



UvA-DARE (Digital Academic Repository)

Detecting the temporal structure of sound sequences in newborn infants

Háden, G.P.; Honing, H.; Török, M.; Winkler, I.

DOI

[10.1016/j.ijpsycho.2015.02.024](https://doi.org/10.1016/j.ijpsycho.2015.02.024)

Publication date

2015

Document Version

Final published version

Published in

International Journal of Psychophysiology

License

Article 25fa Dutch Copyright Act

[Link to publication](#)

Citation for published version (APA):

Háden, G. P., Honing, H., Török, M., & Winkler, I. (2015). Detecting the temporal structure of sound sequences in newborn infants. *International Journal of Psychophysiology*, *96*(1), 23-28. <https://doi.org/10.1016/j.ijpsycho.2015.02.024>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.



Research Article

Detecting the temporal structure of sound sequences in newborn infants ^{☆, ☆ ☆}



Gábor P. Háden ^{a,*}, Henkjan Honing ^{b,c}, Miklós Török ^d, István Winkler ^{a,e}

^a Institute of Cognitive Neuroscience and Psychology, Research Centre for Natural Sciences, Hungarian Academy of Sciences, Magyar Tudósok körútja 2, H-1117 Budapest, Hungary

^b Institute for Logic, Language and Computation, University of Amsterdam, Science Park 107, 1098 XH Amsterdam, The Netherlands

^c Amsterdam Brain and Cognition, Nieuwe Achtergracht 129, 1018 WS Amsterdam, The Netherlands

^d Department of Obstetrics–Gynaecology and Perinatal Intensive Care Unit, Military Hospital, Podmaniczky u. 111, H-1062 Budapest, Hungary

^e Institute of Psychology, University of Szeged, Egyetem u. 2, H-6722 Szeged, Hungary

ARTICLE INFO

Article history:

Received 30 September 2014

Received in revised form 16 February 2015

Accepted 17 February 2015

Available online 24 February 2015

Keywords:

Event related potentials

Development

Auditory processing

Presentation rate

Onset

Offset

ABSTRACT

Most high-level auditory functions require one to detect the onset and offset of sound sequences as well as registering the rate at which sounds are presented within the sound trains. By recording event-related brain potentials to onsets and offsets of tone trains as well as to changes in the presentation rate, we tested whether these fundamental auditory capabilities are functional at birth. Each of these events elicited significant event-related potential components in sleeping healthy neonates. The data thus demonstrate that the newborn brain is sensitive to these acoustic features suggesting that infants are geared towards the temporal aspects of segregating sound sources, speech and music perception already at birth.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Extracting temporal regularities from sound sequences and detecting their violations are fundamental capabilities for correctly perceiving objects in the acoustic environment (Griffiths and Warren, 2004; Winkler et al., 2009a), including interpreting speech and music (Honing, 2013; Patel, 2008) and also form the basis of synchronized communication with others (Jaffe et al., 2001; Jungers et al., 2002). Speech dynamics provide information about the emotional state and intents of the speaker, and structure information within and between sentences and allow marking agreement and turn taking (O'Connell and Kowal, 2008). In music, temporal information defines beat, metrical structure, and tempo, allows the extraction of expressive timing in a performance, helps coordination between players, and conveys

emotions (Honing, 2013). Because of its fundamental nature, one may assume that this capability appears early during infancy. The current study tested in newborn infants the detection of the three most basic temporal features of sound sequences: onset, presentation rate change, and offset.

Behavioral studies testing the processing of temporal features of sound sequences found that 2-month old infants detect 15% tempo accelerations in isochronous sequences at the base rate of 600 ms inter-onset interval (IOI), but not at faster or slower IOIs (Baruch and Drake, 1997). They also coordinate movements with the tempo of external sounds (Bobin-Bègue et al., 2006). By 6 months, infants form long-term memories of tempo (Trainor et al., 2004) and by 9 months they can distinguish between happy and sad music (Flom et al., 2008). The sensitivity to temporal features of the stimulation was also proposed to be an important predictor of later performance in tests of verbal development in young infants (Benasich and Tallal, 2002; Chonchaiya et al., 2013) though the mechanisms underlying these effects are unclear (Protopapas, 2014). Much less is known about auditory temporal processing in newborns. Previous studies showed that neonates segregate interleaved tonal sequences by pitch (Winkler et al., 2003), prefer infant-directed to adult-directed speech and singing (Cooper and Aslin, 1990; Masataka, 1999), and discriminate languages based on rhythmic class (Nazzi et al., 1998; Ramus et al., 2000; Nazzi and Ramus, 2003). These capabilities probably involve detecting auditory temporal cues.

Abbreviations: ANOVA, Analysis of variance; EEG, Electroencephalogram; ERP, Event-related potential; IOI, Inter-onset interval; MMN, Mismatch negativity.

