

## Auditory ERP response to successive stimuli in infancy

Ao Chen, Varghese Peter, Denis Burnham

**Background.** Auditory event-related potentials (ERPs) are useful for understanding early auditory development among infants, as it allows the collection of a relatively large amount of data in a short time. So far, studies that have investigated development in auditory ERPs in infancy have mainly used single sounds as stimuli. Yet in real life, infants must decode successive rather than single acoustic events. In the present study, we tested 4-, 8-, and 12-month-old infants' auditory ERPs to musical melodies comprising three piano notes, and examined ERPs to each individual note in the melody.

**Methods.** Infants were presented with 360 repetitions of a three-note melody while EEG was recorded from 128 channels on the scalp through a Geodesic Sensor Net. For each infant both latency and amplitude of auditory components P1 and N2 were measured from averaged ERPs for each individual note.

**Results.** Analysis was restricted to response collected at frontal central site. For all three notes there was an overall reduction in latency for both P1 and N2 over age. For P1, latency reduction was significant from 4 to 8 months, but not from 8 to 12 months. N2 latency, on the other hand, decreased significantly from 4 to 8 to 12 months. With regard to amplitude, no significant change was found for either P1 or N2. Nevertheless, the waveforms of the three age groups were qualitatively different: for the 4-month-olds, the P1-N2 deflection was attenuated for the second and the third notes; for the 8-month-olds, such attenuation was observed only for the middle note; for the 12-month-olds, the P1 and N2 peaks show relatively equivalent amplitude and peak width across all three notes.

**Conclusion.** Our findings indicate that the infant brain is able to register successive acoustic events in a stream, and ERPs become better time-locked to each composite event over age. Younger infants may have difficulties in responding to late occurring events in a stream, and the onset response to the late events may overlap with the incomplete response to preceding events.

1                                    Auditory ERP Response to Successive Stimuli in Infancy

2                                    Ao Chen<sup>1,2,3</sup>, Varghese Peter<sup>1</sup>, & Denis Burnham<sup>1</sup>

3                                    <sup>1</sup>. MARCS Institute, Western Sydney University, Australia

4                                    <sup>2</sup>. Utrecht Institute of Linguistics, Utrecht University, the Netherlands

5                                    <sup>3</sup>. Department of Psychiatry, Brain Center Rudolf Magnus, University

6                                    Medical Center Utrecht, the Netherlands

7  
8                                    Corresponding author: Ao Chen, [irischen71@hotmail.com](mailto:irischen71@hotmail.com), Trans 10, 3512JK, Utrecht,  
9                                    the Netherlands. Tel: +31-30-253-6407

21

22

23 **Abstract**

24 **Background.** Auditory event-related potentials (ERPs) are useful for understanding early  
25 auditory development among infants, as it allows the collection of a relatively large amount of  
26 data in a short time. So far, studies that have investigated development in auditory ERPs in  
27 infancy have mainly used single sounds as stimuli. Yet in real life, infants must decode  
28 successive rather than single acoustic events. In the present study, we tested 4-, 8-, and 12-  
29 month-old infants' auditory ERPs to musical melodies comprising three piano notes, and  
30 examined ERPs to each individual note in the melody.

31 **Methods.** Infants were presented with 360 repetitions of a three-note melody while EEG was  
32 recorded from 128 channels on the scalp through a Geodesic Sensor Net. For each infant both  
33 latency and amplitude of auditory components P1 and N2 were measured from averaged ERPs  
34 for each individual note.

35 **Results.** Analysis was restricted to response collected at frontal central site. For all three notes  
36 there was an overall reduction in latency for both P1 and N2 over age. For P1, latency reduction  
37 was significant from 4 to 8 months, but not from 8 to 12 months. N2 latency, on the other hand,  
38 decreased significantly from 4 to 8 to 12 months. With regard to amplitude, no significant  
39 change was found for either P1 or N2. Nevertheless, the waveforms of the three age groups were  
40 qualitatively different: for the 4-month-olds, the P1-N2 deflection was attenuated for the second

41 and the third notes; for the 8-month-olds, such attenuation was observed only for the middle note;  
42 for the 12-month-olds, the P1 and N2 peaks show relatively equivalent amplitude and peak width  
43 across all three notes.

44 **Conclusion.** Our findings indicate that the infant brain is able to register successive acoustic  
45 events in a stream, and ERPs become better time-locked to each composite event over age.  
46 Younger infants may have difficulties in responding to late occurring events in a stream, and the  
47 onset response to the late events may overlap with the incomplete response to preceding events.

48

49

50

51

52

53

54

55

56

57

58

59

60

61

62

63

64

## Introduction

65           Electrical responses in the brain to external or internal events are called event-related  
66 potentials (ERPs). While various behavioural methods are available, ERPs are especially useful  
67 for studying auditory development early in infancy. Behavioral methods measure infants'  
68 auditory perception indirectly, using for example, looking time as a proxy for auditory attention,  
69 but as active attention is required, such methods are limited by infants' attentional capacity and  
70 so typically only relatively small amounts of data can be collected from each infant. In contrast,  
71 auditory ERP data collection does not require active participation, and a fairly large amount of  
72 data can be obtained within a short time period, which makes it a useful tool for studying  
73 auditory development in the early years of life (Dehaene-Lambertz & Baillet, 1998; He, Hotson,  
74 & Trainor, 2007; Kushnerenko, Ceponiene, Balan, Fellman, & Näätänen, 2002; Trainor, 2012)

75           Auditory ERPs mature from birth to childhood. The usual peaks in adult auditory ERPs,  
76 P1 (positive deflection between 40-60 ms after stimulus onset), N1 (negative deflection between  
77 90-110 ms after stimulus onset), P2 (positive deflection between 140-170 ms after stimulus  
78 onset), N2 (negative deflection between 220-280 ms after stimulus onset), may not be visible in  
79 infant auditory ERPs (Barnet, Ohlrich, Weiss, & Shanks, 1975; Ceponiene, Cheour, & Näätänen,  
80 1998; Kushnerenko et al., 2002; Näätänen, 1992; Ponton, Eggermont, Kwong, & Don, 2000;

