound type would be statistically reliable in either condition alone, we repeated the ANOVA *[sound type* *(tuned vs. mistuned) x anteriority (frontal vs. central) x laterality (left vs. middle vs. right]* separately for the two conditions. There was a significant main effect of sound type in the double-cue condition [$F(1,40) = 4.884$, $p = 0.033$] and a tendency for the same effect in the single-cue condition [$F(1,40) = 3.748$, $p = 0.060$]. These similar patterns correspond with the non-significant interaction of the effects of Sound type and Condition in the omnibus ANOVA ($p = .919$), ensuring that the main effect of Sound type was not carried by one of the conditions alone.

--- insert Table 2 at about here ---

Discussion

The present study was designed to investigate whether newborn infants may be capable of segregating concurrent sounds based on inharmonic relations between different frequency components (*mistuning*) or with additional support from delaying the mistuned component relative to others (*onset asynchrony*). Results of the electrophysiological measurements suggest that at least some of the cues underlying concurrent sound segregation are indeed processed at birth. The brain responses elicited by the manipulated sounds (with mistuned-only or mistuned-plus-delayed partials) markedly differed from those elicited by the corresponding tuned sounds.

The ERP differences between tuned and manipulated sounds appeared in two distinct latency ranges. In the early latency range (160-200 ms), a frontally dominant negative deflection was elicited by mistuned-plus-delayed sounds, but not by mistuned-only sounds. Although this corresponds to the typical ORN latency in adults when the mistuned partial starts with sound onset [15,19], the present negativity cannot yet be a mistuning-related ORN because the mistuned partial did not start until 100 ms after sound onset. With identical stimulus material, a double-peaked ERP response to mistuned-plus-delayed sounds was also observed in adults [34]. The early deflection, interpreted as an N1 effect in the adult data, probably reflects the automatic registration and processing of the delayed sound component. Although no analogue to the adult

N1 is present in neonatal ERP responses, it is possible that the ERP response reflecting the detection of delay in newborns is a subcomponent of the multi-source N1 wave that is present before the main contributors of the adult N1 reach maturity [14]. However, as the analogy with N1 cannot be tested with the present data, we used the more neutral term *eN* (early negativity).

The present *eN* results suggest that the detection of onset asynchrony as one important cue for concurrent sound segregation [24,25,34] is functional in newborns. This is consistent with other reports of adult-like temporal discrimination in neonates [38,39]. However, the data do not permit the inference that the negative deflection in the *eN* latency range reflects actual sound segregation based on onset asynchrony. It is possible that the onset asynchrony response is simply driven by the rise in acoustic energy provided by the onset of the delayed partial. In other words, we cannot exclude the possibility that the delay would be processed by the neonates' auditory system as a loudness increase in one and the same stimulus, rather than as a marker of the onset of a separate sound.

During the later latency range (360-400 ms), a negative deflection was elicited by both types of manipulated sounds (mistuned-only and mistuned-plus-delayed). As it occurred in both conditions, this response is very likely related to processing the inharmonic relation between the mistuned partial and the rest of the harmonic complex. We interpret it to be an analogue of the adult ORN. Although it occurred about 200 ms later than the typical ORN latency range in adults [15,34], such shifts are not unusual in neonate ERPs [38]. It merits noting that in adults, ORN peaks later when the cues for concurrent sound segregation are weak; for instance, ORN was found between 320 and 360 ms (close to the latency range observed here) in a condition in which only 3% mistuning was employed [25]. It is possible that the present amount of mistuning (8%) was not very salient for the neonatal auditory system, whose frequency resolution is known to be coarser than that in adults [40]. Hence a delayed occurrence of ORN in neonates is reasonable.

On the other hand, the information provided by 8% mistuning was apparently not so weak that an additional cue (i.e., onset asynchrony) could have increased the ORN further. In fact, there was no difference in ORN amplitude or scalp distribution between the conditions with mistuning-only or mistuning-plus-delay. This is consistent with recent results suggesting that ORN shows properties of an all-or-none response [34]; that is, once it is elicited by a given cue of concurrent sound segregation, it cannot be further increased by providing additional congruent cues. The rationale of combining mistuning with onset asynchrony in one of the conditions had been to increase the chances of finding an ORN-like response in neonates, as adult studies have shown that the process indexed by ORN integrates information from different cues [41,42], and that the combination of two weak cues can lead to ORN elicitation in situations in which one of the cues alone would have been insufficient to elicit ORN [5,25]. Such congruent evidence was, however, not necessary with the present stimulus parameters: Mistuning-alone was sufficient, and ORN showed a redundancy effect (i.e., no further increase) with additional cues [5,34].