[☆] Author contributions: GPH, HH, and IW designed the experiment, GPH and MT oversaw the data collection, GPH analyzed the data, and GPH, HH, MT, and IW wrote the paper ^{☆☆} The authors declare no conflict of interests.

* Corresponding author at: Institute of Cognitive Neuroscience and Psychology, Research Centre for Natural Sciences, Hungarian Academy of Sciences, H-1519 Budapest, P. O. Box 286, Hungary. Tel.: +36 1 382 6808; fax: +36 1 3826295.

E-mail addresses: haden.gabor@ttk.mta.hu (G.P. Háden), honing@uva.nl (H. Honing), miklostorok@vipmail.hu (M. Török), winkler.istvan@ttk.mta.hu (I. Winkler).

In neonates, sound processing can be tested by measuring event-related potentials (ERP) elicited by acoustic events. Two parallel event detector systems have been described in adults (Näätänen, 1990; Näätänen et al., 2011): one sensitive to sudden changes in sound energy and another triggered by violations of some regular feature of a sound sequence. The former is based on adaptation/refractoriness of afferent neurons (in adults, the N1 ERP response; Näätänen and Picton, 1987; May and Tiitinen, 2010), the latter probably on prediction errors in the brain (the MMN component; Näätänen and Picton, 1987; Garrido et al., 2009; Winkler, 2007). Although no true equivalent of either of these adult ERP responses has been obtained in newborn infants, neonatal ERP responses to large energy changes and violations of simple acoustic regularities have been described (Alho et al., 1990; for a review, see Kushnerenko et al., 2013). Previous ERP studies testing temporal features of sound sequences in young infants showed that occasional shortenings of a regular 300 ms long pre-stimulus interval to 100 ms are detected at 2 months of age (Otte et al., 2013; for similar results in 10-month olds, see Brannon et al., 2004, 2008) and that newborns can distinguish between the downbeat and other positions within a rhythmic sequence (Winkler et al., 2009a,b).

Thus, whereas we know that young infants use temporal cues while making complex linguistic and musical discriminations, the developmental origins of the underlying processing capabilities have not yet been established. Here we assess the developmental origins of processing the temporal structure of sound sequences by testing whether the neonate brain is sensitive to the onset and offset of sound trains that roughly estimate the structure of sentences or short musical phrases as well as to presentation rate change within them. To this end we compare responses elicited by significant events (onset and rate change) in our sound sequences to events that are physically the same but appear in a different context. We assume that the onset of the train elicits a response; however if the rate change elicits a response it signals that the change was indeed detected. Finally if we see a response at the offset of the train in a position where the continuation of the train could be expected we can assume that the offset itself was detected.

2. Methods

ERPs were recorded from 30 (16 male) healthy, full-term newborn infants during day 1–3 postpartum. The mean gestational age was 39.7 weeks (SD = 1.00), birth weight was 3450 g (SD = 372.46), and the average Apgar score was 9/9.8 (SD = 0.52/0.48). An additional 7 (2 male) participants were recorded, but discarded due to excessive electrical artifacts (<100 artefact free epochs per condition). Informed consent was obtained from one or both parents. The mother of the infant could opt to be present during the recording. The study was conducted in full accordance with the World Medical Association Declaration of Helsinki and all applicable national laws and it was approved by the Medical Research Council—Committee of Scientific and Research Ethics (ETT-TUKEB), Hungary. The experiment was carried out in a dedicated experimental room at the Department of Obstetrics–Gynaecology and Perinatal Intensive Care Unit, Military Hospital, Budapest, Hungary.

Trains of complex tones (Fig. 1) uniform within but varying in pitch (F0) across trains (8 different pitches taken from the C major scale: C3, D3, E3, F3, G3, A3, B3, and C4, viz. 130.80, 147.15, 163.50, 173.96, 196.20, 217.13, 245.25, and 261.60 Hz, respectively) were delivered to the infants at ~65dB_{SPL}. Each tone consisted of the F0 and its first five harmonics, the spectral power of the higher harmonics being set at 1:2, 1:3, 1:4, 1:5, and 1:6, with respect to that of the F0 component. Tone duration was 50 ms, including 5 ms rise and 5 ms fall times (raised cosine ramps). 170 trains were presented in two stimulus blocks. For each train, a pitch was selected randomly (with equal probability; no repetition allowed). Trains consisted of 8–24 (randomly selected, equal probability) tone repetitions and a silent interval. The N element trains were split into two parts: the first part consisting of N/2 (rounded

down) + 0/1 (random) elements, the second part consisting of N-n_(first part) elements. Tones in the first part of the train were presented at the “slow” rate (average IOI = 200 ms, 150 ms offset to onset), and in the second part at the “fast” rate (average IOI = 100 ms; 50 ms offset to onset), followed by a silent interval (average IOI 1050 ms, 1000 ms offset to onset). All time intervals were taken from normal distributions with a standard deviation of 5%. The amount of jitter is below the adult JND for tempo discrimination (Quené, 2007; Grondin et al., 2011) and was chosen to dampen steady state responses arising from the fast stimulation. The length of individual trains was about 1.2–3.6 s. Sounds were presented binaurally using the E-Prime stimulus presentation software (Psychology Software Tools, Inc., Pittsburgh, PA, USA) via ER-1 headphones (Etymotic Research Inc., Elk Grove Village, IL, USA) connected via sound tubes to self-adhesive ear-couplers (Natus Medical Inc., San Carlos, CA, USA) placed over the infants’ ears.