81 Wunderlich, Cone-Wesson, & Shepherd, 2006). When presented with stimuli at inter-stimulus  
82 intervals (ISIs) shorter than 1000 ms, the auditory ERPs of infants and young children exhibit a  
83 positive peak around 150ms (P1) after stimulus onset, a negative peak around 250ms (N2),  
84 another positive peak at around 350ms (P2), and another negative peak at around 450ms (N3)  
85 (Ceponiene et al., 1998; Kushnerenko et al., 2002). Many studies have found decreases in  
86 response latency as infants mature (Barnet, Ohlrich, Weiss, & Shanks, 1975; Ceponiene, Cheour,  
87 & Näätänen, 1998; Kushnerenko et al., 2002; Näätänen, 1992; Ponton, Eggermont, Kwong, &  
88 Don, 2000; Wunderlich, Cone-Wesson, & Shepherd, 2006), probably due to increasing neuronal  
89 myelination over age (Eggermont & Salamy, 1988; Moore & Guan, 2001). Yet, such reductions  
90 of latency are not found for all peaks. Little et al. (1999) tested 5- to 17-week-old infants with  
91 100ms duration tones as well as clicks, and found latency decrease over age for the late peaks  
92 (N2, P2), but not for the earlier P1 (peak around 150ms after the stimuli onset). However, testing  
93 older infants longitudinally at 3, 6, 9, and 12 months on three 100ms harmonic tones,  
94 Kushnerenko et al. (2002) found that P1 latency decreased as the infant grew older, whereas N2  
95 latency did not show significant difference across ages. The different patterns in these two  
96 studies suggest a non-linear development of peak latency, i.e., the latency decrease for certain  
97 peaks might be more evident within a specific age window. In addition, Little et al. used only  
98 one tone (400Hz) whereas Kushnerenko et al. (2002) used three tones (500Hz, 625Hz, and  
99 750Hz) and the ERPs were obtained by averaging the responses to all the tones. These different  
100 stimuli may have influenced developmental patterns.

101           The development of the amplitude of these auditory ERP peaks shows an even more  
102 complex picture. Barnet et al. (1975) tested infants from 10 days to 37 months, using clicks as  
103 stimuli, and found that P1-N2 deflection was the “landmark” to measure auditory ERP, as it is  
104 present in all participants, and the P1-N2 amplitude increased over age. In the Little et al. (1999)  
105 study stimulus specific developments of P1 and N2 amplitudes was found: P1 showed a  
106 significant increase for clicks but not tones; N2 showed a quadratic trend for both the tones and  
107 clicks, but in the opposite direction. For the tones, N2 amplitude first increased and then  
108 decreased from 5 to 17 weeks, whereas the opposite was true for the N2 of the clicks.  
109 Kushnerenko et al. (2002) demonstrated that P1 amplitude at birth was significantly smaller than  
110 at older ages, and remained stable from 3 months to 12 months, and that N2 increased in  
111 amplitude (became more negative) between 3 and 9 months. Jing and Benasich (2006) found that  
112 N2 and P3 increased between 3 and 9 months, and then decreased with age. Together these  
113 findings suggest that the development of auditory ERPs is highly stimulus specific. Similar to  
114 latency, amplitude change might be more evident within a specific age window for certain peaks.  
115 Nevertheless, the P1-N2 deflection seems to be a reliable marker for the auditory onset response.  
116           In these studies, quite often single sounds (e.g., clicks or single tones) have been used as  
117 stimuli. Yet in real life, we often need to decode successive rather than single acoustic events.  
118 For example, speech consists of multi-word utterances (e.g., van de Weijer, 1998), and the  
119 segmentation of words from the continuous speech stream is fundamental to infants’ language  
120 learning. Similarly, music is composed of multi-note bars and phrases, often without pauses

121 between notes (Krumhansl & Jusczyk, 1990). Efficient on-line processing of real life auditory  
122 input requires efficient accurate processing of these rapidly presented successive signals.  
123 Impairments in such successive processing are associated with reading and language  
124 impairments (Tallal, 2004) and are predictive of later language skills when assessed in infancy  
125 (Choudhury & Benasich, 2011). Therefore understanding how the infant's brain responds to  
126 successive stimuli is fundamental to studying high level processing. Mismatched negativity  
127 (MMN), the neural detection of regularity violation (see Näätänen, Paavilainen, Rinne, & Alho,  
128 2007; Paavilainen, 2013), is a method that has been widely used to understand human auditory  
129 discrimination. An MMN is elicited when listeners encounter infrequent sounds embedded in  
130 repeating frequent sounds. Multi-tone melodies have been used as stimuli to understand infants'  
131 mismatched responses (Stefanics et al., 2009; Tew, Fujioka, He, & Trainor, 2009), yet these  
132 studies often incorporate relatively long inter-tone intervals between the notes composing the  
133 melodies. For example, the tones used by Stefanics et al. (2009) were of extremely short duration  
134 (50 ms), with a relatively long inter-tone intervals (150ms) and inter-pair intervals much longer  
135 than the tones themselves (1250 ms).

136 In real life, however, sounds tend to have a much longer duration and occur successively  
137 with much shorter inter-sound intervals, if at all. Speech syllables in running speech tend to have  
138 a duration of around 150-200 ms (Koopmans-van Beinum & van Donzel, 1996), while musical  
139 tempi tend to cluster between 90 and 120 beats per minute, i.e., each beat is longer than 500ms,  
140 and often contains multiple notes (Moelants, 2002). Previous findings have shown that infants

141 and adults respond to fast versus slow temporal modulations differently (Telkemeyer et al., 2009;  
142 Giraud & Poeppel, 2012), hence it is important to investigate how infants' brain responds to  
143 sound streams that represent real life characteristics, as this provides the basis for understanding  
144 higher level neural function. Given that ERPs develop in a stimulus-specific manner (Little et al.,  
145 1999; Wunderlich, Cone-Wessen, & Shepard, 2006), it is important to gain insight into how the  
146 infant brain registers real life successive acoustic events, so that higher level processing such as  
147 MMN can be better understood. In the current study, we used a melody comprised of three  
148 successive piano notes, and examined the development of auditory ERPs, specifically the  
149 infantile auditory onset responses, the P1 and N2. We tested 4-, 8-, and 12-month-old infants,  
150 and investigated how P1 and N2 latency and amplitude change across the three tones, and how  
151 P1 and N2 to successive tones develop in the first year of life. Auditory ERPs reflect neural  
152 firing time-locked to the stimuli, and exhibit high temporal resolution. As the P1-N2 onset  
153 complex is generated with a longer latency among young infants, we suggest that that upon  
154 encountering successive acoustic events, young infants may not be able to process each event  
155 sufficiently before the next one is initiated. Older infants, on the other hand, may be able to  
156 register the sounds much more rapidly, and hence may be more capable of processing each  
157 successive tone in the 3-tone stimulus.