One note of caution is that it is difficult to strictly compare the two conditions in the present study because they differed not only in the presence or absence of delay but also in SOA, tone duration, as well as in the order number of the manipulated harmonic, all of which can affect the ORN response [15,23]. Yet the occurrence of similar responses despite these differences can also be regarded as a strength in that it supports the generalizability of the observed ORN response in newborns. The design of the present study was not optimized towards a comparison between the two conditions, but towards a comparison with prior infant [10] and adult [34] data. On the basis of the present successful demonstration of an adult-like ORN response in neonates, future studies should address the properties of this response more systematically to derive a developmental trajectory of the processes underlying concurrent sound segregation. Similar to previous studies in children, this would include manipulations of the amount of mistuning [10,20] as well as testing whether other cues of concurrent sound segregation lead to ORN-like responses as well [31]. It would also be informative to use a version of the paradigm based on concurrent vowel segregation [42], for which ageing-related deteriorations have been demonstrated [43], suggesting a possible relation with speech processing difficulties that might prove diagnostically relevant in newborn infants.

The present demonstration of processing the cues for concurrent sound *segregation* at birth creates an interesting dissociation with concurrent sound *integration*, which had previously been shown to emerge only between three and four months of age [44]. It may appear contradictory that the infant should be able to detect an inharmonic partial earlier than being able to integrate harmonic frequency components into a single pitch percept. However, concurrent integration is indeed assumed to be a higher-level function in the auditory pathway [45,46] than concurrent segregation [47].

One may object that the ORN-like response in neonates could simply indicate the detection of an inharmonic partial, which does not necessarily translate into the percept of two segregated sounds. Indeed, this argument is difficult to rule out with neonates whose perceptual experience cannot be directly assessed. However, the close correspondence between ORN elicitation and concurrent segregation in adults [48] makes it plausible that the present data indicate sound segregation rather than mere processing of the inharmonic relation. At the very least, the present data permit the inference that newborn infants are able to pick up inharmonicity as a major cue of segregating a mixture of sounds. Importantly, this speaks not only to the neonates' capacity for processing their auditory surroundings, but also to the automaticity of the ORN response, confirming the view that it is a largely attention-independent response [15,18,19,28].

Previous studies have demonstrated that sequential grouping [49] and sequential stream segregation [50] are also functional in newborn infants. Together with these findings, the present results suggest that newborn infants are able to pick up the cues underlying the two main processing mechanisms of auditory scene analysis: sequential and concurrent sound organization [2,3]. Hence from birth onwards, auditory object formation may function similarly to that observed in adults. In other words, we do not need to worry that a human speaker on top of radio background music constitutes a principally inseparable mixture for neonate listeners.

In conclusion, our results suggest that at least some of the cues supporting the segregation of concurrent sounds are processed at birth. Newborn infants differentially processed co-occurring auditory stimuli based on harmonic relations as reflected by ORN-like responses observed in the 360-400 ms latency range. Future studies should perform parametric manipulations of inharmonicity as well as other cues of concurrent sound segregation to assess the simultaneous sound organization capabilities of newborn infants in greater detail.

Acknowledgments

This work was funded by the Hungarian Scientific Research Fund (OTKA K101060 to I.W.), by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG Cluster of Excellence 1077 "Hearing4all"), by the German Academic Exchange Service (Deutscher Akademischer Austauschdienst [DAAD], Project 56265741), and by the Hungarian Scholarship Board (Magyar Ösztöndíj Bizottság [MÖB], Project 39589). The authors are grateful to Judit Roschéné Farkas for collecting the data, and to Prof. Dr. Laurel J. Trainor, McMaster University, Hamilton, Ontario, for providing the sound files from their study in infants to allow for a developmental comparison with identical stimulus material.

References

1. McDermott JH: The cocktail party problem. Curr Biol 2009;19:R1024–R1027.
2. Bregman AS: Auditory scene analysis. The perceptual organization of sound. Cambridge, MA, MIT Press, 1990.
3. Carlyon RP: How the brain separates sounds. Trends Cogn Sci 2004;8:465–471.
4. Snyder JS, Alain C: Toward a neurophysiological theory of auditory stream segregation. Psychol Bull 2007;133:780–799.

5. McDonald KL, Alain C: Contribution of harmonicity and location to auditory object formation in free field: Evidence from event-related brain potentials. J Acoust Soc Am 2005;118:1593–1604.

6. Moore BCJ, Peters RW, Glasberg BR: Thresholds for the detection of inharmonicity in complex tones. J Acoust Soc Am 1985;77:1861–1867.

7. Moore BCJ, Glasberg BR, Peters RW: Thresholds for hearing mistuned partials as separate tones in harmonic complexes. J Acoust Soc Am 1986;80:479–483.

8. Bidet-Caulet A, Fischer C, Bauchet F, Aguera P-E, Bertrand O: Neural substrate of concurrent sound perception: direct electrophysiological recordings from human auditory cortex. Front Hum Neurosci 2008;1:5.

9. Hartmann WM, McAdams S, Smith BK: Hearing a mistuned harmonic in an otherwise periodic complex tone. J Acoust Soc Am 1990;88:1712–1724.

10. Folland NA, Butler BE, Smith NA, Trainor LJ: Processing simultaneous auditory objects: Infants' ability to detect mistuning in harmonic complexes. J Acoust Soc Am 2012;131:993–997.