EEG was recorded with Ag/AgCl electrodes attached to the F3, Fz, F4, C3, Cz, and C4 locations (international 10–20 system) using a direct-coupled amplifier (V-Amp, Brain Products GmbH, Munich, Germany) at 24-bit resolution and a sampling rate of 250 Hz or 1000 Hz (1000 Hz sampling rate was used for 16 participants; due to experimenter error, these recordings were off-line down-sampled to 250 Hz). The reference and ground electrodes were attached to the nose and the forehead respectively. Signals were off-line filtered between 1 and 30 Hz and epochs from –100 to 500 ms with respect to the event onset (tone or expected tone, see next paragraph) were extracted for each sound. The 100 ms pre-stimulus interval served as the baseline for amplitude measurements and illustrations. Epochs with an absolute voltage change outside the 0.1–100 μ V range throughout the epoch were rejected from the analyses as artefacts. Data from infants with less than 100 artefact-free epochs per condition were dropped from the analyses. The mean number of epochs and per condition is given in Table 1.

Responses were measured at the “train onset” (the first element of the train), at “presentation rate change” (the first element after a short, 100 ms IOI interval in the train), and at “expected tone” (100 ms after train offset; see Fig. 1 b.). Responses were also measured for “slow control” tones, tones between train onset and the rate change but separated from both by at least 2 elements; and “fast control” tones, tones between rate change and train offset but separated from both by at least 2 elements. (e.g. in a 20 element train where the rate change occurs at the 11th element, elements 4–8 are slow controls whereas elements 14–17 are fast controls). Responses to train onset were compared to slow control responses whereas rate change and train offset responses were compared to fast control responses. Average response amplitudes were measured from separate time windows for each type of the three events. The latencies of the two highest-amplitude (early and late) difference peaks were determined from the mean group response averaged over all six electrode locations. The corresponding window was defined as the continuous segment of data points on both sides of the respective peak within which the difference amplitude exceeded 30% of the corresponding peak amplitude (see Table 1. and Fig. 2. for the latency ranges). This method allows comparison between responses to event types and their respective controls when the latencies of peaks vary or no clear peaks are visible.

Effects were tested with separate dependent-measures analyses of variance (ANOVA) of the structure Stimulus type [Event vs. Control] \times Frontality [F vs. C electrode line] \times Laterality [left vs. midline vs. right] for the three timing events (Train Onset, Presentation Rate Change, Expected tone) and the two (early and late) measurement windows. Greenhouse–Geisser correction ϵ factors (where appropriate) and the partial η^2 effect sizes are given in Table 1.

3. Results

ERP responses for train onsets, presentation rate changes, and expected tone (train offsets) are shown together with the corresponding