158 **Methods**

159 **Ethics**

160 The ethics committee for human research at Western Sydney University approved all the  
161 experimental methods used in the study (Approval number: H9660). Informed consent was  
162 obtained from parents of all the participants.

### 163 **Participants**

164 Eighteen 4-month-old (range: 4 month 4 days-5 month 5 days, 10 girls), 18 8-month-old  
165 (range: 7 month 22 days-9 month 4 days, 10 girls), and 18 12-month-old (range: 11 month 10  
166 days-13 month 3 days, 11 girls) healthy infants participated in the study. All the infants were  
167 raised in a monolingual or predominantly Australian English environment. The current study is  
168 part of larger project, where we compare the perceptual development of musical versus speech  
169 stimuli. We selected these three age groups because the infants are assumed to tune into the  
170 sound structure of their native language between 4 and 12 months. None of the parents reported  
171 any hearing impairment or any ear infection within the two weeks before the experiment. One  
172 12-month-old girl was tested but excluded from the analysis as she pulled the EEG cap off  
173 shortly after the experiment started, and one 8-month-old boy was excluded due to fussiness.

### 174 **Stimuli**

175 Piano tones F3 (174.61Hz), G3 (196Hz) and A#3 (223.08Hz) were synthesized in  
176 Nyquist, where the frequency of middle A (A4) was the usual 440Hz. The notes were generated  
177 with the default duration of one-16<sup>th</sup> of a note (250ms). Then they were concatenated to form a  
178 rising melody. In order to ensure continuity and naturalness, the whole melody was then adjusted  
179 in Praat (Boersma & Weenink, 2013) using the overlap-add method to a duration of 667.5ms,

180 which resulted in slightly different durations for each note – first note duration = 220ms, second  
181 note duration = 227ms, and third note duration = 221ms. The stimuli were presented through  
182 two audio speakers (Edirol MA-10A) each 1m from the infants, and each 30 cm from the midline  
183 of the infant’s sagittal plane. The stimuli were presented at 75dB SPL, with a random inter-  
184 stimulus interval varying between 450 and 550ms. The stimuli were played through Presentation  
185 14.9 (Neurobehavioral Systems).

## 186 **Procedure**

187 The infants sat on the caregiver’s lap in a sound attenuated room. An infant friendly  
188 movie was played silently on a screen ~1m from the infants (in between the two audio speakers)  
189 to keep them engaged. Parents were instructed not to talk during the experiment. A maximum of  
190 360 trials were presented, but the experiment terminated if the infant became restless. The total  
191 duration of the recording was around 7 minutes. EEG was recorded from 128 channels on the  
192 scalp through a Geodesic Sensor Net (net sizes: 40-42cm, 43-44cm, or 44-47cm, depending on  
193 the head size of the infants), and all electrodes had impedance lower than 50 k $\Omega$  at the beginning  
194 of the experiment. The EEG was recorded at a sampling rate of 1000Hz, and the online recording  
195 was referred to the vertex.

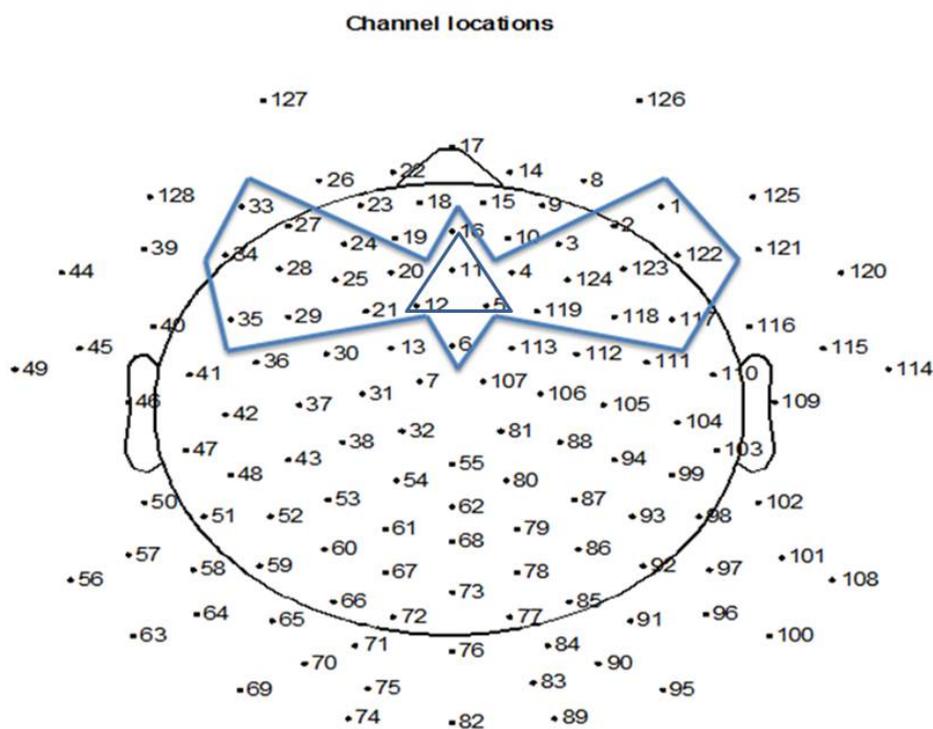
## 196 **EEG analysis**

197 The EEG was analyzed offline using NetStation 4.5.7 software and EEGLAB toolbox  
198 (version 13.1.1b) in Matlab 2011b. The raw recordings were filtered with EGI software between  
199 0.3-20hz. The filtered recordings were down-sampled to 250Hz before further analysis. The

200 continuous recordings were segmented into 1000ms epochs from the 100ms before the onset  
201 (baseline) to 900ms after the onset of the first note. The 15 electrodes at the peripheral positions  
202 of the net were removed from the analysis due to contamination from muscle movements. Next,  
203 for each participant, bad channels were identified by visual inspection and were interpolated  
204 (mean number of interpolated channels = 4, SD = 2). EEG was then re-referenced to an average  
205 reference. As it was impossible for the infants to sit still during the entire experiment, ERPs  
206 recorded from the parietal and occipital electrodes were contaminated by head movements.  
207 Auditory ERPs are mainly central-frontally distributed, hence conducting artifact reduction on all  
208 channels would waste clean signal from the central-frontal site. In order to sufficiently remove  
209 artifacts while retaining sufficient data from each child, we conducted artifact rejection on the 25  
210 frontal channels (Figure 1), where trials having an amplitude larger than 150 microvolts were  
211 removed. The remaining artifact free trials were averaged to obtain the ERP values for each  
212 infant. After artifact removal, the 4-month-olds had a mean of 248 trials (standard deviation (SD)  
213 = 39.87) accepted, the 8-month-olds a mean of 228 trials (SD = 63.64) accepted, and the 12-  
214 month-olds a mean of 275 trials (SD = 38.1) accepted. Channel 5, 11 and 12 were averaged to  
215 represent frontal central scalp ERP responses (FC), which correspond to the location of Fz on a  
216 10-20 system. Figure 1 indicates the 25 channels used for artifact reduction and the channels  
217 averaged to represent FC. For each age, a grand average ERP was computed by averaging the  
218 ERPs of all participants.

