Infant Directed Speech Enhances Statistical Learning in Newborn Infants: An ERP Study

Alexis N. Bosseler[1,2]*, Tuomas Teinonen[1], Mari Tervaniemi[1,3], Minna Huotilainen[1,3]

1 Cognitive Brain Research Unit, Institute of Behavioural Sciences, University of Helsinki, Helsinki, Finland, **2** Institute for Learning and Brain Sciences, University of Washington, Seattle, Washington, United States of America, **3** Cicero Learning, University of Helsinki, Helsinki, Finland

* bosseler@u.washington.edu

Abstract

Statistical learning and the social contexts of language addressed to infants are hypothesized to play important roles in early language development. Previous behavioral work has found that the exaggerated prosodic contours of infant-directed speech (IDS) facilitate statistical learning in 8-month-old infants. Here we examined the neural processes involved in on-line statistical learning and investigated whether the use of IDS facilitates statistical learning in sleeping newborns. Event-related potentials (ERPs) were recorded while newborns were exposed to 12 pseudo-words, six spoken with exaggerated pitch contours of IDS and six spoken without exaggerated pitch contours (ADS) in ten alternating blocks. We examined whether ERP amplitudes for *syllable position* within a pseudo-word (word-initial vs. word-medial vs. word-final, indicating statistical word learning) and *speech register* (ADS vs. IDS) would interact. The ADS and IDS registers elicited similar ERP patterns for syllable position in an early 0–100 ms component but elicited different ERP effects in both the polarity and topographical distribution at 200–400 ms and 450–650 ms. These results provide the first evidence that the exaggerated pitch contours of IDS result in differences in brain activity linked to on-line statistical learning in sleeping newborns.

Introduction

A long-standing question in cognitive neuroscience concerns the learning processes that guide language acquisition. Infants begin life with perceptual abilities that allow them to learn any language, and their perception is shaped by experience with their native language [1–3]. Previous research indicates that both infant-directed speech (IDS) and "statistical learning" (the ability to detect the distributional and statistical patterns of phonetic units in language input) play important roles in this process, influencing both phonetic learning and early word learning [4–8]. There is also evidence that the prosodic characteristics of IDS, as compared to the less varying prosody of adult-directed speech (ADS), may promote statistical learning by enhancing infants' attention to speech as early as 8-month of age [9]. Exaggerated pitch contours of IDS may benefit early word learning by heightening attention to the input, which in turn expedites the detection of statistical regularities (see [10] for discussion). Given that

electrophysiological studies have shown newborn infants can track statistical regularities in the ambient language [7] and are sensitive to pitch variation [11], a relevant question for developmental science is whether newborn infant use both prosodic fluctuations and the statistical regularities simultaneously to learn language, or whether these cues serve distinct functions early in development.

Studies on statistical learning show that attention [12] as well as the listener's experience with auditory input impact the rate of learning, as well as the polarity and distribution of brain responses [13–16]. In the present study, we examined the hypothesis that exaggerated pitch will affect the efficiency of computational learning in sleeping newborns and assessed the effect of both exaggerated pitch and statistical regularities on their brain responses. We specifically examined whether the pattern of brain responses to a statistical learning paradigm would vary as a function of speech register. We hypothesized that brain activity could be influenced by speech register in at least two ways.

First, differences in ERP response to ADS and IDS may be related to acoustic processing. Previous research in infants reported enhanced brain activity for IDS when compared to ADS [17–21]. A functional imaging study has shown an increase in blood flow over the frontal area of newborn brains as they listened to their mother speak in IDS, as opposed to ADS [18]. A similar increase in frontal activity to IDS has been reported using electroencephalography (EEG) power in 9-month-old infants [19]. Event-related potential studies have also revealed an enhanced response to IDS versus ADS at both the phonetic [21] and word levels [20]. We also hypothesized differences in ERP response to ADS and IDS may be related to processing efficiency. In this regard, the pattern of effects linked to statistical learning would be more broadly distributed for the ADS register as compared to the IDS register. For example, experiments with adults have shown that attention [12] as well as the listener's experience with auditory input impacts both the rate of learning and distribution of brain responses within a statistical learning paradigm [13–16]. However, it is likely that experience with stimuli and the allocation of cognitive resources, such as attention, are linked [20, 22, 23]. In this view, processing familiar or more salient auditory input frees attentional resources for other tasks involved in language processing, such as detecting the transitional probabilities between syllables (see also [24]). Recent functional imaging results in adults support this view. Tremblay et al. [13] compared segmentation accuracy and the corresponding neural responses for segmenting speech and birdsong and found that the brain activity linked to computational learning for speech input was more focal and significantly smaller in magnitude as compared to non-speech input. A similar parallel between familiarity and the extent of brain activity has been corroborated in studies examining adult processing of native versus non-native phonemes [25–29], and the processing of known versus unknown words in young children [20, 22, 30, 31].

We were also interested in identifying the similarities in the pattern of brain responses that arise from tracking the statistics. In adults, word-initial syllables elicit both an early N100 component and a later N400 component [12, 14–16, 32–34]. The N400 has traditionally been linked to semantic expectancy [35], word category violations, or unexpected but semantically acceptable words (e.g. [36]); however, within the context of a statistical learning paradigm this response is thought to relate to the identification of recently segmented pseudo-words [14]. There is evidence suggesting the N100 response reflects cognitive processes arising from the predictive dependencies of word onset (see[34]). Whether this mechanism also contributes to newborn segmentation abilities has not been examined. The two previous electrophysiological studies investigating statistical learning in newborn infants report different patterns of brain activity [37, 38]. In the first of these studies, Teinonen et al. [38] demonstrated that exposing newborn infants to tri-syllabic pseudo-words embedded within a continuous speech stream results in a larger negative deflection to word-initial syllables of each pseudo-word compared

to word-medial or word-final syllables, beginning after 300 ms. Using tri-tone pseudo-words, Kudo et al. [37] reported a broad positive deflection spanning 550 ms from stimulus onset that was significant only over frontal electrode sites. These studies differed in the type of stimuli used (speech vs. non-speech), the amount of exposure to the input the newborns received, and approach to analysis, making the relative influences of experience on brain activity linked to learning across investigations difficult to assess.

