

# Auditory evoked potentials to speech and nonspeech stimuli are associated with verbal skills in preschoolers



Soila Kuuluvainen<sup>a,\*</sup>, Alina Leminen<sup>a,b</sup>, Teija Kujala<sup>a</sup>

<sup>a</sup> Cognitive Brain Research Unit, Institute of Behavioral Sciences, University of Helsinki, Finland

<sup>b</sup> Center of Functionally Integrative Neuroscience, Department of Clinical Medicine, Aarhus University, Denmark

## ARTICLE INFO

### Article history:

Received 14 July 2015

Received in revised form

23 December 2015

Accepted 1 April 2016

Available online 14 April 2016

### Keywords:

Auditory

Event-related potential

Speech

Nonspeech

Verbal skills

Children

## ABSTRACT

Children's obligatory auditory event-related potentials (ERPs) to speech and nonspeech sounds have been shown to associate with reading performance in children at risk or with dyslexia and their controls. However, very little is known of the cognitive processes these responses reflect. To investigate this question, we recorded ERPs to semisynthetic syllables and their acoustically matched nonspeech counterparts in 63 typically developed preschoolers, and assessed their verbal skills with an extensive set of neurocognitive tests. P1 and N2 amplitudes were larger for nonspeech than speech stimuli, whereas the opposite was true for N4. Furthermore, left-lateralized P1s were associated with better phonological and prereading skills, and larger P1s to nonspeech than speech stimuli with poorer verbal reasoning performance. Moreover, left-lateralized N2s, and equal-sized N4s to both speech and nonspeech stimuli were associated with slower naming. In contrast, children with equal-sized N2 amplitudes at left and right scalp locations, and larger N4s for speech than nonspeech stimuli, performed fastest. We discuss the possibility that children's ERPs reflect not only neural encoding of sounds, but also sound quality processing, memory-trace construction, and lexical access. The results also corroborate previous findings that speech and nonspeech sounds are processed by at least partially distinct neural substrates.

© 2016 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. Introduction

Even before the development of vocabulary, newborns and infants are biased towards listening to speech over equally complex nonspeech sounds (Vouloumanos and Werker, 2004, 2007). This bias lays the foundation to the development of later language skills via the increasing specialization of the cortex in processing speech (for reviews, see Kuhl, 2004; Kuhl et al., 2008). By preschool age, typically developing children distinguish speech and nonspeech sounds effortlessly, and master basic language skills necessary for learning in a formal school setting. However, little is known of the underlying processes of speech versus nonspeech sound encoding in preschoolers, as no comprehensive studies have been conducted in this age group. This study aims to investigate speech and nonspeech sound processing using cortical auditory event-related potentials (ERPs) and their association with neurocognitive task performance.

In children under 11 years of age, sounds typically elicit a pattern of so-called 'obligatory' ERPs labeled P1-N2-N4 or P100-N250-N450 according to their polarity (positive or negative) and latency (100, 250 or 450 ms; e.g. Pihko et al., 2005; Ponton et al., 2000; Shafer et al., 2015). They are identifiable already in neonates to harmonic tones presented at a slow rate, with P1 increasing in amplitude during the first three months, and N2 becoming increasingly robust between six and nine months of age (Kushnerenko et al., 2002). For syllables, P1 is identifiable already at the youngest age group of three-month-olds, whereas N2 emerges at around six months of age, both stabilizing in amplitude and latency by the age of two years (Shafer et al., 2015). P1 amplitude to syllables increases again at the age of five, remaining stable after that until the age of eight years, whereas N2 amplitude to syllables shows no clear developmental tendencies between ages two and eight years (Shafer et al., 2015).

In contrast to syllables, P1 amplitude to harmonic tones is of similar magnitude at ages four and nine years, decreasing by adulthood (Čeponienė et al., 2002) and it decreases steadily for pure tones from age seven to adulthood (Bishop et al., 2011; Sussman et al., 2008). N2 amplitude to harmonic tones decreases between ages four and nine (Čeponienė et al., 2002) and is stable for pure

\* Corresponding author.