219 The P1s for each of the three notes were detected in the windows 0-200ms, 200-400ms,  
220 and 500-700ms respectively from the onset of the melody. The P1 was defined as the highest  
221 positive peak in the above time windows. The N2s to each note were detected in the windows  
222 200-300ms, 400-600ms, and 600-700ms respectively from the onset of the melody. N2 was the  
223 lowest negative peak in the above time windows. Next, for each individual participant, the P1  
224 and N2 peaks for each note were identified in the  $\pm 25$ ms window of the corresponding grand  
225 average peaks.

226

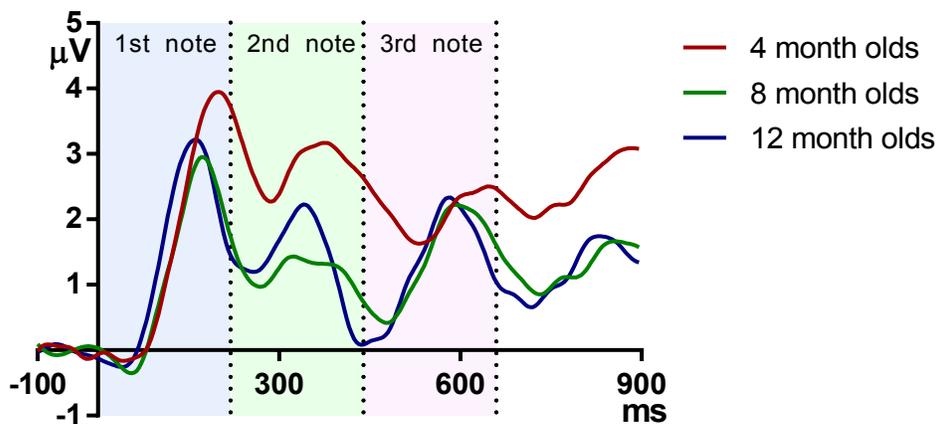


227

228 Figure 1. The 25 channels used for artifact reduction (the area circumscribed by the bow-tie  
 229 shape) and the three channels averaged for representing response at FC (circumscribed by the  
 230 triangle shape).

## 231 Results

232 Figure 2 plots the grand average ERPs for the three age groups across time. As can be seen, the  
 233 onset responses for each note, namely the infantile P1 and N2, were easily identifiable at each  
 234 age.



235  
 236 Figure 2. Grand average responses of the three age groups. The vertical axis indicates the onset  
 237 of the melody, and the vertical dotted lines indicate the onsets of the second and the third note.

238 Table 1 shows the latency of P1 and N2 for each note at FC. For each note separately, the  
 239 latency measurements of P1 and N2 at FC were submitted to a univariate ANOVA, with age as  
 240 the between-subjects variable. When age showed a significant result, Bonferroni-corrected post-  
 241 hoc tests were conducted for pair-wise comparisons. Table 2 shows the ANOVA results for the  
 242 latency measurements.

| ERP peaks | 1 <sup>st</sup> note | 2 <sup>nd</sup> note | 3 <sup>rd</sup> note |
|-----------|----------------------|----------------------|----------------------|
| 4m P1     | 195.78 (16.49)       | 375.55 (20.86)       | 648.89 (18.59)       |
| 4m N2     | 288.44(17.35)        | 532.89 (19.43)       | 722.89 (19.30)       |
| 8m P1     | 169.41 (13.34)       | 325.88 (20.84)       | 587.76 (17.00)       |
| 8m N2     | 268.94 (20.76)       | 470.12 (18.72)       | 732.00 (17.83)       |
| 12m P1    | 162.12 (16.68)       | 339.53 (16.36)       | 581.65 (12.50)       |
| 12m N2    | 250.12 (19.60)       | 439.76 (19.60)       | 712.14 (17.41)       |

243 Table 1. Mean P1 and N2 latency of each note with regard to the onset of the melody separated

244 by age groups. Standard deviations are given in parentheses.

|                         | Effect of age                 | Post-hoc test  |
|-------------------------|-------------------------------|--|
| 1 <sup>st</sup> note P1 | $F(2, 49) = 22.72, p < 0.01$  | 4m v.s. 8m, $P < 0.01$<br>4m v.s. 12m, $p < 0.01$<br>8m v.s. 12m, n.s.       |
| 1 <sup>st</sup> note N2 | $F(2, 49) = 17.33, p < 0.01$  | 4m v.s. 8m, $p < 0.05$<br>4m v.s. 12m, $p < 0.01$<br>8m v.s. 12m, $p < 0.05$ |
| 2 <sup>nd</sup> note P1 | $F(2, 49) = 30.2, p < 0.01$   | 4m v.s. 8m, $p < 0.01$<br>4m v.s. 12m, $p < 0.01$<br>8m v.s. 12m, n.s.       |
| 2 <sup>nd</sup> note N2 | $F(2, 49) = 117.92, p < 0.01$ | 4m v.s. 8m, $p < 0.01$<br>4m v.s. 12m, $p < 0.01$<br>8m v.s. 12m, $p < 0.01$ |

|     |                         |                          |                     |
|-----|-------------------------|--------------------------|---------------------|
| 245 | 3 <sup>rd</sup> note P1 | F(2, 49) = 92.08, p<0.01 | 4m v.s. 8m, p<0.01  |
| 246 |                         |                          | 4m v.s. 12m, p<0.01 |
| 247 |                         |                          | 8m v.s. 12m, n.s.   |
| 248 | 3 <sup>rd</sup> note N2 | F(2, 49) = 5.01, p<0.05  | 4m v.s. 8m, n.s.    |
| 249 |                         |                          | 4m v.s. 12m, n.s.   |
| 250 |                         |                          | 8m v.s. 12m, p<0.01 |