Several studies of ERP responses to native and non-native stress patterns in infants have reported a mismatch response (MMR) with a positive polarity for a non-native stress pattern [39, 40]. The positive polarity of the MMR may be dependent on the stimulus characteristics, presentation speed, or reflect an enhanced effort in processing less familiar patterns due to the involvement of weaker or less activated (immature) brain processes (see [41, 42], however, see [43, 44] for a different view). Moreover, 7-month-old infants whose brain response showed a negative deflection to a familiarized word embedded in continuous speech showed more advanced language skills at 3 years of age as compared to 7-month-old infants whose brain response showed a more distributed, positive deflection, to the same stimuli [45]. Experiments with adults have shown that attention [12] as well as the listener's experience with auditory input impacts both the rate of learning and distribution of brain responses within a statistical learning paradigm [13–16]. However, it is likely that experience with stimuli and the allocation of cognitive resources, such as attention, are linked [20, 30, 46]. In this view, processing familiar or more salient auditory input frees attentional resources for other tasks involved in language processing, such as detecting the transitional probabilities between syllables (see also [24]). The current study asked whether evidence of sensitivity to predictive dependencies is reflected in newborn ERPs. We had two hypotheses. First, based on reports that newborns track the conditional probabilities between syllables and tones [37, 38], we hypothesized that a predictive response would be present for both the ADS and IDS. Second, based on studies showing an enhanced response for predicted input (e.g. [47, 48–50]), we hypothesized that a predictive response would manifest as an enhanced response to word medial and word final syllables for both the ADS and IDS. Infants' and adults' computation of statistical probabilities coincides with experimental learning research showing that both human and non-human animals are sensitive to predictive dependencies of environmental input, and that this sensitivity guides learning ([51–53], see also [54]). Although the majority of the studies on sensitivity to predictive dependencies focus on the reduction in brain activity that occurs to predictable input [55–59], in newborns, frequently presented stimuli will elicit an enhanced negative ERP within 100 ms of onset, peaking at around 50 ms [60]. A similar negative deflection has been recently reported to familiar versus unfamiliar words in 7-month-old infants [45]. Attending to the regularities in the environment is efficient in promoting learning during infancy, because probability statistics can reveal information that assists category formation across domains (see [24]).

In summary, we hypothesized that when presented with a statistical learning paradigm, ERP amplitudes as a function of *syllable position* within a pseudo-word (word-initial vs. word-medial vs. word-final, reflecting word learning) and *speech register* (ADS vs. IDS) would interact. First, a *predictive* response would occur prior to 100 ms for syllable positions with high transitional probabilities (word-medial and word-final syllables) (for review, see [61, 62, 63]), for both the IDS and ADS registers. Second, we expected an enhanced *acoustic processing* response for the IDS, but not the ADS register in the 200–400 ms measurement window (see [33]). Third, we hypothesized that differences in brain activity in response to segmented pseudo-words would be evident in the latency window linked to successful segmentation for speech in newborn infants (i.e, after 300 ms, see 38), and sensitive to the saliency of the speech register (see [13, 16]) in terms of both the topography and the polarity of the effect (see [12,

20]). For the less salient ADS register, we hypothesized that the effect of word-initial syllables would be broadly distributed across electrode sites, whereas the effect of word-initial syllables for the highly salient IDS register would occur over a small subset of electrodes, consistent with the literature on neural efficiency. In this respect, the more diffuse ERPs would presumably reflect greater cognitive effort.

Materials and Methods

Twenty-five healthy full-term newborns were recruited at Jorvi Hospital, Espoo, Finland (11 boys, 14 girls). Of the 25 newborns, 2 were omitted from the analysis, one due to experimenter error in the recording procedure and the second due to excessive movements during the measurement. The infants were recorded 0–3 days after birth, with a mean gestational age of 40 weeks and 2.64 days (39 weeks 1 day– 42 weeks 3 days), a mean birth weight of 3,552 kg (2875–4375 kg) and a mean Apgar score of 9.47 (6–10). The study protocol was approved by the Ethics Committee for Pediatrics, Adolescent Medicine, and Psychiatry, Hospital District of Helsinki and Uusimaa, and a written informed consent was obtained from one or both parents of the newborns.

The mean number of accepted epochs for the ADS register was 587, 612, and 614 for word-initial, word-medial, and word-final syllables, respectively. The mean number of accepted epochs for the IDS register was 611, 613, and 609 for word-initial, word-medial, and word-final syllables respectively.

Stimuli

Each speech register (ADS and IDS) consisted of 18 natural Finnish syllables, 600 ms in duration, separated by 150 ms of silence (inter-stimulus interval) throughout the entire stream. A total of 12 pseudo-words were created from the syllables, with 6 pseudo-words in each condition, and presented so that each pseudo-word was never immediately repeated, and every pseudo-word followed every other pseudo-word equally often (transitional probability from a word to any other word being 1/5), keeping the word order otherwise random. There were ten 3.55-minute-long presentation blocks consisting of this, seemingly random, stream of pseudo-words. Every other block consisted of ADS pseudo-words and the rest were IDS blocks. The order of the blocks was counterbalanced across participants. The total duration of the experiment was approximately 40 minutes.

Speech stimuli were cut from natural utterances of a female speaker recorded in an anechoic chamber [38]. Four different types of syllables were used: /k/ + vowel, /s/ + vowel, long vowel, and diphthong. The syllables were chosen so that the fundamental frequency of the voice remained relatively stable throughout the syllables. From these syllables, IDS was created by overlaying the prosodic contours excised from naturally spoken IDS registers onto the syllables using PRAAT [64].

The ADS register average F0 was 191 Hz (range = 181–212 Hz) and the average F0 in the IDS was 212 Hz (range = 180–235 Hz). The larger range in the IDS register reflects the exaggerated pitch peaks, which reached an average of 381 Hz whereas the ADS register reached an average of 228 Hz. To make sure that pitch peaks did not mark word boundaries, we ensured that no syllables were consistently stressed for any given syllable position, and pitch peaks were distributed evenly across the syllables in words based on the fundamental frequency contours. Fig 1 shows an example of the stimuli used.

EEG recording

The EEG was recorded in a quiet room from 8 standard electrode sites spanning the scalp. Single-use electrodes were used for recording the EEG (electrodes F3, F4, C3, C4, T3, T4, P3, and
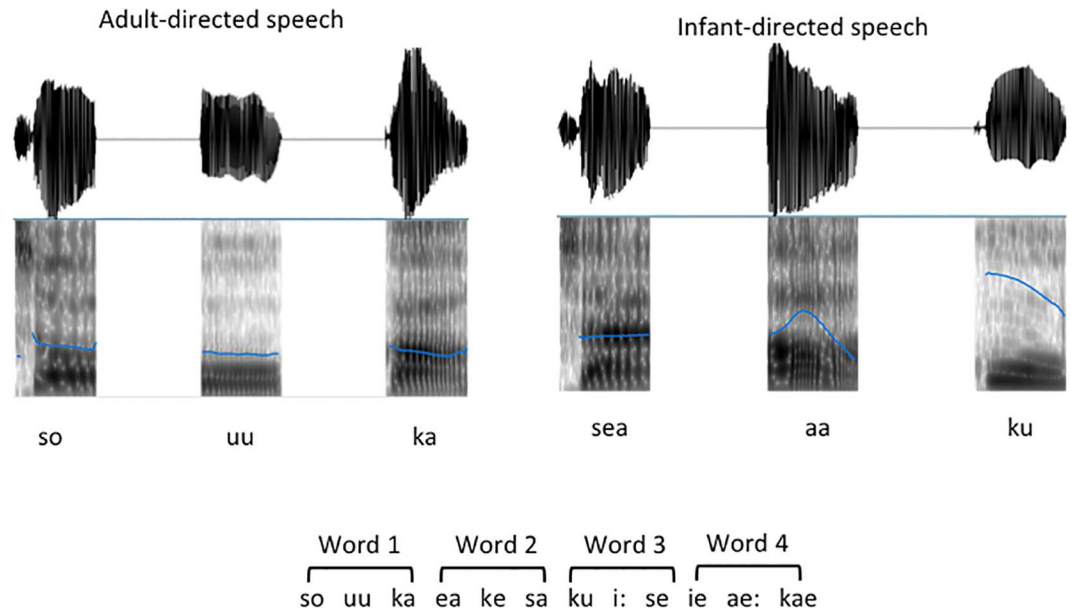
Fig 1. Example F0 contours, spectrograms of syllables, and schematic of the experimental procedure. Top panel. Example F0 contours and spectrograms of syllables presented in the ADS (left) and IDS (right) registers. Bottom panel. Schematic of the experimental procedure, i.e., 4 pseudo-words from the speech stream.