11. He C, Hotson L, Trainor LJ: Maturation of cortical mismatch responses to occasional pitch change in early infancy: Effects of presentation rate and magnitude of change. Neuropsychologia 2009;47:218–229.

12. He C, Hotson L, Trainor LJ: Development of infant mismatch responses to auditory pattern changes between 2 and 4 months old. Eur J Neurosci 2009;29:861–867.

13. Kushnerenko EV, Van den Bergh BRH, Winkler I: Separating acoustic deviance from novelty during the first year of life: a review of event-related potential evidence. Front Psychol 2013;4:595.

14. Wunderlich JL, Cone-Wesson BK, Shepherd R: Maturation of the cortical auditory evoked potential in infants and young children. Hear Res 2006;212:185–202.

15. Alain C, Arnott SR, Picton TW: Bottom-up and top-down influences on auditory scene analysis: Evidence from event-related brain potentials. J Exp Psychol Hum Percept Perform 2001;27:1072–1089.

16. Alain C, McDonald KL: Age-related differences in neuromagnetic brain activity underlying concurrent sound perception. J Neurosci 2007;27:1308–1314.

17. Clapp WC, Johnson BW, Hautus MJ: Graded cue information in dichotic pitch: effects on event-related potentials. NeuroReport 2007;18:365–368.

18. Alain C, Schuler BM, McDonald KL: Neural activity associated with distinguishing concurrent auditory objects. J Acoust Soc Am 2002;111:990–995.

19. Alain C, Izenberg A: Effects of attentional load on auditory scene analysis. J Cogn Neurosci 2003;15:1063–1073.

20. Alain C, Theunissen EL, Chevalier H, Batty M, Taylor MJ: Developmental changes in distinguishing concurrent auditory objects. Cogn Brain Res 2003;16:210–218.

21. Hiraumi H, Nagamine T, Morita T, Naito Y, Fukuyama H, Ito J: Right hemispheric predominance in the segregation of mistuned partials. Eur J Neurosci 2005;22:1821–1824.

22. Bendixen A, Jones SJ, Klump GM, Winkler I: Probability dependence and functional separation of the object-related and mismatch negativity event-related potential components. NeuroImage 2010;50:285–290.

23. Alain C, McDonald KL, Van Roon P: Effects of age and background noise on processing a mistuned harmonic in an otherwise periodic complex sound. Hear Res 2012;283:126–135.

24. Lipp R, Kitterick PT, Summerfield AQ, Bailey PJ, Paul-Jordanov I: Concurrent sound segregation based on inharmonicity and onset asynchrony. Neuropsychologia 2010;48:1417–1425.

25. Weise A, Schröger E, Bendixen A: The processing of concurrent sounds based on inharmonicity and asynchronous onsets: An object-related negativity (ORN) study. Brain Res 2012;1439:73–81.

26. Butcher A, Govenlock SW, Tata MS: A lateralized auditory evoked potential elicited when auditory objects are defined by spatial motion. Hear Res 2011;272:58–68.

27. Johnson BW, Hautus M, Clapp WC: Neural activity associated with binaural processes for the perceptual segregation of pitch. Clin Neurophysiol 2003;114:2245–2250.

28. Hautus MJ, Johnson BW: Object-related brain potentials associated with the perceptual segregation of a dichotically embedded pitch. J Acoust Soc Am 2005;117:275–280.

29. Johnson BW, Hautus MJ, Duff DJ, Clapp WC: Sequential processing of interaural timing differences for sound source segregation and spatial localization: Evidence from event-related cortical potentials. Psychophysiology 2007;44:541–551.

30. Johnson BW, Hautus MJ: Processing of binaural spatial information in human auditory cortex: Neuromagnetic responses to interaural timing and level differences. Neuropsychologia 2010;48:2610–2619.

31. Brock J, Bzishvili S, Reid M, Hautus M, Johnson BW: Brief report: Atypical neuromagnetic responses to illusory auditory pitch in children with Autism spectrum disorders. J Autism Dev Disord 2013;43:2726–2731.

32. Sanders LD, Joh AS, Keen RE, Freyman RL: One sound or two? Object-related negativity indexes echo perception. Percept Psychophys 2008;70:1558–1570.

33. Sanders LD, Zobel BH, Freyman RL, Keen RE: Manipulations of listeners' echo perception are reflected in event-related potentials. J Acoust Soc Am 2011;129:301–309.

34. Kocsis Z, Winkler I, Szalárdy O, Bendixen A: Effects of multiple congruent cues on concurrent sound segregation during passive and active listening: An event-related potential (ERP) study. Biol Psychol 2014;100:20–33.

35. Otte RA, Winkler I, Braeken MAKA, Stekelenburg JJ, van der Stelt O, Van den Bergh BRH: Detecting violations of temporal regularities in waking and sleeping two-month-old infants. Biol Psychol 2013;92:315–322.