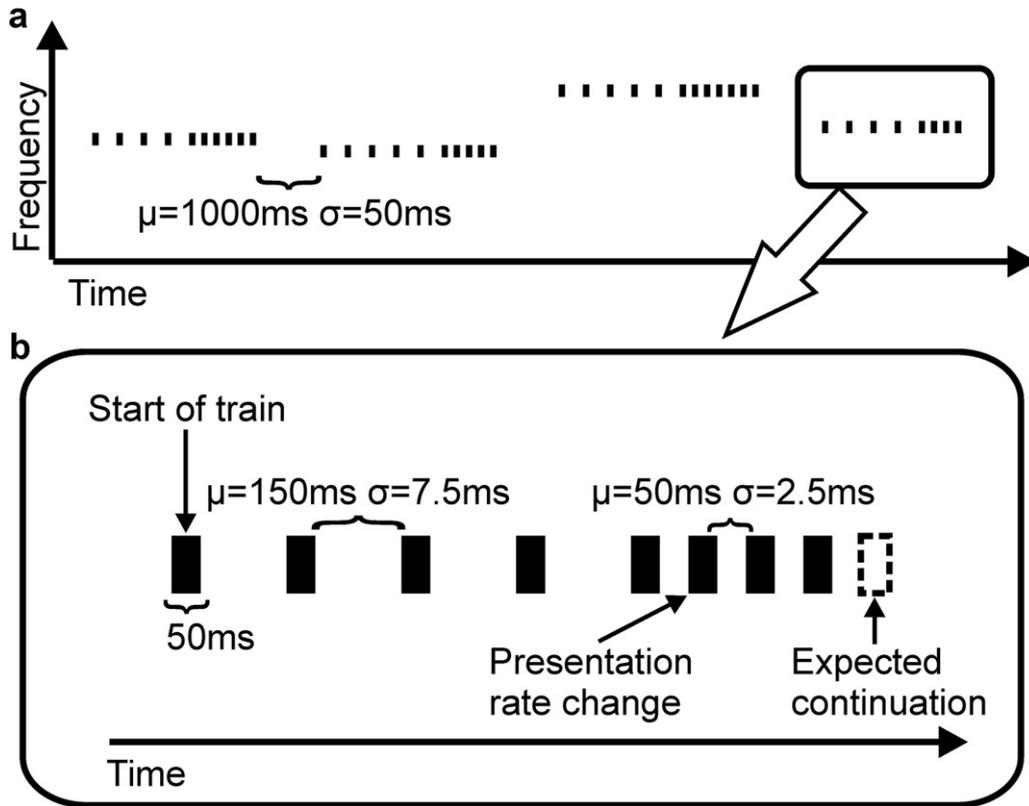


Fig. 1. Panel a) gives a schematic overview of the experimental paradigm. Panel b) shows the temporal relations within a train. The events relevant to the analyses are highlighted. Tones are denoted by black rectangles; the expected continuation of the tone sequence is marked with a dashed rectangle; time flows from left to right. Note that the inter-onset intervals were jittered in the experiment (see [Methods](#)).

control responses and difference waveforms on [Fig. 2a, b, and c](#), respectively. Each of these events elicited significantly different ERP responses in both time windows (except for the late window of the presentation rate change) compared with the corresponding control event (see [Table 1](#)).

The control responses show a much adapted response, hardly displaying any discernible waveforms. In contrast, all three stimulus train events elicited unique waveforms with detectable component structure. Thus the significant interactions with the scalp-distribution factors (frontality and laterality) mainly represent the scalp distribution of the ERPs elicited by the stimulus train events. Train onsets elicited a large early negative followed by a positive response with frontocentral maxima. This pattern is typical for large spectral energy changes

([Kushnerenko et al., 2007](#)), as is the case for sounds appearing after a relatively long silent interval having a different pitch from that appearing in the previous train. The laterality effect found in the late latency time window was caused by the left deviant response being slightly smaller than the central deviant response (Tukey HSD, $df = 58$, $p < 0.01$). Presentation-rate changes elicited only an early frontocentral negative response. Finally, the response synchronized to the expected onset of the tone that would have continued the train in an isochronous manner shows an early broadly distributed positive waveform followed by a negative one. The three-way interaction found in the late time window was caused by the left central standard response being smaller than the left and right central deviant response (Tukey HSD, $df = 58$, $p < .05$).

Table 1

Significant effects obtained in the ANOVAs of the structure Stimulus type [Event vs. Control] \times Frontality [F vs. C electrode line] \times Laterality [left vs. midline vs. right], separately for the three temporal events (Train Onset, Presentation Rate Change, Train Offset) and the two (early and late) measurement windows (see [Methods](#)). In addition to the F, df, and p values, effect sizes (η^2), and, where applicable, Greenhouse–Geisser correction factors (ϵ) are also shown. The mean and standard deviations (in parentheses) of the number of artefact free epochs are given in the second column.

ANOVA results								
Condition	μ epochs (σ)	Measurement window	Effect	F	df	p	ϵ	η^2
Train Onset vs. Slow Control	150 (19)	24–192 ms	Stimulus type	6.48	1, 29	0.016	–	0.18
			Frontality	5.35	1, 29	0.028	–	0.16
			Stimulus type	13.24	1, 29	0.001	–	0.31
			Frontality	14.03	1, 29	0.001	–	0.33
			Stimulus type \times frontality	6.12	1, 29	0.019	–	0.17
			Stimulus type \times laterality	4.11	2, 58	0.028	0.83	0.12
Presentation Rate Change vs. Fast Control	146 (17)	56–120 ms	Stimulus type	4.97	1, 29	0.034	–	0.15
			Frontality	9.60	1, 29	0.004	–	0.25
			Frontality \times laterality	6.00	2, 58	0.005	0.95	0.17
Train Offset vs. Fast Control	150 (17)	0–164 ms	Stimulus type	4.46	1, 29	0.037	–	0.14
			Stimulus type \times frontality \times laterality	4.43	2, 58	0.021	0.88	0.13