P4 according to the 10–20 system), mastoids, and EOG from the canthus and below the eye. Linked mastoids were used as a reference. Sounds were presented through two loudspeakers placed 20 cm from both sides of the infant's head. The EEG had a sampling rate of 250 Hz, and was digitally filtered offline (bandpass 0.2–20 Hz).

The EEG measurement was divided into 4 recording sessions, each approximately 10 minutes in duration. The blocks were further divided into epochs between -100 ms pre-stimulus onset to 750 ms, i.e., the duration of one syllable including the silent intervals after the syllables. After baseline correction to the pre-stimulus interval, the epochs with artifacts exceeding ±150 μV were discarded. Due to low signal-to-noise ratio (SNR), the data obtained from temporal electrode sites T3 and T4 were omitted from the statistical analysis.

Measurement windows for the 0–100 ms, 200–400 ms and 450–650 ms components were based on previous studies of statistical learning [15, 37, 38], inspection of individual averages for each newborn at each electrode site and grand averages in order to capture effects across conditions. The use of measurement windows was also a conservative choice due to variability in individual peak latencies.

Statistical analyses

To assess the effects of exaggerated pitch on statistical learning, three separate 4-way repeated measures ANOVAs were conducted for each measurement window. Each ANOVA consisted of 4 within subject factors: speech register (ADS vs. IDS), syllable position (word-initial vs. word-medial vs. word-final), electrode site (frontal vs. central vs. parietal), and hemisphere (left vs. right) as within subject factors. The main effects and interactions for each 4-way ANOVA are presented separately for each measurement window. In our design, all prospective amplitude differences in the ERPs between the two conditions would reflect differences in learning as a function of exposure, because the stimuli were counterbalanced across speech registers and participants.

In a second set of analyses (see supplemental data), we directly compared ERP amplitudes and amplitude changes as a function of exposure. Separate statistical analyses were conducted cumulatively for each of the 3.55-minute exposure blocks for the ADS and IDS registers. This analysis allowed us to examine the contribution of each successive block to the cumulative grand averaged ERPs as a function of exposure. It also provided information on the time course of segmentation.

For each ANOVA, Greenhouse-Geisser sphericity corrections were applied when appropriate. Partial-eta-squared (η_p^2) was calculated for each main effect and interaction. The Bonferroni correction was applied to multiple within-subject comparisons. Post hoc tests were conducted using Tukey's HSD method. Planned comparisons were reported as significant at the .05 level and Cohen's d effect sizes were calculated using means and original standard deviations to determine the proportion of total variance attributed to each significant effect.

Results

Fig 2 shows grand averaged ERP response to each syllable in the ADS (top panel) and IDS (bottom panel) registers collapsed across the 10 measurement blocks and all subjects. Our ERP results show clear responses in the 0–100 ms, 200–400 ms, and 450–650 ms measurement windows for the 3 syllable positions, and these responses differ for the ADS and IDS registers. Compared with the ADS register, the IDS register elicited more negative ERPs in the 0–100 ms measurement window, a larger positive response to word-medial syllables in the 200–400 ms measurement window, and larger negative responses to word-initial syllables in the 450–650 ms measurement window. For the ADS register, ERPs were dominated by a larger positive deflection spanning the entire measured response.

0–100 ms measurement window

Fig 3 shows the overall mean amplitude for the ADS and IDS registers in the 0–100 ms measurement window.

Main effect: A four-way repeated ANOVA (2 register x 3 syllable location x 3 electrode site x 2 hemisphere) revealed a significant main effect for speech register, [$F_{(1,22)} = 4.5$, $p < 0.05$,
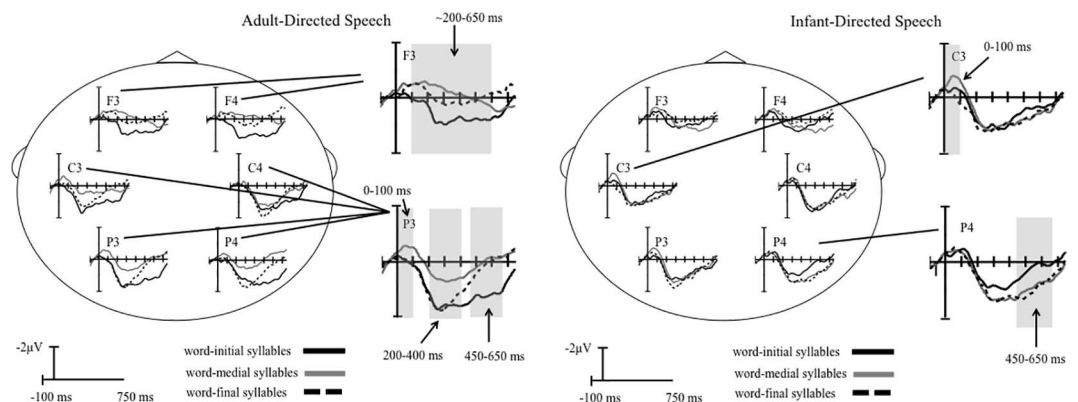

Fig 2. Grand-averaged ERPs for word-initial, word-medial, and word-final syllables for ADS and IDS registers collapsed across the 10 exposure blocks. Grand-averaged ERPs to the word-initial (black line), word-medial (light gray), and word-final (dashed line) syllables in the tri-syllabic pseudo-words for the ADS (top panel) and IDS (right panel) registers collapsed across the 10 exposure blocks. Infants heard each syllable 111 times. Enlarged area displays results at represented electrode sites for each measurement window. Grey bars denote significant differences in mean amplitudes between syllable positions. Negative voltages (microvolts) are plotted upward.