36. Jasper HH: The ten-twenty electrode system of the International Federation. Electroencephalogr Clin Neurophysiol 1958;10:371–375.

37. Delorme A, Makeig S: EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods 2004;134:9–21.

38. Cheour M, Kushnerenko E, Čeponienė R, Fellman V, Näätänen R: Electric brain responses obtained from newborn infants to changes in duration in complex harmonic tones. Dev Neuropsychol 2002;22:471–479.

39. Kushnerenko E, Čeponienė R, Fellman V, Huotilainen M, Winkler I: Event-related potential correlates of sound duration: similar pattern from birth to adulthood. NeuroReport 2001;12:3777–3781.

40. Novitski N, Huotilainen M, Tervaniemi M, Näätänen R, Fellman V: Neonatal frequency discrimination in 250–4000-Hz range: Electrophysiological evidence. Clin Neurophysiol 2007;118:412–419.

41. Hautus MJ, Johnson BW, Colling LJ: Event-related potentials for interaural time differences and spectral cues: NeuroReport 2009;20:951–956.

42. Du Y, He Y, Ross B, Bardouille T, Wu X, Li L, Alain C: Human auditory cortex activity shows additive effects of spectral and spatial cues during speech segregation. Cereb Cortex 2011;21:698–707.

43. Snyder JS, Alain C: Age-related changes in neural activity associated with concurrent vowel segregation. Cogn Brain Res 2005;24:492–499.

44. He C, Trainor LJ: Finding the pitch of the missing fundamental in infants. J Neurosci 2009;29:7718–7722.

45. Penagos H, Melcher JR, Oxenham AJ: A neural representation of pitch salience in nonprimary human auditory cortex revealed with functional magnetic resonance imaging. J Neurosci 2004;24:6810–6815.

46. Schönwiesner M, Zatorre RJ: Depth electrode recordings show double dissociation between pitch processing in lateral Heschl's gyrus and sound onset processing in medial Heschl's gyrus. Exp Brain Res 2008;187:97–105.

47. Dyson BJ, Alain C: Representation of concurrent acoustic objects in primary auditory cortex. J Acoust Soc Am 2004;115:280–288.

48. Alain C: Breaking the wave: Effects of attention and learning on concurrent sound perception. Hear Res 2007;229:225–236.

49. Stefanics G, Háden G, Huotilainen M, Balázs L, Sziller I, Beke A, Fellman V, Winkler I: Auditory temporal grouping in newborn infants. Psychophysiology 2007;44:697–702.

50. Winkler I, Kushnerenko E, Horváth J, Čeponienė R, Fellman V, Huotilainen M, Näätänen R, Sussman ES: Newborn infants can organize the auditory world. Proc Natl Acad Sci U S A 2003;100:11812–11815.

Tables

Table 1: Results of the omnibus and follow-up repeated-measures ANOVAs of the mean amplitudes measured during the *eN* window (160-200 ms).