E-mail address: [soila.kuuluvainen@helsinki.fi](mailto:soila.kuuluvainen@helsinki.fi) (S. Kuuluvainen).

tones between ages eight and eleven years (Sussman et al., 2008). No results were reported for N4 in these studies. Taken together, the results suggest that speech and nonspeech sound processing have different developmental trajectories, with turning points at around ages two and five years. However, for a more complete picture, the processing of speech and nonspeech sounds would have to be compared within the same participants. This has been done previously with school-aged children, but not with preschoolers.

Previous studies of speech and nonspeech processing in 8–10-year-old children have given variable results depending on stimulus properties. P1 amplitude was found to be larger for vowels than complex or simple tones (Bruder et al., 2011; Čeponienė et al., 2001) but smaller for syllables than nonspeech analogues (Čeponienė et al., 2005, 2008). The studies of Čeponienė et al. (2005, 2008) suggest that the child P1 is, in the absence of N1, fused together with P2, which in adults is enhanced to discrimination training (Tremblay et al., 2001) suggesting it reflects neural tuning to newly learned sound contrasts. Furthermore, the amplitude of P1 to prototypical vowels was found to correlate inversely with behavioral same-assessment of vowels and with reading speed, so that children with smaller P1s were more accurate in assessing two identical vowels “the same” and could also read more words per minute (Bruder et al., 2011). Therefore, the child P1 was suggested to reflect both sound detection and speech-nonspeech as well as the wideness of neural tuning curves to vowel prototypes (Bruder et al., 2011).

In the same studies, results for N2 amplitude were similarly variable. N2 was smaller (Čeponienė et al., 2001) or equal in size (Bruder et al., 2011) for vowels and simple tones when compared to complex tones, but larger for syllables than nonspeech analogues (Čeponienė et al., 2005, 2008). Since the amplitude of N2 elicited by tone pips was found to increase with repetition in nine-year-olds (Karhu et al., 1997), larger N2s to complex sounds than vowels were interpreted as memory-trace build-up for the unfamiliar stimuli (Čeponienė et al., 2001). In the studies using syllables, N2 and N4 behaved similarly, and were suggested to reflect higher-order sound analysis, such as the content recognition of syllables, scanning for access to semantic representations, or short-term memory retrieval (Čeponienė et al., 2001, 2005, 2008). As N4 was also larger for vowels than simple or complex tones, it is the only component, which has consistently had larger amplitude for speech than nonspeech sounds, and was thus interpreted as an index of sound “speechness” (Čeponienė et al., 2001, 2005, 2008).

A few studies of preschool children with clinical groups also stress the usefulness of ERPs as indexes of language development. For example, Lovio et al. (2010) reported diminished P1 peaks to syllables in 6-year-old children at risk for dyslexia, whereas Hämäläinen et al. (2013) reported abnormally large N2s to a short pseudo-word and its nonspeech counterpart in 6-year-old children who three years later had reading problems. Furthermore, in a longitudinal study, Espy et al. (2004) presented syllables and sinusoidal tones with long, 2.5–4.0 s inter-stimulus intervals (ISI), which produces the child N1 in addition to the P1–N2–N4 complex. Increased N1 amplitudes to both speech and nonspeech stimuli between ages 1 and 4 years were related to poorer pseudo-word reading at school, whereas decreased N2 amplitudes to nonspeech stimuli between ages 4 and 8 years predicted poorer word reading at school.

Here, our goal was to fill a gap in research by contrasting speech and nonspeech sound processing in preschoolers, using syllables and nonspeech stimuli that were carefully matched for acoustic properties with the speech stimuli. As, to our knowledge, there are no such previous studies in six-year-olds, our hypotheses are only tentative. If sound detection quality processing in preschoolers is akin to school-aged children, we will observe smaller P1 but larger N2 and N4 responses to syllables than nonspeech sounds (Bruder

et al., 2011; Čeponienė et al., 2001, 2005, 2008). We will also analyze the relationship between cortical responses and neurocognitive task performance, expecting P1 amplitude to be associated with better phonological skills (Bruder et al., 2011), and larger speech than nonspeech N2/N4s to be associated with better cognitive functioning.