Effect of Speech Register
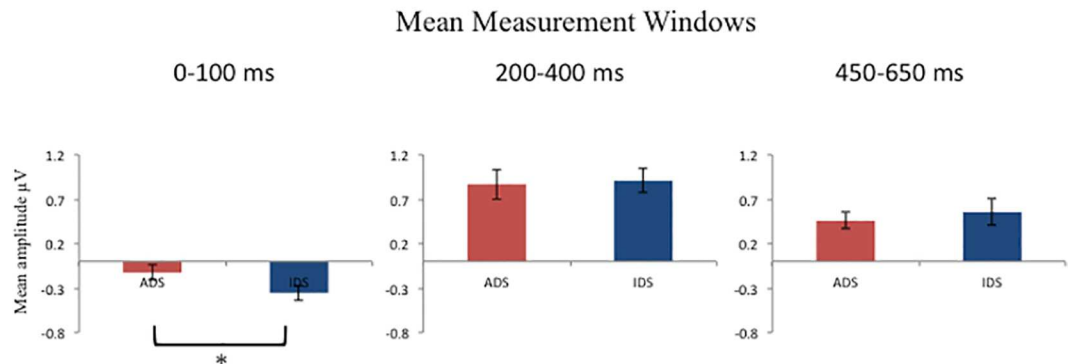
Mean Measurement Windows

Fig 3. Mean amplitude for the ADS and IDS registers in the 0–100 ms, 200–400 ms and 450–650 ms measurement windows. Mean amplitude (in microvolts) for the ADS and IDS registers in the 0–100 ms (left panel), 200–400 ms (middle panel) and 450–650 ms (right panel) measurement windows averaged across the 10 exposure blocks, 3 syllable positions and 6 electrode sites. Asterisks indicate significant differences.

doi:10.1371/journal.pone.0162177.g003

$\eta_p^2 = 0.17$, observed power = 0.53], reflecting larger negative mean amplitudes for the IDS (M = -0.354μv, $S.E.$ = 0.085) than the ADS (M = -0.127μv, $S.E.$ = 0.082) register.

Interaction: To further explore the marginally significant position x hemisphere interaction ($F_{(2,44)}$ = 2.586, p = 0.054, η_p^2 = 0.124, observed power = 0.570), two 3-way (syllable position x electrode site x hemisphere) repeated measures ANOVAs for the ADS and IDS registers were conducted. These tests indicated the trend for the syllable position x hemisphere interaction was driven by the significant interaction in the IDS register [$F_{(2,44)}$ = 3.548, $p < 0.05$, η_p^2 = 0.139, observed power = 0.630]. Post-hoc tests for the IDS register indicated larger negative mean amplitudes over the left hemisphere to word-medial syllables (M = -0.553μv, $S.E.$ = 0.153) than word-initial (M = -0.254μv, $S.E.$ = 0.132, $p < 0.05$, d = 0.762) and word-final sylla-bles (M = -0.277μv, $S.E.$ = 0.123, $p < 0.05$, d = 0.642).

The syllable position x hemisphere interaction was not significant for ADS [$F_{(2,44)}$ = 0.748, p = 0.48]; however, there was a significant main effect for hemisphere [$F_{(1,22)}$ = 4.435, $p < 0.05$, η_p^2 = 0.168, observed power = 0.521], reflecting significantly larger negative mean amplitudes over the left (M = -0.180μv, $S.E.$ = 0.09) than the right (M = -0.046μv, $S.E.$ = 0.098) hemisphere. No other main effects or interactions were significant.

200–400 ms measurement window

Fig 3 shows the overall mean amplitude for the ADS and IDS registers in the 200–400 ms mea-surement window.

Main effect: The same four way repeated-measures ANOVA was conducted in this time win-dow. In contrast to the 0–100 ms measurement window, the main effect of speech register was not significant, [$F_{(1,22)}$ = 0.314, p = 0.843]. There was a main effect of electrode site, [$F_{(2,30)}$ = 15.744, $p < 0.05$, η_p^2 = 0.417, observed power = 0.989], indicating that overall, mean amplitudes were larger over parietal (M = 0.99μv, $S.E.$ = 0.168) and central (M = 1.095μv, $S.E.$ = 0.143) than frontal (M = 0.13μv, $S.E.$ = 0.135, frontal vs. parietal: $p < 0.05$, d = 1.15; and frontal vs. central: $p < 0.05$, d = 1.29) electrode sites.

Interaction: As predicted, there was a significant speech register x syllable position interac-tion [$F_{(2,44)}$ = 3.259, $p < 0.05$, η_p^2 = 0.129, observed power = 0.591], reflecting a significant effect of syllable position for the ADS [$F_{(2,44)}$ = 5.037, $p < 0.05$, η_p^2 = 0.186, observed power = 0.790],

but not the IDS [$F_{(2,44)}$ = 0.072, p = 0.931] register. Post-hoc tests for the ADS register indicated smaller mean amplitudes to word-medial syllables (M = -0.277μv, $S.E.$ = 0.123) than to word-initial (M = 1.238μv, $S.E.$ = 0.267, p<0.05, d = 0.716) and word-final syllables (M = 1.014μv, $S.E.$ = 0.194, p<0.05, d = 0.620). The electrode site x hemisphere interaction was also significant [$F_{(2,44)}$ = 5.002, p<0.05, η_p^2 = 0.185, observed power = 0.787], indicated larger mean amplitudes over central than parietal electrode sites in the left hemisphere (p = 0.011, d = 0.64).

450–650 ms measurement window

As seen in Fig 3, the 450–650 ms mean amplitude response differs in both the topographical distribution and polarity for the ADS and IDS registers.

Main effect: There were no significant main effects in this measurement window.

Interaction: As predicted, a 4-way repeated-measures ANOVA revealed a significant speech register x syllable position interaction [$F_{(2,44)}$ = 7.813, p<0.05, η_p^2 = 0.262, observed power = 0.938], reflecting significant speech register effects for syllable position. To further explore this interaction, two separate 3-way (syllable position x electrode site x hemisphere) ANOVAs were conducted for the ADS and IDS registers. These tests indicated that the effect for syllable position was significant for the ADS, [$F_{(2,44)}$ = 6.959, p<0.05, η_p^2 = 0.240, observed power = 0.908], but not the IDS register [$F_{(2,44)}$ = 1.019, p = 0.369]. As shown in Fig 3, for the ADS register in the 450–650 ms measurement window, amplitudes to word-initial syllables (M = 1.059μv, $S.E.$ = 0.196) were larger than word-medial (M = 0.138μv, $S.E.$ = 0.197, p = 0.025, d = 0.977) and word-final (M = 0.174μv, $S.E.$ = 0.166, p = 0.007, d = 1.016) syllables. Amplitudes for word-medial and word-final syllables did not differ significantly from each other (p = 1.00).