250

251

252

253

254

255

256

257

258

259

260

261

262 Table 2. Effects of age for P1 and N2 latency measurements for each note.

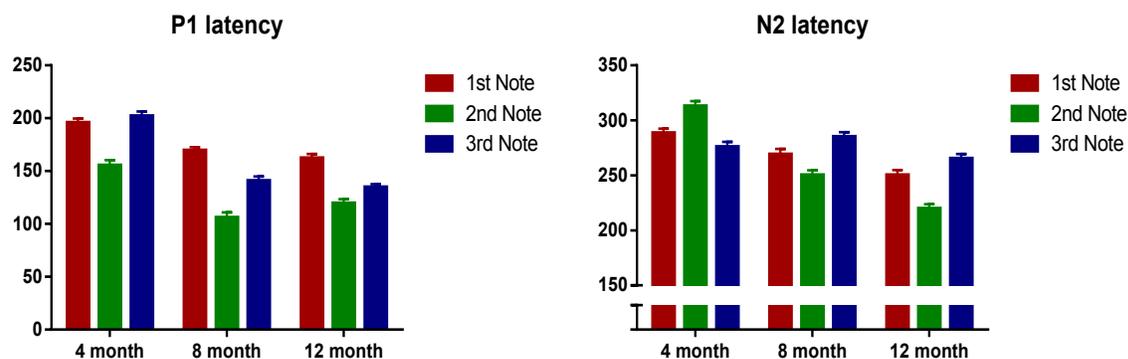
263 These results demonstrate that in general, the latency of ERP peaks is lower for older

264 infants. More specifically, reduction of P1 latency is mainly observed between 4 and 8 months,

265 whereas the latency of N2 decreases monotonically and consistently from 4 to 8 to 12 months. It  
266 can be observed that at the onsets of the second and the third note, the 12-month-olds had a more  
267 negative deflection than the 4- and 8-month-olds, which suggests that the younger groups have a  
268 later complement of N2 than the older ones.

269         Next we examined the P1 and N2 latency with regard to the onset of each note. Figure 3  
270 plots the P1 and N2 latency respectively with regard to note onset for each age group. Mixed  
271 design ANOVAs were conducted for P1 and N2 latencies with notes (Note1, Note2, Note3) as  
272 the within-subjects variable, and age as between-subjects variable. For P1, there was an overall  
273 significant main effect of note,  $F(2, 98) = 122.49, p < 0.01$ . Bonferroni-corrected post-hoc  
274 analysis indicated that P1 was significantly later for Note 1 than for Notes 2 and 3, both  $p < 0.001$ ;  
275 and P1 of Note 2 was significantly earlier than both Note 1 and 3, both  $p < 0.001$ . Age also  
276 showed a significant main effect,  $F(2, 49) = 100.48, p < 0.01$ . The 4-month-olds had a  
277 significantly later P1 compared to the 8- and 12-month-olds ( $p < 0.001$ ) with no significant  
278 difference between the 8- and 12-month-olds. A significant interaction was found between notes  
279 and age,  $F(2, 98) = 8.96, p < 0.01$ . Bonferroni post-hoc tests indicated that: for the 4-month-olds,  
280 P1 of Note 2 was significantly earlier than that of Notes 1 and 3 (both  $p < 0.01$ ) with no  
281 significant difference between Notes 1 and 3; for the 8- and 12-month-olds, P1 of Note 2 was  
282 significantly earlier than that of Notes 1 and 3 (both  $p < 0.01$ ), and P1 for Note 1 was significantly  
283 later than that of Note 3 (both  $p < 0.05$ ).

284 For N2, notes had a significant main effect, where  $F(2, 98) = 10.15, p < 0.01$ , and  
 285 Bonferroni-corrected post-hoc tests indicated that the N2 of Note 2 was significantly earlier than  
 286 that of Note 1 and 3, with no significant difference between these two. Age was also significant,  
 287  $F(2, 49) = 58.83, p < 0.01$ . Bonferroni-corrected post-hoc tests indicated that the N2 latency  
 288 decreased significantly from 4 to 8 months, and from 8 to 12 months, both  $p < 0.05$ . A significant  
 289 interaction was found between notes and age,  $F(2, 98) = 34.00, p < 0.01$ , Bonferroni corrected  
 290 post-hoc tests indicated that: for 4-month-olds, N2 latency of Note 2 was significantly later than  
 291 that of Notes 1 and 3 (both  $p < 0.01$ ); whereas on the other hand, for the 8- and 12-month-olds, N2  
 292 latency of Note 2 was significantly earlier than that of Notes 1 and 3, with N2 latency of Note 3  
 293 being significantly later than that of Note 1 (all  $p < 0.05$ ).



294  
 295 Figure 3. Mean P1 and N2 latency with regard to each note's onset of the three age groups. Error  
 296 bars represent standard errors.

297 In summary, P1 latency for each note decreases as the infants grow older. Interestingly,  
 298 for all the three ages, the P1 of the middle note, Note 2, had the shortest latency compared to P1

299 of the edge notes. For the 4-month-olds, the P1 latency of Note 3 is longer than that of Note 1,  
300 whereas the 8- and 12-month-olds showed the opposite pattern. For the N2, the 4-month-olds had  
301 a longer N2 latency for Note 2, whereas for the other two groups, N2 latency was the shortest for  
302 Note 2.