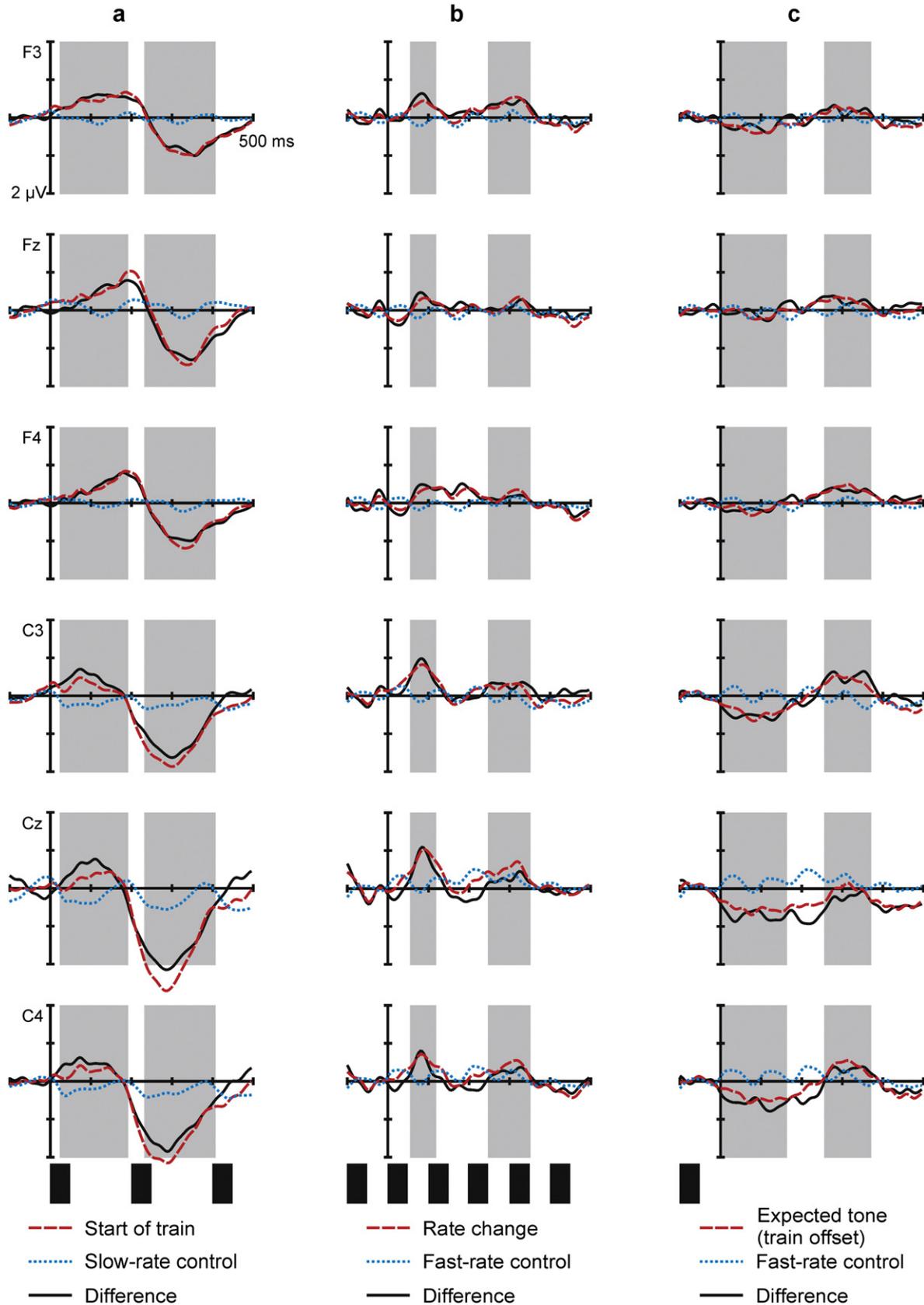


Fig. 2. Group average ($n = 30$) ERP and difference waveforms on all six electrodes (F3, Fz, F4, C3, Cz, C4) for the a) Start of train vs. Slow control; b) Presentation rate change vs. Fast control; c) Expected tone vs. Fast control events. Stimulus onset is at the crossing of the axes. Note that in c), the crossing of the axes is at the onset of the stimulus that would seamlessly continue the train. Stimuli are marked in black rectangles under each column. The stimulus mark in c) does not correspond to fast control. Amplitude measurement windows are marked with grey shading.

4. Discussion

Results showed that the neonate brain detects the onset and offset of sound trains as well as changes in the presentation rate. Thus the abilities for detecting and processing these primary temporal events are functional already at birth.

Train onsets elicited the largest responses, which were similar to the N1–P2-like components found in newborns (Wunderlich et al., 2006; Kushnerenko et al., 2002, 2007). The P2-like positive waveform has been shown to reflect the detection of sound onsets (Telkemeyer et al., 2009). This is similar to adults, in whom the auditory N1 is the most prominent response elicited by abrupt changes in spectral sound energy (Näätänen and Picton, 1987) and it likely reflects an adaptation/refractoriness based response increment. That is, in contrast to the highly refracted response elicited within the fast-paced train, the neuronal elements are relatively fresh after a longer silent interval and change of spectral contents. Such responses help one to detect the emergence of a new auditory object.