Paired comparisons were conducted to assess the distribution of the ERP effects for the ADS and IDS registers. These tests revealed significant speech register x syllable position interactions over the right frontal [$F_{(2,44)}$ = 4.146, p<0.05, η_p^2 = 0.159, observed power = 0.702], left parietal [$F_{(2,44)}$ = 3.241, p<0.05, η_p^2 = 0.128, observed power = 0.588], and right parietal [$F_{(2,44)}$ = 11.681, p<0.05, η_p^2 = 0.347, observed power = 0.991] electrode sites.

As predicted, distribution of the ERP effects differed for the ADS and IDS registers, with larger positive mean amplitudes elicited to word-initial syllables over the right frontal and the left and right parietal electrode sites as compared to word-medial and word-final syllables (p<0.05, d = 0.71–1.05) in response to the ADS register. In contrast, for the IDS register, an effect of syllable position was observed only over the right parietal electrode site and was driven by significantly larger negative mean amplitudes for word-initial syllables as compared to word-medial and word-final syllables (p<0.05, d = 0.60 and 0.66, respectively).

Discussion

The current study examined how exaggerated pitch, typical of IDS, affects the brain response in a statistical learning paradigm in newborns. We hypothesized that presenting the speech with exaggerated pitch would facilitate statistical learning in sleeping newborns, and more importantly, that the ERPs would vary as a function of two different aspects of the speech stimuli that were manipulated in the current experiment: (1) the transitional probabilities between syllables (syllable position), and (2) speech register (ADS vs. IDS). Interactions between speech register and syllable position were predicted in 3 measurement windows: at the very early 0–100 ms window, differences were expected to show an enhanced ERP for syllables occurring in predictable over unpredictable positions within pseudo-words, regardless of speech register; at the 200–400 ms window, the IDS (but not ADS) register was expected to show enhanced ERP responses to the 3 syllable positions based on the acoustic saliency of

individual phonemes; and at the late 450–650 ms window, differences were expected to show patterns consistent with processing efficiency, with effects being more broadly distributed for the ADS register versus the IDS register.

Our results were consistent with these hypotheses. Event-related brain potentials differed as a function of syllable position within a pseudo-word and speech register. Our results showed that although overall ERP responses were larger in the earliest 0–100 ms measurement window for the IDS over ADS register, within each register the ERPs were larger for the syllable position with the highest transitional probability. Detecting the most predictable syllable in continuous speech may promote segmentation during infancy, because attending to the probabilistic information in language input identifies the critical elements (phonemes and words) and thus supports learning. The brain formulates predictions based on the incoming statistical regularities, allowing for more efficient processing and pattern recognition. In the current study, increasing the saliency of the input, for example by varying the prosodic contours of individual phonemes, appears to enhance statistical processing, making the response to critical features of the stimuli more robust. We suggest interpreting the early 0–100 ms effect as reflecting efficient memory trace formation for the statistical regularities that results from the same prediction based mechanism linked to predictive coding (see also [32, 38]). This early response may function to allocate cognitive resources to processing the raw statistics of the input, resulting in more efficient processing of the statistical input. The more focal effect of syllable position for IDS versus ADS pseudo-words seen in the 450–650 ms measurement window could be interpreted as more efficient processing of statistical patterns over the 40-minute exposure period during the IDS blocks. Selective attention has been shown to play a role in successful segmentation in awake adults [12, 32], and in natural learning environments IDS may be beneficial in early word learning by heightening attention to the input more generally, which in turn expedites the detection of statistical regularities.

We also observed an interaction between speech register and syllable position in the 200–400 ms measurement window (see [33]). A previous study has shown an enhanced positivity to nonnative language contrasts for 11- month old infants with a significantly larger vocabulary size at 18, 22, 25, 27 and 30 months compared to infants with smaller vocabularies [42]. Rivera-Gaxiola et al. [42] attribute the positive ERP to enhanced acoustic processing. Work with adults is consistent with this interpretation, showing an enhanced P200 response occurs with intense auditory discrimination training [65–70] or with the implantation of cochlear implants in congenitally deaf patients [71]. Other work has shown that presenting both statistical information (transitional probabilities between syllables) and the prosodic cue of increased pitch to word-initial syllables results in an enhanced positivity that peaks at approximately 225 ms (P200) for predictable, but not unpredictable, word streams in adults [33]. The authors of this later study posit that the enhanced P200 reflects enhanced auditory learning, with pitch cues functioning as attentional cues that prime language segmentation [33]. Although speculative, the significantly larger positivity in the 200–400 ms measurement window for the IDS register in this study may also be linked to auditory learning, reflecting enhanced, or more in-depth processing of the acoustic properties for the IDS register that occurs independent of their position within the pseudo-words. Importantly, our data suggest that such an enhancement may be seen even when sleeping newborns are exposed to both exaggerated pitch and statistical regularities.

Lastly, we hypothesized that the observed broader distribution of brain activity across electrode sites after 300 ms for ADS versus IDS register is linked to less successful pattern recognition, or segmentation processes. Previous work has argued that effects of syllable position in this latency window reflect the process of statistical learning with regard to word segmentation in both infants [38] and adults [15, 32, 33]. When averaged over the 10 exposure blocks, word-

initial syllables elicited larger mean amplitudes for the ADS register that were broadly distributed over the right frontal and bilateral parietal electrodes. In contrast, the IDS register showed an effect of syllable position that was specific to the right parietal electrode site, and driven by more negative mean amplitudes to the word-initial syllables as compared to word-medial and final-syllables. The greater bilateral and anterior distribution of this effect for the ADS register suggests that processing statistical regularities of sounds with exaggerated pitch results in the recruitment of different brain areas than processing statistical regularities without exaggerated pitch. Research with older infants and adults have found links between processing efficiency and the distribution of the brain response, with more focal activity linked to more efficient processing for familiar and/or known words [23, 31, 46, 72]. Because infants in this experiment received equal amounts of exposure to the ADS and IDS inputs, these results cannot reflect differences in experience with the statistical regularities, but rather suggest that speech register modulates activity based on syllable position within the pseudo-word.

Our findings differ from previous studies in both the polarity and topographical distribution of effects. Using the same ADS stimuli as the current study, Teinonen and colleagues [38] reported an ERP response that differentiated between syllable positions, however, the effect was driven by a larger negative response for word-initial syllables that began after 300 ms from syllable onset. The differences between our findings and those reported elsewhere may reflect processing differences between speech and non-speech input [37], differences in baseline correction, stimulus duration, interstimulus interval, and our choice of using an alternating block design in the current study in which the speech register switched during the ERP measurement. It is possible that switching between IDS and ADS registers during testing in this study may have had an effect on processing by creating a greater cognitive load for the ADS register. Kudo et al. [37] hypothesized that the broad positive deflection observed in their results may reflect the immaturity of neural and glial cells in the newborn, leading to slower perceptual processing.