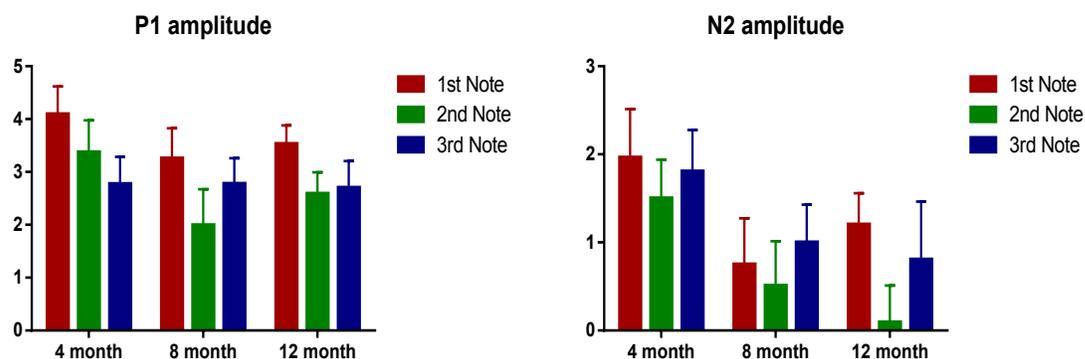
303 Mean P1 and N2 amplitudes of each note for the three age groups are provided in Table 3  
304 and Figure 4 plots the mean P1 and N2 amplitude of each note of the three age groups. ANOVAs  
305 were conducted for the absolute amplitude of P1 and N2 with notes as the within-subjects  
306 variable and age as the between-subjects variable. For P1, a significant main effect of notes was  
307 found,  $F(2, 98) = 9.01, p < 0.01$ , but no significant main effect of age. Bonferroni-corrected pair-  
308 wise analyses indicated that P1 amplitude was significantly larger at Note 1, compared to that at  
309 Notes 2 and 3 (both  $p < 0.01$ ), with no difference between these two. For N2, notes had a  
310 marginally significant effect,  $F(2, 98) = 2.95, p = 0.057$ , and no significant difference was found  
311 between any pair of the three notes, and there was no significant main effect of age. For neither  
312 P1 nor N2, was there a significant interaction between age and notes. Similarly, for P1-N2 peak-  
313 to-peak amplitude, notes again showed a significant main effect,  $F(2, 98) = 3.66, p < 0.01$ ; the P1-  
314 N2 peak-to-peak amplitude of Note 3 was significantly smaller than that of Note 1 ( $p < 0.05$ ),  
315 whereas no significant difference was found between either Note 1 and 2, or Note 2 and 3. As  
316 notes had a significant main effect for all the three amplitude measurements, we examined the  
317 effect of age for each note separately. For both the P1 and N2 amplitudes, age did not show a  
318 significant effect for any of the notes, and no pair-wise difference was significant. For P1-N2

319 amplitude, no significant effect of age was found for either Note 1 or 2, but for Note 3, there was  
 320 a significant age effect,  $F(2, 49) = 3.45$ ,  $p < 0.05$ ; Bonferroni-corrected pair-wise analyses showed  
 321 a marginally significant difference between the 4- and 12-month-olds, with the 4-month-olds  
 322 having a smaller amplitude ( $p = 0.06$ ).

323

|        | 1 <sup>st</sup> note | 2 <sup>nd</sup> note | 3 <sup>rd</sup> note |
|--------|----------------------|----------------------|----------------------|
| 4m P1  | 4.09 (2.23)          | 3.37 (2.56)          | 2.77 (2.18)          |
| N2     | 1.96 (2.34)          | 1.50 (1.86)          | 1.81 (1.99)          |
| 8m P1  | 3.26 (2.36)          | 1.99 (2.81)          | 2.78 (1.99)          |
| N2     | 1.76 (2.15)          | 0.51 (2.09)          | 1.00 (1.76)          |
| 12m P1 | 3.53 (1.45)          | 2.59 (1.67)          | 2.70 (2.07)          |
| N2     | 1.20 (1.47)          | 0.09 (1.72)          | 0.81 (2.71)          |

324 Table 3. Mean P1 and N2 amplitude (in  $\mu\text{V}$ ) of each note of each age group. Standard deviations  
 325 are given in parenthesis.



326

327 Figure 4. Mean P1 and N2 amplitude of each note of each age group. Error bars represent  
328 standard errors.

### 329 **Discussion**

330 In the present study, we tested 4-, 8- and 12-month-old infants on their auditory ERPs to  
331 a three-note melody. For all the three successive notes, there was an overall decrease in P1 and  
332 N2 latency over age. This indicates an increase in the speed of neural transmission, which is  
333 presumably related to the increased myelination of neurons over age. Latency decrease has been  
334 previously found in the response to single acoustic stimuli (e.g. Barnet et al., 1975; Novak et al.,  
335 1989; Kushnerenko et al., 2002), but as the results here show decreasing latencies to successive  
336 stimuli in a stream, this implies that infants are able to segment a sound stream into its composite  
337 elements, with individual registration of each element at the neural level. Even at 4 months,  
338 although the N2 was not fully expressed when the next note occurred, P1 for the next note was  
339 still visibly evident. Nevertheless, as the onset response to the late events in a stream may  
340 overlap with the response from preceding events, the non-initial response is possibly more  
341 variable across participants. Kushnerenko et al. (2002) found a decrease in P1 but not in N2  
342 latency from 6 to 12 months, and N2 was not visibly evident before 6 months. Little et al. (1999)  
343 found a decrease in N2 but not in P1 from birth to 17 weeks. In the present study, we found  
344 consistent latency decrease across age for both peaks. These different findings are possibly due  
345 to the different stimuli: Little et al. (1999) used one single harmonic tone, Kushnerenko et al.  
346 (2002) used three single harmonic tones, and we used a stream containing three piano tones. It is

347 possible that the relatively rich spectral content of our stimuli allowed N2 to be visibly evident at  
348 a younger age than in the Kushnerenko et al. (2002) study.

349         Interestingly, the middle Note 2 elicited an earlier P1 latency compared to the edge notes,  
350 1 and 3. The early P1 of the middle note is unlikely to be solely due to the incomplete response  
351 to the first note. If this were the case, then Note 3 should have had an equally early P1 peak. It  
352 seems that infant brain does not respond to the acoustic events at different positions of a stream  
353 in an identical manner. For the moment, it is difficult to ascertain what might cause the change in  
354 P1 latency across successive notes, yet it seems that the response to sounds at a central position  
355 in a stream tend to be more variable. In addition, for the 8-month-olds, the amplitude of the  
356 middle note seems to be more attenuated compared to the response to the first and third notes.  
357 Whether the neuronal processes are different for edge events and for medial events in a stream  
358 should be further tested, preferably with streams containing more than three notes. The majority  
359 words in English have a strong-weak pattern (Cutler & Carter, 1987), and whether such metrical  
360 structure might lead to a more prominent response to the onset event in an acoustic stream should  
361 be investigated by testing participants whose native language exhibits the opposite metrical  
362 pattern (e.g., Turkish).