Responses elicited by train offsets show that the infants were set for the regular continuation of sound sequences. The very early onset of the difference response militates against a passive effect (i.e., that the observed waveform would represent the late response to the previous stimulus, which was cut off within the train by the arrival of next tone), because even the earliest effects of the next sound take a little more time to reach the cortex. Rather, this response is quite similar to that observed when a predictable (as opposed to an unpredictable) tone is omitted from a sequence (Bendixen et al., 2009), suggesting that this response represents the brain's preparation for a predicted sound event. This prediction error could allow one to detect the end of stimulus trains. The current response was morphologically different from that obtained in neonates for sound-omissions violating the rhythmic structure of the sound sequence (Winkler et al., 2009a,b). Predictability was lower in Winkler et al. (2009b) study due to the variability set up to distinguish the detection of a repeating pattern (Stefanics et al., 2007) from that of the rhythmic structure. The difference in the morphology of onset and offset responses is unlike the corresponding adult responses where onsets and offsets elicit similar N1 responses (Yamashiro et al., 2009). This suggests that the continuation of the train was indeed expected and the response cannot be described in terms of an offset response.

Finally, the presentation rate change elicited an early negative response. As the change of rate brings no spectral change and nor can it be explained by lower levels of refractoriness (because the inter-tone interval was shortened), this response is not likely to originate from differential refractoriness. On the other hand, the observed response is quite different from the later positive response observed for occasional early sound delivery in 2-month olds (Otte et al., 2013). This difference suggests that the processing of temporal changes is context-dependent, as was also found for spectral changes (Háden et al., 2013). One possibility is that after several trials of the same structure, the neonatal brain learned that a switch to a faster presentation rate can be expected and the response marks the detection of the onset of the change. This is supported by the similarity of the responses to that obtained in 3 month olds to the onsets of sequences made up from of either 12 or 25 ms long snippets of modulated noise (Telkemeyer et al., 2011). However, newborns did not show this type of response to the same stimuli (Telkemeyer et al., 2009). Thus the analogy may not be perfect.

We investigated the developmental origins of processing the gross temporal structure of short sound sequences. In general, we found that newborn infants have similar capabilities as adults for processing the cues that allow one to form a rough description of auditory objects. Although we have suggested in Introduction that such fundamental capabilities are required for infants for learning from others, the finding is still surprising on one sense: Research in young infants has consistently shown that when it comes to simple discrimination abilities, infantile capabilities are far from the adult level (see, e.g., pitch discrimination;

Novitski et al., 2007; for a review, see Werner, 2007). Regarding temporal features, for example, the sensitivity of detecting changes in sound duration (Kushnerenko et al., 2001; Čeponienė et al., 2002; Cheour et al., 2002) or gaps between sounds is much lower than that in adults even at 6–7 months of age (Smith et al., 2006; Trainor et al., 2001, 2003; Werner et al., 1992). In sharp contrast, the abilities required for structuring the auditory environment, such as auditory stream segregation (Winkler et al., 2003), source identification (Vestergaard et al., 2009), pattern detection (Stefanics et al., 2007), or extracting the temporal structure of sound sequences (the current study as well as Winkler et al., 2009b) appear to be functional already at birth. These abilities found already at birth could allow them to access information encoded in the tempo of both speech and music and to enter into a dialogue with others later in development where timing is crucial to achieve synchrony and facilitates even preverbal communication (Jaffe et al., 2001). The neonatal auditory processing capabilities found in the current study are amongst those serving cognitive development through helping to learn speech and music perception and bootstrapping communication by sound.

Acknowledgments

GPH is supported by a post-doctoral research fellowship of the Hungarian Academy of Sciences; HH is supported by the Research Priority Area “Brain & Cognition” of the University of Amsterdam and the Hendrik Muller chair designated on behalf of the Royal Netherlands Academy of Arts and Sciences (KNAW); IW is supported by the National Research Fund of Hungary (OTKA K101060). The authors thank research nurse Judit Roschéné Farkas for collecting the data.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.ijpsycho.2015.02.024>.