Time course analysis: Effect of statistical input as a function of exposure

Since newborn ERPs show tremendous variability between individuals and also within an individual infant across experiments, we conducted a time-course analysis. This analysis examined the accumulated response to assess the effects of learning and stability of the response as a function of increased exposure to the statistical input. This approach allowed us to observe the temporal unfolding of ERP activity for the ADS and IDS registers (see Supplement). Our accumulative analysis revealed that the effect of syllable position for the ADS register was initially a broad positivity, similar to the effect observed by Kudo et al. [37]. However, the current study found that this initial broad positivity evolved into three distinct components with increased exposure. For the ADS register, the effect of syllable position was initially observed as a broad positivity that spanned the three measurement windows and multiple electrode sites; however, within the first 3 exposure blocks, the response evolved into three distinct components that differed in spatial distribution. In the 0–100 ms measurement window, responses to word-medial syllables elicited significantly larger negative responses over the left frontal, central and parietal electrode sites in the first exposure block. With continued exposure, this effect became attenuated at the frontal electrode site and was no longer significant when averaged over the first 5 exposure blocks. In the 200–400 ms and 450–650 ms measurement windows, the effect of syllable position differed in both polarity and distribution. Over anterior and posterior electrode sites, these effects began with larger mean amplitudes to word-initial syllables, an effect that continued over the right frontal electrode sites throughout the 10 exposure blocks, whereas effects over central and parietal electrode sites evolved into two distinct components over the course of exposure, suggesting distinct processes. Importantly, we observed that the outcomes

of experience with statistical regularities differed for the ADS and IDS registers, and resulted in a 'narrowing' or differentiation of the ERP into distinct components with different spatial distributions. Our data, combined with the results reported in Kudo et al. [37] suggest that the initial dominant positive response is comprised of multiple generators that underlie distinct processes that become more specialized with experience.

Conclusions

These data illustrate the speed with which the newborn brain encodes both the acoustic and statistical regularities contained in the ambient language input, thus reflecting a dynamic learning mechanism that is sensitive to input quantity and quality (see [2]). Importantly, our results show that the speech register used when addressing infants, even sleeping newborns, is an important factor in determining the patterns of brain activity, even during the earliest stages of language acquisition and provide some evidence that learning about the statistical regularities in speech is more robust when the speech is produced with exaggerated pitch contours. Our results, taken together with previous reports, are suggestive of a facilitative effect of exaggerated pitch for detecting the statistical regularities in the auditory language input that can be observed from the earliest stages of language acquisition [10].

The results from the current study lend support to the view that the same learning mechanism can yield different results, and is sensitive to both experience [13, 73] and input [74]. Our results suggest that even in a lowered arousal state i.e., during sleep, some aspect of the prosodic characteristics of IDS facilitate newborns' access to the statistical structure of speech; however, they do not address the exact mechanism through which this occurs. One possibility is that IDS is better at attracting and sustaining infants' processing resources as compared to ADS. There is a large body of research consistent with the view that IDS is more likely to hold infants' attention than ADS ([75], see [76] for discussion). Follow-up studies will assess the long-term impact of IDS on language exposure. Continuing research on how attention and memory for the acoustic modifications of IDS shift as a function of language experience will advance our understanding of the interaction between the factors that guide language learning.

Supporting Information

S1 Fig. ERPs averaged cumulatively over the 10 blocks of exposure. Mean amplitudes (in microvolts) for the cumulative responses across the 10 exposure blocks for the ADS (left panel) and IDS (right panel) registers over the left and right parietal electrode sites. Grey bars denote significance differences between syllable positions ($p < 0.05$).
(TIF)

S1 File. Time course analysis examining the cumulative responses across the 10 exposure blocks for the ADS and IDS registers.
(DOCX)

Acknowledgments

Author Contributions

Conceptualization: ANB TT MH.

Formal analysis: ANB.

Funding acquisition: ANB.

Investigation: ANB.

Methodology: ANB TT MH.

Software: ANB.

Supervision: MH MT.

Visualization: ANB TT MH MT.

Writing – original draft: ANB.

Writing – review & editing: ANB TT MH MT.

References

1. Best CC, McRoberts GW. Infant perception of non-native consonant contrasts that adults assimilate in different ways. Lang Speech. 2003; 46(Pt 2–3):183–216. PMID: 14748444

2. Kuhl PK, Conboy BT, Coffey-Corina S, Padden D, Rivera-Gaxiola M, Nelson T. Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). Philos Trans R Soc Lond B Biol Sci. 2008; 363(1493):979–1000. PMID: 17846016

3. Werker JF, Curtin S. PRIMIR: a developmental model of speech processing. Lang Learn Dev. 2005; 1 (2):197–234.

4. Goodsitt JV, Morgan JL, Kuhl PK. Perceptual strategies in prelingual speech segmentation. J Child Lang. 1993; 20(2):229–52. PMID: 8376468

5. Maye J, Weiss DJ, Aslin RN. Statistical phonetic learning in infants: facilitation and feature generalization. Dev Sci. 2008; 11(1):122–34. doi: 10.1111/j.1467-7687.2007.00653.x PMID: 18171374

6. Maye J, Werker JF, Gerken L. Infant sensitivity to distributional information can affect phonetic discrimination. Cognition. 2002; 82(3):B101–B11. PMID: 11747867

7. Saffran JR, Aslin RN, Newport EL. Statistical learning by 8-month-old infants. Science. 1996; 274 (5294):1926–8. PMID: 8943209

8. Kuhl PK, Williams KA, Lacerda F, Stevens KN, Lindblom B. Linguistic experience alters phonetic perception in infants by 6 months of age. Science. 1992; 255(5044):606–8. PMID: 1736364

9. Thiessen ED, Saffran JR. When cues collide: Use of stress and statistical cues to word boundaries by 7- to 9- month-old infants. Dev Psychol. 2003; 39(4):706–16. PMID: 12859124

10. Thiessen ED, Hill EA, Saffran JR. Infant-directed speech facilitates word segmentation. Infancy. 2005; 7(1):53–71. doi: 10.1207/s15327078in0701_5

11. Alho K, Sainio K, Sajaniemi N, Reinikainen K, Näätänen R. Event-related brain potential of human newborns to pitch change of an acoustic stimulus. Electroencephalogr Clin Neurophysiol. 1990; 77(2):151–5. PMID: 1690117

12. Toro JM, Sinnett S, Soto-Faraco S. Speech segmentation by statistical learning depends on attention. Cognition. 2005; 97(2):B25–B34. PMID: 16226557

13. Tremblay P, Baroni M, Hasson U. Processing of speech and non-speech sounds in the supratemporal plane: Auditory input preference does not predict sensitivity to statistical structure. Neuroimage. 2013; 66(1):318–32. doi: 10.1016/j.neuroimage.2012.10.055

14. Cunillera T, Gomila A, Rodriguez-Fornells A. Beneficial effects of word final stress in segmenting a new language: evidence from ERPs. BMC Neurosci. 2008; 9:23. doi: 10.1186/1471-2202-9-23 PMID: 18282274

15. Sanders LD, Newport EL, Neville HJ. Segmenting nonsense: an event-related potential index of perceived onsets in continuous speech. Nat Neurosci. 2002; 5(7):700–3. PMID: 12068301

16. Sanders LD, Neville H. An ERP study of continuous speech processing II. Segmentation, semantics, and syntax in non-native speakers. Brain Res Cogn Brain Res. 2003; 15(3):214–37. PMID: 12527096

17. Saito Y, Aoyama S, Kondo T, Fukumoto R, Konishi N, Nakamura K, et al. Frontal cerebral blood flow change associated with infant-directed speech. Arch Dis Child Fetal and Neonatal Ed. 2007; 92(2): F113–F6.