363         With regard to amplitude, P1 was larger for the first than for the subsequent notes.  
364 Neither P1 nor N2 amplitude changed significantly across age, although it should be  
365 acknowledged that individual variation is large for amplitude measurements, and such variation  
366 may have masked any age group differences. Even though the statistical tests failed to show any

367 consistent age effect, the grand average waveforms are qualitatively different for each age (see  
368 Figure 2). For the 4-month-olds, clear P1s and N2s can be observed for Note 1 whereas those to  
369 the two subsequent notes are much more attenuated. For the 8-month-olds, compared to the edge  
370 notes, the central Note 2 has smaller P1 and N2 amplitudes. In addition, the P1 peak of the  
371 central note shows a plateau-like morphology, whereas the responses to the edge notes have  
372 sharper P1 peaks. The P1 and N2 peaks of the 12-month-olds show a comparably equivalent  
373 amplitude and peak width across all three notes. Thus in summary, compared to older 8- and 12-  
374 month-old infants, the younger 4-month-old infants had a smaller P1-N2 deflection for Note 3,  
375 even this was the final note, with no other acoustic event following it. This suggests that young  
376 infants may have difficulties in responding to late occurring events in a stream. Such attenuated  
377 amplitude may be due to neural refractoriness (Budd, Barry, Gordon, Rennie, & Michie, 1998;  
378 Gilley, Sharma, Dorman, & Martin, 2005; Ritter, Vaughan, & Costa, 1968; Sable, Low, Maclin,  
379 Fabiani, & Gratton, 2004; Sussman, Steinschneider, Gumenyuk, Grushko, & Lawson, 2008) with  
380 processing of late events being hindered by unfinished processing of preceding events. In other  
381 words, more time may be required for neurons to reset and respond to following acoustic events  
382 for young infants.

### 383 **Conclusion**

384 In this study we examined auditory cortical responses to successive musical notes of  
385 infant at ages of 4, 8, and 12 months. Clear P1 and N2 peaks were identified for all three notes at  
386 all three ages, indicating successful neural registration of individual components in a continuous

387 stream. While there was a significant decrease in P1 and N2 latencies all three notes as infant age  
388 increased, these results show, for the first time, individual neural registration of components of  
389 continuous multi-element events, even in relatively young infants. Over and above this general  
390 finding age-related nuances were evident: P1 and N2 amplitudes were larger for the first note in  
391 the stream than for subsequent notes, and additionally, for the final Note 3, the 4-month-olds had  
392 a smaller amplitude compared to the two older groups, suggesting that, despite individual  
393 registration, young infants may have greater difficulty registering late occurring events in a  
394 stream. This attenuated amplitude for late events at the younger age points to the possible nature  
395 of developmental progression in auditory event registration in infancy; there could well be neural  
396 registration of individual elements in an acoustic stream, but, possibly due to incomplete  
397 processing of each event, there may be cumulative reduction of processing resources within a  
398 certain refractory period. This would imply temporal and numerical limits on successive  
399 component event processing, which are possibly related to memory limits. Further studies of  
400 such limits and their reduction over age will greatly improve the understanding of speech  
401 perception and language development as well as the development of other complex abilities such  
402 as music and other event processing. For young infants, P1 and N2 latencies change over the  
403 course of the experiment. This study provides the basis for such future studies by showing, for  
404 the first time, that as young as 4 months, infants show specific neural registration of successive  
405 frequency-modulated events in an otherwise continuous stream, and continuing maturation of the  
406 brain response to successive sounds in the first year of life.

## 407 References

- 408 Barnet, A. B., Ohlrich, E. S., Weiss, I. P., & Shanks, B. (1975). Auditory evoked potentials  
409 during sleep in normal children from ten days to three years of age. *Electroencephalography  
410 and Clinical Neurophysiology*, 39(1), 29-41. doi:[10.1016/0013-4694\(75\)90124-8](https://doi.org/10.1016/0013-4694(75)90124-8)
- 411 Budd, T. W., Barry, R. J., Gordon, E., Rennie, C., & Michie, P. T. (1998). Decrement of the N1  
412 auditory event-related potential with stimulus repetition: habituation vs. refractoriness.  
413 *International Journal of Psychophysiology*, 31(1), 51-68. doi:[10.1016/S0167-  
414 8760\(98\)00040-3](https://doi.org/10.1016/S0167-8760(98)00040-3)
- 415 Ceponiene, R., Cheour, M., & Näätänen, R. (1998). Interstimulus interval and auditory event-  
416 related potentials in children: Evidence for multiple generators. *Electroencephalography  
417 and Clinical Neurophysiology - Evoked Potentials*, 108(4), 345-354. doi:[10.1016/S0168-  
418 5597\(97\)00081-6](https://doi.org/10.1016/S0168-5597(97)00081-6)
- 419 Choudhury, N., & Benasich, A. A. (2011). Maturation of auditory evoked potentials from 6 to 48  
420 months: Prediction to 3 and 4 year language and cognitive abilities. *Clinical  
421 Neurophysiology*, 122(2), 320-338. doi:[10.1016/j.clinph.2010.05.035](https://doi.org/10.1016/j.clinph.2010.05.035)
- 422 Cutler, A., & Carter, D. M. (1987). The predominance of strong initial syllables in the English  
423 vocabulary. *Computer Speech & Language*, 2(3-4), 133-142. doi:[10.1016/0885-  
424 2308\(87\)90004-0](https://doi.org/10.1016/0885-2308(87)90004-0)
- 425 Dehaene-Lambertz, G., & Baillet, S. (1998). A phonological representation in the infant brain.  
426 *NeuroReport*, 9(8), 1885-1888. doi: [10.1097/00001756-199806010-00040](https://doi.org/10.1097/00001756-199806010-00040)

- 427 Gilley, P. M., Sharma, A., Dorman, M., & Martin, K. (2005). Developmental changes in  
428 refractoriness of the cortical auditory evoked potential. *Clinical Neurophysiology*, *116*(3),  
429 648-657. doi:[10.1016/j.clinph.2004.09.009](https://doi.org/10.1016/j.clinph.2004.09.009)
- 430 Giraud, A., & Poeppel, D. (2012). Cortical oscillations and speech processing: Emerging  
431 computational principles and operations. *Nature Neuroscience*, *15*(4), 511-517. Retrieved  
432 from [10.1038/nn.3063](https://doi.org/10.1038/nn.3063)
- 433 He, C., Hotson, L., & Trainor, L. J. (2007). Mismatch responses to pitch changes in early infancy.  
434 *Journal of Cognitive Neuroscience*, *19*(5), 878-892. doi:[10.1162/jocn.2007.19.5.878](https://doi.org/10.1162/jocn.2007.19.5.878)
- 435 Jing, H., & Benasich, A. A. (2006). Brain responses to tonal changes in the first two years of life.  
436 *Brain and Development*, *28*(4), 247-256. doi:[10.1016/j.braindev.2005.09.002](https://doi.org/10.1016/j.braindev.2005.09.002)
- 437 Kushnerenko, E., Ceponiene, R., Balan, P., Fellman, V., & Näätänen, R. (2002). Maturation of  
438 the auditory change detection response in infants: a longitudinal ERP study. *Neuroreport*,  
439 *13*(15), 1843-1848. doi: [10.1097/00001756-200210280-00002](https://doi.org/10.1097/00001756-200210280-00002)
- 440 Little, V. M., Thomas, D. G., & Letterman, M. R. (1999). Single-trial analyses of developmental  
441 trends in infant auditory event-related potentials. *Developmental Neuropsychology*, *16*(3),  
442 455-478. doi:[10.1207/S15326942DN1603\\_26](https://doi.org/10.1207/S15326942DN1603_26)
- 443 Moelants, D. (2002). Preferred tempo reconsidered. *Proceedings of the 7th International*  
444 *Conference on Music Perception and Cognition*, pp. 580-583.
- 445 Näätänen, R. (1992). *Attention and brain function*. New Jersey: Erlbaum.