References

- Alho, K., Sainio, K., Sajaniemi, N., Reinikainen, K., Näätänen, R., 1990. Event-related brain potential of human newborns to pitch change of an acoustic stimulus. *Electroencephalogr. Clin. Neurophysiol.* 77, 151–155.
- Baruch, C., Drake, C., 1997. Tempo discrimination in infants. *Infant Behav. Dev.* 20, 573–577.
- Benasich, A.A., Tallal, P., 2002. Infant discrimination of rapid auditory cues predicts later language impairment. *Behav. Brain Res.* 136, 31–49.
- Bendixen, A., Schröger, E., Winkler, I., 2009. I heard that coming: event-related potential evidence for stimulus-driven prediction in the auditory system. *J. Neurosci.* 29, 8447–8451.
- Bobin-Bègue, A., Provasi, J., Marks, A., Pouthas, V., 2006. Influence of auditory tempo on the endogenous rhythm of non-nutritive sucking. *Eur. Rev. Appl. Psychol.* 56, 239–245.
- Brannon, E.M., Wolfe, L., Meck, W.H., Woldorff, M., 2004. Timing in the baby brain. *Brain Res. Cogn. Brain Res.* 21, 227–233.
- Brannon, E.M., Libertus, M.E., Meck, W.H., Woldorff, M.G., 2008. Electrophysiological measures of time processing in infant and adult brains: Weber's law holds. *J. Cogn. Neurosci.* 22, 193–203.
- Čeponienė, R., Kushnerenko, E., Fellman, V., Renlund, M., Suominen, K., Näätänen, R., 2002. Event-related potential features indexing central auditory discrimination by newborns. *Brain Res. Cogn. Brain Res.* 13, 101–113.
- Cheour, M., Kushnerenko, E., Čeponienė, R., Fellman, V., Näätänen, R., 2002. Electric brain responses obtained from newborn infants to changes in duration in complex harmonic tones. *Dev. Neuropsychol.* 22, 471–479.
- Chonchaiya, W., Tardif, T., Mai, X., Xu, L., Li, M., Kaciroti, N., Lozoff, B., 2013. Developmental trends in auditory processing can provide early predictions of language acquisition in young infants. *Dev. Sci.* 16, 159–172.
- Cooper, R., Aslin, R., 1990. Preference for infant directed speech in the first month after birth. *Child Dev.* 61, 1584–1595.
- Flom, R., Gentile, D.A., Pick, A.D., 2008. Infants' discrimination of happy and sad music. *Infant Behav. Dev.* 31, 716–728.
- Garrido, M.L., Kilner, J.M., Stephan, K.E., Friston, K.J., 2009. The mismatch negativity: a review of underlying mechanisms. *Clin. Neurophysiol.* 120, 453–463.
- Griffiths, T.D., Warren, J.D., 2004. Opinion: what is an auditory object? *Nat. Rev. Neurosci.* 5, 887–892.
- Grondin, S., Bisson, N., Gagnon, C., 2011. Sensitivity to time interval changes in speech and tone conditions. *Atten. Percept. Psychophys.* 73, 720–728.