18. Saito Y, Kondo T, Aoyama S, Fukumoto R, Konishi N, Nakamura K, et al. The function of the frontal lobe in neonates for response to a prosodic voice. Early Hum Dev. 2007; 83(4):225–30. doi: 10.1016/j. earlhumdev.2006.05.017 PMID: 16839715

19. Santesso DL, Schmidt LA, Trainor LJ. Frontal brain electrical activity (EEG) and heart rate in response to affective infant-directed (ID) speech in 9-month-old infants. Cognition. 2007; 65(1):14–21.

20. Zangl R, Mills DL. Increased brain activity to infant-directed speech in 6-and 13-month-old infants. Infancy. 2007; 11(1):31–62. doi: 10.1207/s15327078in1101_2

21. Zhang Y, Koerner T, Miller S, Grice-Patil Z, Svec A, Akbari D, et al. Neural coding of formant-exaggerated speech in the infant brain. Dev Sci. 2011; 14(3):566–81. doi: 10.1111/j.1467-7687.2010.01004.x PMID: 21477195

22. Mills DL, Coffey-Corina S, Neville HJ. Language comprehension and cerebral specialization from 13 to 20 months. Dev Neuropsychol. 1997; 13(3):397–445.

23. Mills DL, Prat C, Stager C, Zangl R, Neville H, Werker J. Language experience and the organization of brain activity to phonetically similar words: ERP evidence from 14- and 20-month-olds. J Cogn Neurosci. 2004; 16(8):1452–64. PMID: 15509390

24. Bosseler AN, Taulu S, Pihko E, Mäkelä JP, Imada T, Ahonen A, et al. Theta brain rhythms index perceptual narrowing in infant speech perception. Front Psychol. 2013; 4:690. doi: 10.3389/fpsyg.2013. 00690 PMID: 24130536

25. Callan DE, Jones JA, Callan AM, Akahane-Yamada RN-. Phonetic perceptual identification by native- and second-language speakers differentially activates brain regions involved with acoustic phonetic processing and those involved with articulatory-auditory/orosensory internal models. Neuroimage. 2004; 22(3):1182–1194. PMID: 15219590

26. Golestani N, Zatorre RJ. Learning new sounds of speech: reallocation of neural substrates. NeuroImage. 21(2): 494–506. PMID: 14980552

27. Guenther FH, Nieto-Castanon A, Ghosh SS, Tourville JA. Representation of sound categories in auditory cortical maps. J Speech Lang Hear Res. 2004; 47(1):46–57. PMID: 15072527

28. Zhang Y, Kuhl PK, Imada T, Iverson P, Pruitt J, Stevens EB, et al. Neural signatures of phonetic learning in adulthood: a magnetoencephalography study. Neuroimage. 2009; 46(1):226–40. doi: 10.1016/j. neuroimage.2009.01.028 PMID: 19457395

29. Zhang Y, Kuhl PK, Imada T, Kotani M, Tohkura Y. Effects of language experience: neural commitment to language-specific auditory patterns. Neuroimage. 2005; 26(3):703–20. PMID: 15955480

30. Mills DL, Prat C, Zangl R, Stager CL, Neville HJ, Werker JF. Language experience and the organization of brain activity to phonetically similar words: ERP evidence from 14-and 20-month-olds. J Cogn Neurosci. 2004; 16(8):1452–64. doi: 10.1162/0898929042304697 PMID: 15509390

31. Conboy BT, Mills DL. Two languages, on developing brain: event-related potentials to words in bilingual toddlers. Dev Sci. 2006; 9(1):F1–F12. PMID: 16445386

32. Abla D, Katahira K, Okanoya K. On-line assessment of statistical learning by event-related potentials. J Cogn Neurosci. 2008; 20(6):952–64. doi: 10.1162/jocn.2008.20058 PMID: 18211232

33. Cunillera T, Toro JM, Sebastián-Gallés N, Rodríguez-Fornells A. The effects of stress and statistical cues on continuous speech segmentation: an event-related brain potential study. Brain Res 2006; 1123(1):168–78. PMID: 17064672

34. Teinonen T, Huotilainen M. Implicit Segmentation of a Stream of Syllables Based on Transitional Probabilities: an MEG Study. J Psycholinguist Res. 2012; 41(1):71–82. doi: 10.1007/s10936-011-9182-2 PMID: 21993901

35. Kutas M, Hillyard SA. Brain potentials during reading reflect word expectancy and semantic association. Nature. 1984; 307(5947):161–3. PMID: 6690995

36. Hinojosa JA, Moreno EM, Casado P, Muñoz F, Pozo MA. Syntactic expectancy: An event-related potentials study. Neurosci Lett. 2005; 378(1):34–9. PMID: 15763168

37. Kudo N, Nonaka Y, Mizuno N, Mizuno K, Okanoya K. On-line statistical segmentation of a non-speech auditory stream in neonates as demonstrated by event-related brain potentials. Dev Sci. 2011; 14 (5):1100–6. doi: 10.1111/j.1467-7687.2011.01056.x PMID: 21884325

38. Teinonen T, Fellman V, Näätänen R, Alku P, Huotilainen M. Statistical language learning in neonates revealed by event-related brain potentials. BMC Neurosci. 2009; 10:21. doi: 10.1186/1471-2202-10-21 PMID: 19284661

39. Friederici AD, Friedrich M, Christophe A. Brain responses in 4-month-old infants are already language specific. Curr Biol. 2007; 17(14):1208–11. PMID: 17583508

40. Friedrich M, Herold B, Friederici AD. ERP correlates of processing native and non-native language word stress in infants with different language outcomes. Cortex. 2009; 45(5):662–76. doi: 10.1016/j.cortex.2008.06.014 PMID: 19100528

41. Garcia-Sierra A, Rivera-Gaxiola M, Conboy BT, Romo H, Percaccio CR, Klarman L, et al. Bilingual language learning: An ERP study relating early brain responses to speech, language input and later word production. J Phon. 2011; 39:456–557.