- 446 Nääätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in  
447 basic research of central auditory processing: A review. *Clinical Neurophysiology*, *118*(12),  
448 2544-2590. doi:[10.1016/j.clinph.2007.04.026](https://doi.org/10.1016/j.clinph.2007.04.026)
- 449 Novak, G. P., Kurtzberg, D., Kreuzer, J. A., & Vaughan Jr., H. G. (1989). Cortical responses to  
450 speech sounds and their formants in normal infants: maturational sequence and  
451 spatiotemporal analysis. *Electroencephalography and Clinical Neurophysiology*, *73*(4), 295-  
452 305. doi:[10.1016/0013-4694\(89\)90108-9](https://doi.org/10.1016/0013-4694(89)90108-9)
- 453 Paavilainen, P. (2013). The mismatch-negativity (MMN) component of the auditory event-  
454 related potential to violations of abstract regularities: A review. *International Journal of*  
455 *Psychophysiology*, *88*(2), 109-123. doi:[10.1016/j.ijpsycho.2013.03.015](https://doi.org/10.1016/j.ijpsycho.2013.03.015)
- 456 Ponton, C. W., Eggermont, J. J., Kwong, B., & Don, M. (2000). Maturation of human central  
457 auditory system activity: evidence from multi-channel evoked potentials. *Clinical*  
458 *Neurophysiology*, *111*(2), 220-236. doi:[10.1016/S1388-2457\(99\)00236-9](https://doi.org/10.1016/S1388-2457(99)00236-9)
- 459 Ritter, W., Vaughan, H. G., & Costa, L. D. (1968). Orienting and habituation to auditory stimuli:  
460 A study of short terms changes in average evoked responses. *Electroencephalography and*  
461 *Clinical Neurophysiology*, *25*(6), 550-556. doi:[10.1016/0013-4694\(68\)90234-4](https://doi.org/10.1016/0013-4694(68)90234-4)
- 462 Sable, J. J., Low, K. A., Maclin, E. L., Fabiani, M., & Gratton, G. (2004). Latent inhibition  
463 mediates N1 attenuation to repeating sounds. *Psychophysiology*, *41*(4), 636-642.  
464 doi:[10.1111/j.1469-8986.2004.00192.x](https://doi.org/10.1111/j.1469-8986.2004.00192.x)

- 465 Shucard, D. W., Shucard, J. L., & Thomas, D. G. (1987). Auditory event-related potentials in  
466 waking infants and adults: a developmental perspective. *Electroencephalography and*  
467 *Clinical Neurophysiology*, 68(4), 303-310. doi:[10.1016/0168-5597\(87\)90051-7](https://doi.org/10.1016/0168-5597(87)90051-7)
- 468 Stefanics, G., Háden, G. P., Sziller, I., Balázs, L., Beke, A., & Winkler, I. (2009). Newborn  
469 infants process pitch intervals. *Clinical Neurophysiology*, 120(2), 304-308.  
470 doi:[10.1016/j.clinph.2008.11.020](https://doi.org/10.1016/j.clinph.2008.11.020)
- 471 Sussman, E., Steinschneider, M., Gumenyuk, V., Grushko, J., & Lawson, K. (2008). The  
472 maturation of human evoked brain potentials to sounds presented at different stimulus rates.  
473 *Hearing Research*, 236(1-2), 61-79. doi:[10.1016/j.heares.2007.12.001](https://doi.org/10.1016/j.heares.2007.12.001)
- 474 Tallal, P. (2004). Improving language and literacy is a matter of time. *Nature Reviews*  
475 *Neuroscience*, 5(9), 721-728. doi: [10.1038/nrn1499](https://doi.org/10.1038/nrn1499)
- 476 Telkemeyer, S., Rossi, S., Koch, S. P., Nierhaus, T., Steinbrink, J., Poeppel, D., . . .  
477 Wartenburger, I. (2009). Sensitivity of newborn auditory cortex to the temporal structure of  
478 sounds. *The Journal of Neuroscience*, 29(47), 14726-14733.  
479 doi:[10.1523/JNEUROSCI.1246-09.2009](https://doi.org/10.1523/JNEUROSCI.1246-09.2009)
- 480 Tew, S., Fujioka, T., He, C., & Trainor, L. (2009). Neural representation of transposed melody in  
481 infants at 6 months of age. *Annals of the New York Academy of Sciences*, 1169(1), 287-290.  
482 doi:[10.1111/j.1749-6632.2009.04845.x](https://doi.org/10.1111/j.1749-6632.2009.04845.x)

- 483 Trainor, L. J. (2012). Musical experience, plasticity, and maturation: issues in measuring  
484 developmental change using EEG and MEG. *Annals of the New York Academy of Sciences*,  
485 *1252*(1), 25-36. doi:10.1111/j.1749-6632.2012.06444.x
- 486 Weitzman, E. D., & Graziani, L. J. (1968). Maturation and topography of the auditory evoked  
487 response of the prematurely born infant. *Developmental Psychobiology*, *1*(2), 79-89.  
488 doi:10.1002/dev.420010203
- 489 Wunderlich, J. L., Cone-Wesson, B. K., & Shepherd, R. (2006). Maturation of the cortical  
490 auditory evoked potential in infants and young children. *Hearing Research*, *212*(1-2), 185-  
491 202. doi:[10.1016/j.heares.2005.11.010](https://doi.org/10.1016/j.heares.2005.11.010)
- 492