Factor	Effect on mean ERP amplitudes
Omnibus ANOVA (Condition x Sound type x Anteriority x Laterality)	
Condition *(2 levels: single vs. double)*	$F(1,40) = 0.045$, $p = .833$, $\eta^2 = 0.001$
Sound type *(2 levels: tuned vs. mistuned)*	$F(1,40) = 2.925$, $p = .095$, $\eta^2 = 0.068$
Anteriority (2 *levels: frontal vs. central)*	$\boldsymbol{F(1,40) = 6.566}$, $\boldsymbol{p = .014}$, $\boldsymbol{\eta^2 = 0.141}$
Laterality *(3 levels: left / middle / right)*	$F(2,80) = 2.099$, $p = .137$, $\eta^2 = 0.050$, $\varepsilon_{GG} = 0.865$
Condition x Sound type	$F(1,40) = 1.169$, $p = .286$, $\eta^2 = 0.028$
Condition x Anteriority	$F(1,40) = 0.984$, $p = .327$, $\eta^2 = 0.024$
Sound type x Anteriority	$F(1,40) = 0.000$, $p = .991$, $\eta^2 = 0.000$
Condition x Sound type x Anteriority	$F(1,40) = 0.026$, $p = .872$, $\eta^2 = 0.001$
Condition x Laterality	$F(2,80) = 2.060$, $p = .134$, $\eta^2 = 0.049$
Sound type x Laterality	$F(2,80) = 0.580$, $p = .529$, $\eta^2 = 0.014$, $\varepsilon_{GG} = 0.814$
Condition x Sound type x Laterality	$F(2,80) = 0.297$, $p = .744$, $\eta^2 = 0.007$
Anteriority x Laterality	$F(2,80) = 2.568$, $p = .094$, $\eta^2 = 0.060$, $\varepsilon_{GG} = 0.826$
Condition x Anteriority x Laterality	$F(2,80) = 0.516$, $p = .599$, $\eta^2 = 0.013$
Sound type x Anteriority x Laterality	$F(2,80) = 1.853$, $p = .163$, $\eta^2 = 0.044$
Condition x Sound type x Anteriority x Laterality	$\boldsymbol{F(2,80) = 3.873}$, $\boldsymbol{p = .025}$, $\boldsymbol{\eta^2 = 0.088}$
Follow-up ANOVA: Sound type x Anteriority x Laterality for the double-cue condition	
Sound type *(2 levels: tuned vs. mistuned)*	$F(1,40) = 3.960$, $p = .053$, $\eta^2 = 0.090$
Anteriority (2 *levels: frontal vs. central)*	$\boldsymbol{F(1,40) = 6.016}$, $\boldsymbol{p = .019}$, $\boldsymbol{\eta^2 = 0.131}$
Laterality *(3 levels: left / middle / right)*	$\boldsymbol{F(2,80) = 4.286}$, $\boldsymbol{p = .017}$, $\boldsymbol{\eta^2 = 0.097}$
Sound type x Anteriority	$F(1,40) = 0.014$, $p = .906$, $\eta^2 = 0.000$
Sound type x Laterality	$F(2,80) = 0.776$, $p = .464$, $\eta^2 = 0.019$
Anteriority x Laterality	$F(2,80) = 2.056$, $p = .146$, $\eta^2 = 0.049$, $\varepsilon_{GG} = 0.797$
Sound type x Anteriority x Laterality	$\boldsymbol{F(2,80) = 5.110}$, $\boldsymbol{p = .008}$, $\boldsymbol{\eta^2 = 0.113}$
Follow-up ANOVA: Sound type x Anteriority x Laterality for the single-cue condition	
Sound type *(2 levels: tuned vs. mistuned)*	$F(1,40) = 0.348$, $p = .558$, $\eta^2 = 0.009$
Anteriority (2 *levels: frontal vs. central)*	$F(1,40) = 1.420$, $p = .240$, $\eta^2 = 0.034$
Laterality *(3 levels: left / middle / right)*	$F(2,80) = 0.204$, $p = .816$, $\eta^2 = 0.005$
Sound type x Anteriority	$F(1,40) = 0.014$, $p = .905$, $\eta^2 = 0.000$
Sound type x Laterality	$F(2,80) = 0.193$, $p = .787$, $\eta^2 = 0.005$, $\varepsilon_{GG} = 0.841$
Anteriority x Laterality	$F(2,80) = 0.914$, $p = .405$, $\eta^2 = 0.022$
Sound type x Anteriority x Laterality	$F(2,80) = 0.423$, $p = .616$, $\eta^2 = 0.010$, $\varepsilon_{GG} = 0.819$
Follow-up ANOVA: Sound type x Laterality for frontal channels in the double-cue condition	
Sound type *(2 levels: tuned vs. mistuned)*	$\boldsymbol{F(1,40) = 4.493}$, $\boldsymbol{p = .040}$, $\boldsymbol{\eta^2 = 0.101}$
Laterality *(3 levels: left / middle / right)*	$F(2,80) = 2.569$, $p = .096$, $\eta^2 = 0.060$, $\varepsilon_{GG} = 0.801$
Sound type x Laterality	$\boldsymbol{F(2,80) = 3.971}$, $\boldsymbol{p = .023}$, $\boldsymbol{\eta^2 = 0.090}$
Follow-up ANOVA: Sound type x Laterality for central channels in the double-cue condition	
Sound type *(2 levels: tuned vs. mistuned)*	$F(1,40) = 2.722$, $p = .107$, $\eta^2 = 0.064$
Laterality *(3 levels: left / middle / right)*	$\boldsymbol{F(2,80) = 3.766}$, $\boldsymbol{p = .027}$, $\boldsymbol{\eta^2 = 0.086}$
Sound type x Laterality	$F(2,80) = 0.323$, $p = .725$, $\eta^2 = 0.008$

Note: The ANOVA factor structure is repeated in brackets behind each main effect for clarity. Results are reported with the partial η^2 effect size measure. The Greenhouse-Geisser epsilon correction value is indicated in those cases for which the sphericity assumption was violated. Significant results are marked in bold font.

Table 2: Results of the repeated-measures ANOVA of the mean amplitudes measured during the ORN window (360-400 ms).