42. Rivera-Gaxiola M, Klarman L, Garcia-Sierra A, Kuhl PK. Neural patterns to speech and vocabulary growth in American infants. Neuroreport. 2005; 16(5):495–8. PMID: 15770158

43. Morr ML, Shafer VL, Kreuzer JA, Kurtzberg D. Maturation of mismatch negativity in typicaly developing infants and preschool children. Ear Hear. 2002; 23(2):118–36. PMID: 11951848

44. Shafer LV, Yu HY, Datta H. The development of English vowel perception in monolingual and bilingual infants: Neurophysiological correlates. J Phon. 2011; 39(4):527–45. PMID: 22046059

45. Kooijman V, Junge C, Johnson EK, Hagoort P, Cutler A. Predictive brain signals of linguistic development. Front Psychol. 2013; 4:25. doi: 10.3389/fpsyg.2013.00025 PMID: 23404161

46. Mills DL, Neville HJ. Electrophysiological studies of language and language impairment in children. Semin Pediatr Neurol 1997; 4(2):125–34. PMID: 9195670

47. de Gardelle V, Waszczuk M, Egner T, Summerfield C. Concurrent repetition enhancement and suppression responses in extrastriate visual cortex. Cereb Cortex. 2013; 23(9):2235–44. doi: 10.1093/cercor/bhs211 PMID: 22811008

48. Turk-Browne NB, Yi DJ, Leber AB, Chun MM. Visual quality determines the direction of neural repetition effects. Cereb Cortex. 2007; 17(2):425–33. PMID: 16565294

49. Henson R, Shallice T, Dolan R. Neuroimaging evidence for dissociable forms of repetition priming. Science. 2000; 287(5456):1269–72. PMID: 10678834

50. Müller NG, Strumpf H, Scholz M, Baier B, Melloni L. Repetition suppression versus enhancement- it's quantity that matters. Cereb Cortex. 2012; 23(2):315–22. doi: 10.1093/cercor/bhs009 PMID: 22314047

51. Massaro DW. Perceiving Talking Faces: From Speech Perception to a Behavioral Principle. Cambridge, MA: MIT Press; 1993.

52. Massaro DW, Cohen MM. Phonological context in speech perception. Percept Psychophys. 1983; 34(4):338–48. PMID: 6657435

53. Rescorla RA. Predictability and number of pairings in Pavlovian fear conditioning. Psychon Sci. 1966; 4(11):383–4. doi: 10.3758/BF03342350

54. Saffran JR. Constraints on statistical language learning. J Mem Lang. 2002; 47:172–96. doi: 10.1006/jmla.2001.2839

55. Friston K. A theory of cortical responses. Philos Trans R Soc Lond B Biol Sci. 2005; 360(1456):815–36. PMID: 15937014

56. Wacongne C, Changeux JP, Dehaene S. A neuronal model of predictive coding accounting for the mismatch negativity. J Neurosci. 2012; 32(11):3665–78. doi: 10.1523/JNEUROSCI.5003-11.2012 PMID: 22423089

57. Egner T, Monti JM, Summerfield C. Expectation and surprise determine neural population responses in the ventral visual stream. J Neurosci. 2010; 30(49):16601–8. doi: 10.1523/JNEUROSCI.2770-10.2010 PMID: 21147999

58. Keller GB, Bonhoeffer T, Hübener M. Sensorimotor mismatch signals in primary visual cortex of the behaving mouse. Neuron. 2012; 74(5):809–15. doi: 10.1016/j.neuron.2012.03.040 PMID: 22681686

59. Meyer M, Sauerland U. A pragmatic constraint on ambiguity detection. Nat Lang Linguist Theory. 2009; 27:139–50. doi: 10.1007/s11049-008-9060-2

60. Cheour-Luhtanen M, Alho K, Kujala T, Saino K, Reinikainen K, Renlund M, et al. Mismatch negativity indicates vowel discrimination in newborns. Hear Res. 1995; 82(1):53–8. PMID: 7744713

61. Baldeweg T. Repetition effects to sounds: evidence for predictive coding in the auditory system. Trends Cogn Sci. 2006; 10(3):93–4. PMID: 16460994

62. Summerfield C, Egner T. Expectation (and attention) in visual cognition. Trends Cogn Sci. 2009; 13(9):403–9. doi: 10.1016/j.tics.2009.06.003 PMID: 19716752

63. Grimm S, Schröger E. The processing of frequency deviations within sounds: evidence for the predictive nature of the mismatch negativity (MMN) system. Restor Neurol Neurosci. 2007; 25(3–4):241–9. PMID: 17943002

64. Boersma P, Weenink D. Praat: doning phonetics by computer. 5.3.51, retrieved 2 June 2013 from http://wwwlpraat.org/ed2013.

65. Atienza M, Cantero JL, Dominguez-Marin E. The time course of neural changes underlying auditory perceptual learning. Learn Mem. 2002; 9(3):138–50. PMID: 12075002

66. Bosnyak DJ, Eaton RA, Roberts LE. Distributed auditory cortical representations are modified when non-musicians are trained at pitch discrimination with 40 Hz amplitude modulated tones. Cereb Cortex. 2004; 14(10):1088–99. PMID: 15115745

67. Reinke KS, He Y, Wang CH, Alain C. Perceptual learning modulates sensory evoked response during vowel segregation. Brain Res Cogn Brain Res. 2003; 17(3):781–91. PMID: 14561463

68. Tremblay KL, Kraus N. Auditory training induces asymmetrical changes in cortical neural activity. J Speech Lang Hear Res. 2002; 45(3):564–72. PMID: 12069008

69. Tremblay KL, Kraus N, Mcgee T, Ponton C, Otis B. Central auditory plasticity: changes in the N1-P2 complex after speech-sound training. Ear Hear. 2001; 22(2):79–90. PMID: 11324846

70. Shahin A, Bosnyak DJ, Trainor LJ, Roberts LE. Enhancement of neuroplastic P2 and N1c auditory evoked potentials in musicians. J Neurosci. 2003; 23(13):5545–52. PMID: 12843255

71. Purdy SC, Kelly AS, Thorne PR. Auditory evoked potentials as measures of plasticity in humans. Audiol Neurootol. 2001; 6(4):211–5. PMID: 11694730

72. Mills DL, Coffey-Corina S, Neville HJ. Language acquisition and cerebral specialization in 20-month-old infants. J Cogn Neurosci. 1993; 5(3):317–34. doi: 10.1162/jocn.1993.5.3.317 PMID: 23972220

73. Saffran JR. Words in a sea of sounds: the output of infant statistical learning. Cognition. 2001; 81 (2):149–69. doi: 10.1016/s0010-0277(01)00132-9 PMID: 11376640

74. Saffran JR, Reeck K, Niehbur A, Wilson DP. Changing the tune: the structure of the input affects infants' use of absolute and relative pitch. Dev Sci. 2005; 8(1):1–7. PMID: 15647061

75. Werker JF, Pegg JE, McLeod PJ. A cross-language investigation of infant preference for infant-directed communication. Infant Behav Dev. 1994; 17(3):323–33. doi: 10.1016/0163-6383(94)90012-4

76. Fernald A, Kuhl P. Acoustic determinants of infant preference for motherese speech. Infant Behav Dev. 1987; 10(3):279–93. doi: 10.1016/0163-6383(87)90017-8