- Háden, G.P., Németh, R., Török, M., Drávucz, S., Winkler, I., 2013. Context effects on processing widely deviant sounds in newborn infants. *Front. Psychol.* 4, 674.
- Honing, H., 2013. Structure and interpretation of rhythm in music. In: Deutsch, D. (Ed.), *The Psychology of Music*. Academic Press, London, pp. 369–404.
- Jaffe, J., Beebe, B., Feldstein, S., Crown, C.L., Jasnow, M.D., 2001. Rhythms of dialogue in infancy: coordinated timing in development. *Monogr. Soc. Res. Child Dev.* 66, 1–143.
- Jungers, M.K., Palmer, C., Speer, S.R., 2002. Time after time: the coordinating influence of tempo in music and speech. *Cogn. Process* 1 (2), 21–35.
- Kushnerenko, E., Čeponienė, R., Fellman, V., Huotilainen, M., Winkler, I., 2001. Event-related potential correlates of sound duration: similar pattern from birth to adulthood. *NeuroReport* 12, 3777–3781.
- Kushnerenko, E., Čeponienė, R., Balan, P., Fellman, V., Huotilainen, M., Nääätänen, R., 2002. Maturation of the auditory event-related potentials during the first year of life. *NeuroReport* 13, 47–51.
- Kushnerenko, E., Winkler, I., Horváth, J., Nääätänen, R., Pavlov, I., Fellman, V., Huotilainen, M., 2007. Processing acoustic change and novelty in newborn infants. *Eur. J. Neurosci.* 26, 265–274.
- Kushnerenko, E.V., Van den Bergh, B.R.H., Winkler, I., 2013. Separating acoustic deviance from novelty during the first year of life: a review of event-related potential evidence. *Front. Psychol.* 4, 595.
- Masataka, N., 1999. Preference for infant-directed singing in 2-day-old hearing infants of deaf parents. *Dev. Psychol.* 35, 1001–1005.
- May, P.J.C., Tiitinen, H., 2010. Mismatch negativity (MMN), the deviance-elicited auditory deflection, explained. *Psychophysiology* 47, 66–122.
- Nääätänen, R., 1990. The role of attention in auditory information processing as revealed by event related potentials and other brain measures of cognitive function. *Behav. Brain Sci.* 13, 201–288.
- Nääätänen, R., Picton, T.W., 1987. The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology* 24, 375–425.
- Nääätänen, R., Kujala, T., Winkler, I., 2011. Auditory processing that leads to conscious perception: a unique window to central auditory processing opened by the mismatch negativity (MMN) and related responses. *Psychophysiology* 48, 4–22.
- Nazzi, T., Ramus, F., 2003. Perception and acquisition of linguistic rhythm by infants. *Speech Comm.* 41, 233–243.
- Nazzi, T., Bertocini, J., Mehler, J., 1998. Language discrimination by newborns: towards an understanding of the role of rhythm. *J. Exp. Psychol. Hum. Percept. Perform.* 24, 756–766.
- Novitski, N., Huotilainen, M., Tervaniemi, M., Nääätänen, R., Fellman, V., 2007. *Neurophysiology* 118, 412–419.
- O'Connell, D.C., Kowal, S., 2008. *Communicating with One Another*. Springer, New York.
- Otte, R.A., Winkler, I., Braeken, M.A.K.A., Stekelenburg, J.J., van der Stelt, O., Van den Bergh, B.R.H., 2013. Detecting violations of temporal regularities in waking and sleeping two-month-old infants. *Biol. Psychol.* 92, 315–322.
- Patel, A.D., 2008. *Music, Language, and the Brain*. Oxford UP, Oxford.
- Protopapas, A., 2014. From temporal processing to developmental language disorders: mind the gap. *Philos. Trans. R. Soc. B* 369, 20130090.
- Quené, H., 2007. On the just noticeable difference for tempo in speech. *J. Phon.* 35, 353–362.
- Ramus, F., Hauser, M.D., Miller, C., Morris, D., Mehler, J., 2000. Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science* 288, 349–351.
- Smith, N.A., Trainor, L.J., Shore, D.I., 2006. *J. Speech Lang. Hear. Res.* 49, 1104–1113.
- Stefanics, G., Háden, G., Huotilainen, M., Balázs, L., Sziller, I., Beke, A., Fellman, V., Winkler, I., 2007. Auditory temporal grouping in newborn infants. *Psychophysiology* 44, 697–702.
- Telkemeier, S., Rossi, S., Koch, S.P., Nierhaus, T., Steinbrink, J., Poeppel, D., Obrig, H., Wartenburger, I., 2009. Sensitivity of newborn auditory cortex to the temporal structure of sounds. *J. Neurosci.* 29, 14726–14733.
- Telkemeier, S., Rossi, S., Nierhaus, T., Steinbrink, J., Obrig, H., Wartenburger, I., 2011. Acoustic processing of temporally modulated sounds in infants: evidence from a combined near-infrared spectroscopy and EEG study. *Front. Psychol.* 2.
- Trainor, L.J., Samuel, S.S., Desjardins, R.N., Sonnadara, R.R., 2001. Measuring temporal resolution in infants using mismatch negativity. *NeuroReport* 12, 2443–2448.
- Trainor, L.J., McFadden, M., Hodgson, L., Darragh, L., Barlow, J., Matsos, L., Sonnadara, R., 2003. Changes in auditory cortex and the development of mismatch negativity between 2 and 6 months of age. *Int. J. Psychophysiol.* 51, 5–15.
- Trainor, L.J., Wu, L., Tsang, C.D., 2004. Long-term memory for music: infants remember tempo and timbre. *Dev. Sci.* 7, 289–296.
- Vestergaard, M.D., Háden, G.P., Shtyrov, Y., Patterson, R.D., Pulvermüller, F., Denham, S.L., Sziller, I., Winkler, I., 2009. Auditory size-deviant detection in adults and newborn infants. *Biol. Psychol.* 82, 169–175.
- Werner, L.A., 2007. Issues in human auditory development. *Commun. Disord.* 40, 275–283.
- 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, 260–272.
- Winkler, I., 2007. Interpreting the mismatch negativity. *J. Psychophysiol.* 21 (3–4), 147–163.
- Winkler, I., Kushnerenko, E., Horváth, J., Čeponienė, R., Fellman, V., Huotilainen, M., Nääätänen, R., Sussman, E., 2003. Newborn infants can organize the auditory world. *Proc. Natl. Acad. Sci. U. S. A.* 100 (20), 11812–11815.
- Winkler, I., Denham, S.L., Nelken, I., 2009a. Modeling the auditory scene: predictive regularity representations and perceptual objects. *Trends Cogn. Sci.* 13, 532–540.
- Winkler, I., Háden, G.P., Ladinig, O., Sziller, I., Honing, H., 2009b. Newborn infants detect the beat in music. *Proc. Natl. Acad. Sci. U. S. A.* 106, 2468–2471.
- Wunderlich, J.L., Cone-Wesson, B.K., Shepherd, R., 2006. Maturation of the cortical auditory evoked potential in infants and young children. *Hear. Res.* 212, 185–202.
- Yamashiro, K., Inui, K., Otsuru, N., Kida, T., Kakigi, R., 2009. Automatic auditory off-response in humans: an MEG study. *Eur. J. Neurosci.* 30, 125–131.