Factor	Effect on mean ERP amplitudes
Condition *(2 levels: single vs. double)*	$F(1,40) = 0.484$, $p = .491$, $\eta^2 = 0.012$
Sound type *(2 levels: tuned vs. mistuned)*	$\boldsymbol{F(1,40) = 8.965}$, $\boldsymbol{p = .005}$, $\boldsymbol{\eta^2 = 0.183}$
Anteriority (2 *levels: frontal vs. central)*	$\boldsymbol{F(1,40) = 7.997}$, $\boldsymbol{p = .007}$, $\boldsymbol{\eta^2 = 0.167}$
Laterality *(3 levels: left vs. middle vs. right)*	$F(2,80) = 1.720$, $p = .186$, $\eta^2 = 0.041$
Condition x Sound type	$F(1,40) = 0.011$, $p = .919$, $\eta^2 = 0.000$
Condition x Anteriority	$\boldsymbol{F(1,40) = 8.340}$, $\boldsymbol{p = .006}$, $\boldsymbol{\eta^2 = 0.173}$
Sound type x Anteriority	$F(1,40) = 1.665$, $p = .204$, $\eta^2 = 0.040$
Condition x Sound type x Anteriority	$F(1,40) = 0.027$, $p = .870$, $\eta^2 = 0.001$
Condition x Laterality	$F(2,80) = 0.819$, $p = .445$, $\eta^2 = 0.020$
Sound type x Laterality	$F(2,80) = 0.979$, $p = .380$, $\eta^2 = 0.024$
Condition x Sound type x Laterality	$F(2,80) = 0.156$, $p = .856$, $\eta^2 = 0.004$
Anteriority x Laterality	$\boldsymbol{F(2,80) = 3.734}$, $\boldsymbol{p = .028}$, $\boldsymbol{\eta^2 = 0.085}$
Condition x Anteriority x Laterality	$F(2,80) = 1.095$, $p = .339$, $\eta^2 = 0.027$
Sound type x Anteriority x Laterality	$F(2,80) = 0.001$, $p = .999$, $\eta^2 = 0.000$
Condition x Sound type x Anteriority x Laterality	$F(2,80) = 1.167$, $p = .317$, $\eta^2 = 0.028$

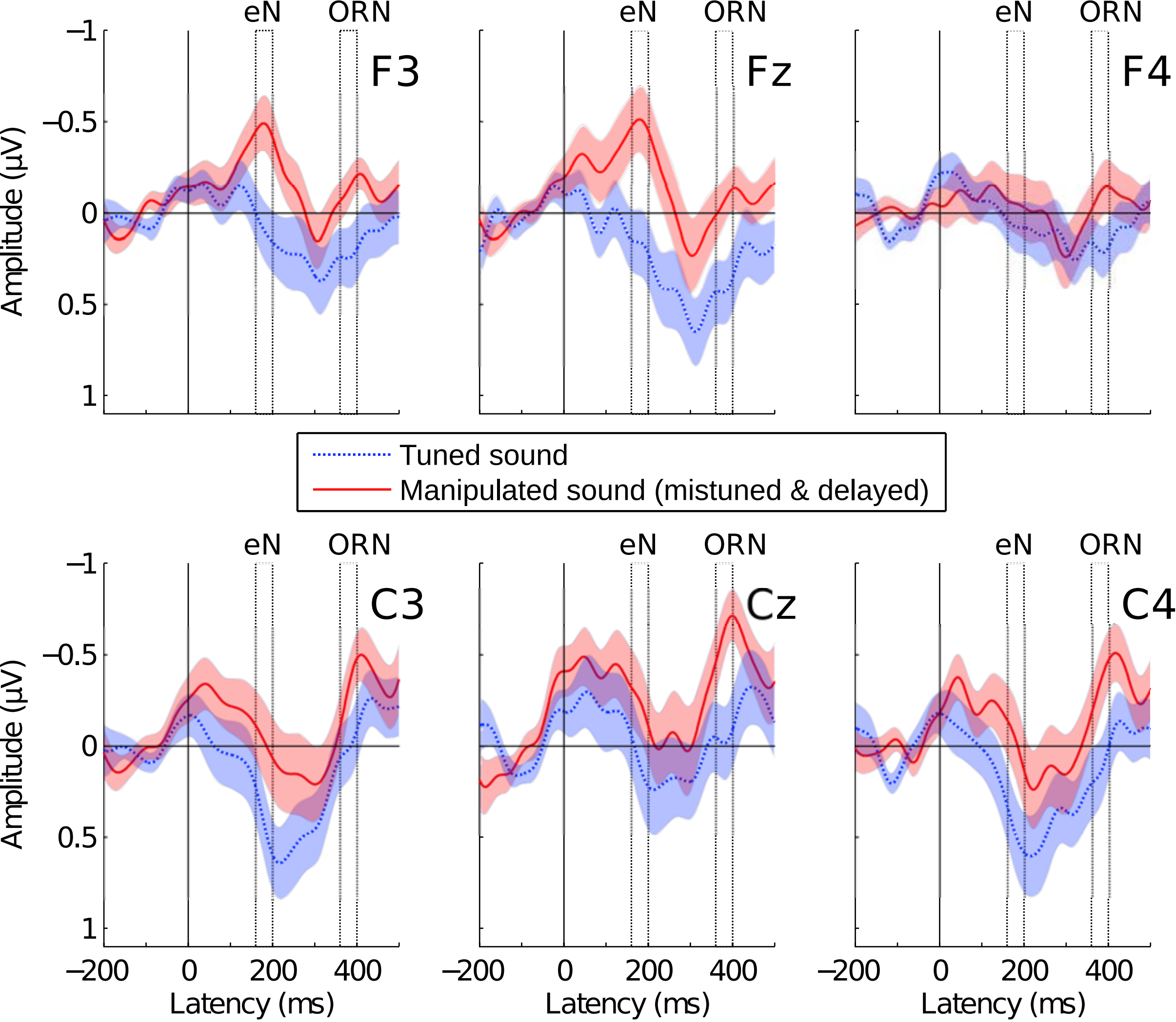
Note: The ANOVA factor structure is repeated in brackets behind each main effect for clarity. Results are reported with the partial η^2 effect size measure. The sphericity assumption was not violated in any case. Significant results are marked in bold font.

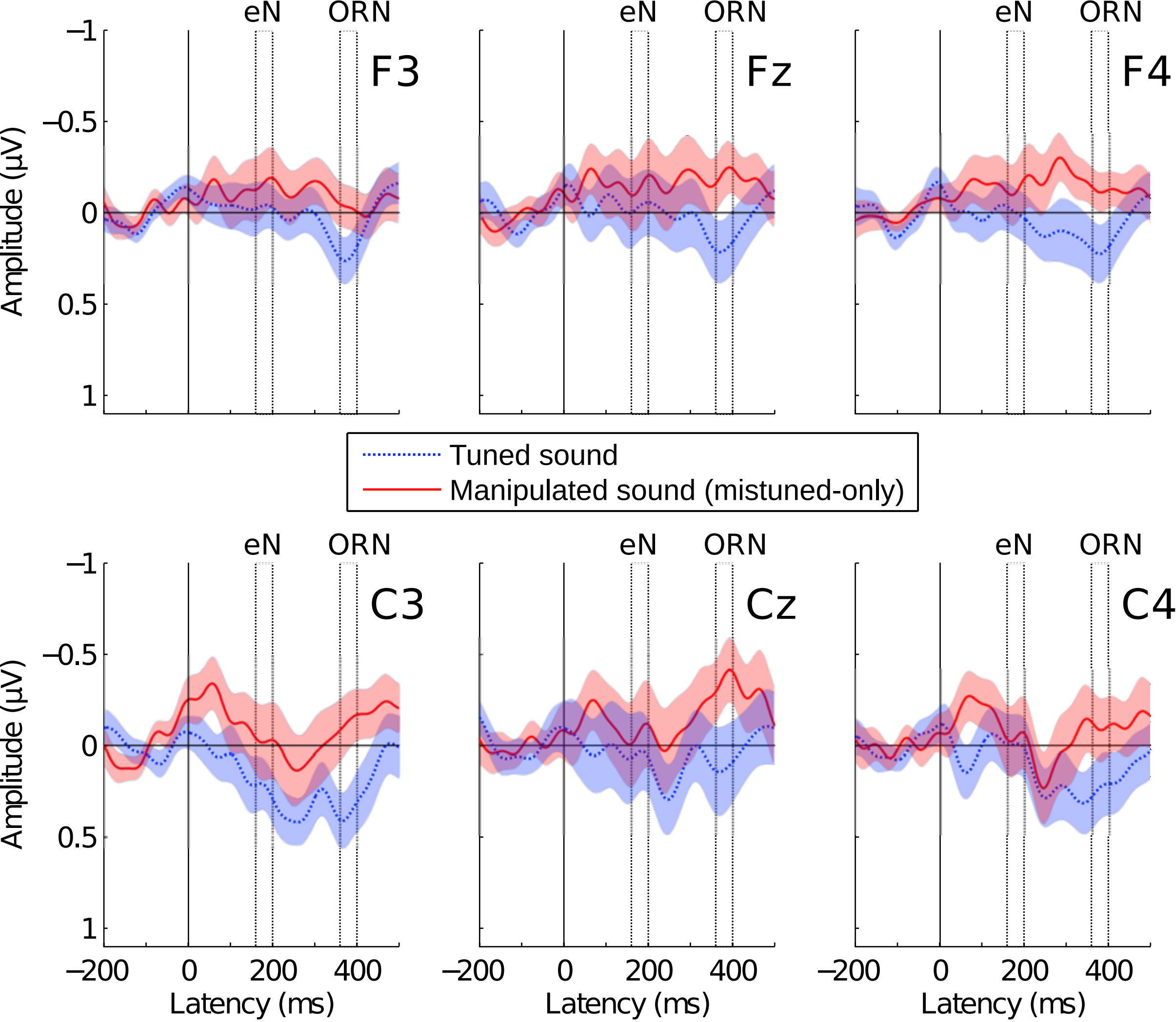
Figure Legends

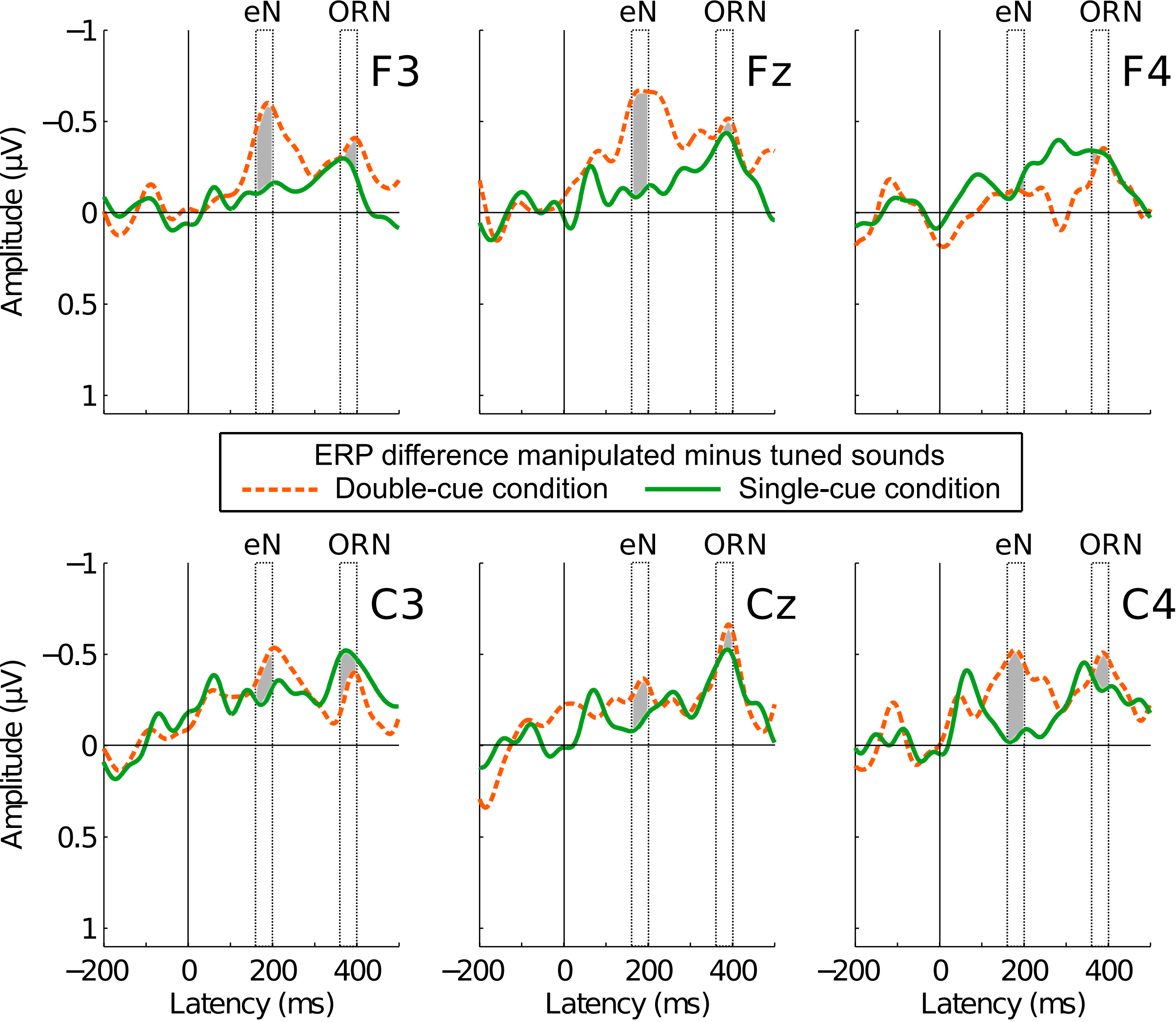
Figure 1: Results of the double-cue condition. Grand-average ERPs (N=41) elicited by tuned sounds (blue dotted line) are contrasted with those elicited by mistuned-plus-delayed sounds (red solid line). Mean ERPs are plotted together with the standard error of mean for assessing the inter-individual variation and noise level in the data.

Figure 2: Results of the single-cue condition. Grand-average ERPs (N=41) elicited by tuned sounds (blue dotted line) are contrasted with those elicited by mistuned sounds (red solid line). Mean ERPs are plotted together with the standard error of mean for assessing the inter-individual variation and noise level in the data.

Figure 3: Manipulated-minus-tuned difference waveforms of the double- and single-cue conditions compared (grand-average ERPs, N=41). The areas shaded in gray reflect amplitude differences between the two conditions. Note that there are significant differences between conditions in the *eN* latency range, but not in the ORN latency range.


eN
ORN
F3
Fz
F4
C3
Cz
C4
Amplitude (µV)
−1
−0.5
0
0.5
1
Tuned sound
Manipulated sound (mistuned & delayed)
−200
0
200
400
Latency (ms)


eN
ORN
F3
Fz
F4
C3
Cz
C4
Amplitude (μV)
Latency (ms)
Tuned sound
Manipulated sound (mistuned-only)
−1
−0.5
0
0.5
1
−200
0
200
400



F3
Fz
F4
C3
Cz
C4
eN
ORN
Amplitude (µV)
Latency (ms)
−1
−0.5
0
0.5
1
−200
0
200
400
ERP difference manipulated minus tuned sounds
Double-cue condition
Single-cue condition