## 2. Methods

### 2.1. Participants

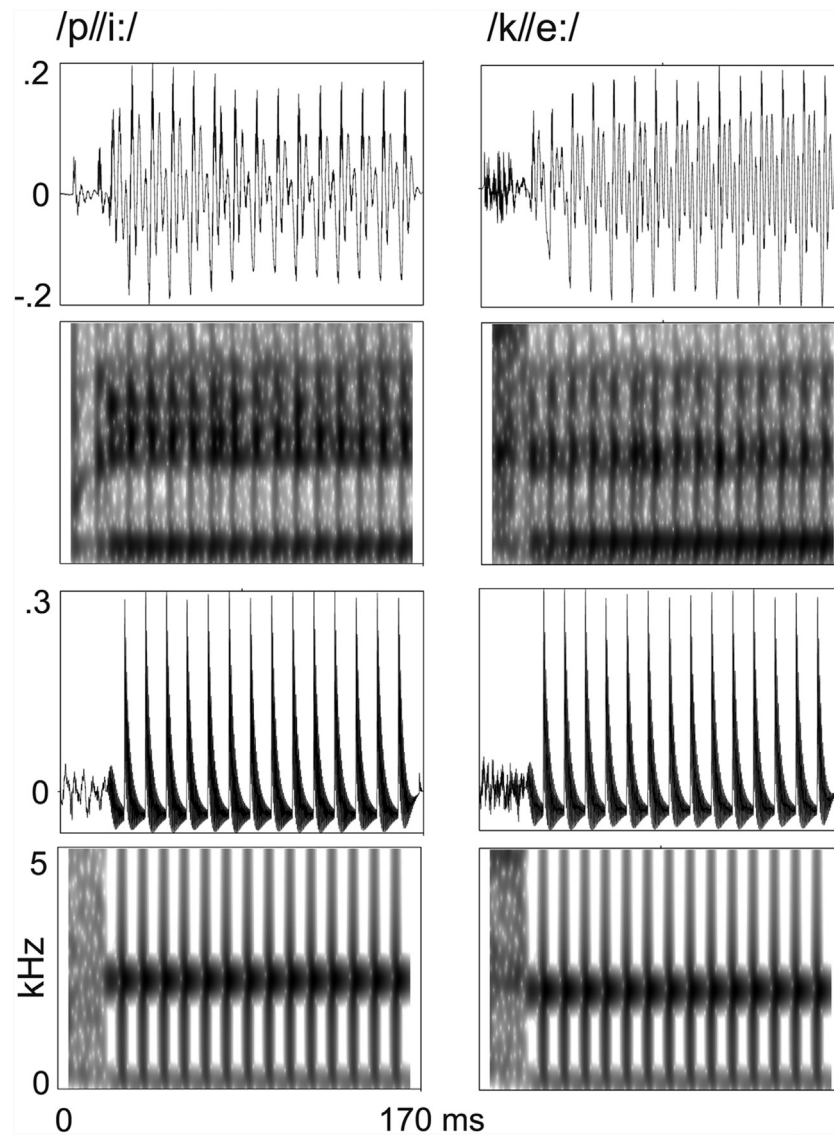
Originally, 94 typically developed monolingual Finnish-speaking children participated in a longitudinal study of preschool language abilities and later reading performance. The current study consists of the preschool data of 63 children (33 boys; 3 left-handed, 1 ambidextrous) that remained after the exclusion of the data of 31 children due to cancellation of participation (N = 12), a PIQ lower than the set limit of 85 (N = 1), later discovery of unclear family history of neurological problems (N = 1), excessive alpha band activity (N = 11) or motor artifacts (N = 8). The mean age of the children was 6 years 6 months (range 6 years 0 months–7 years 0 months), and they had an average of 80 (range 8–156) days of preschool teaching prior to the EEG experiment. All children were born full-term and had reportedly normal hearing. Most parents of the children had completed high school (fathers 73%, mothers 86%), and had college or university education (fathers 59%, mothers 71%), and were employed (fathers 90%, mothers 76%). The family background of these children is typical to the Finnish metropolitan capital area (Official Statistics of Finland (OSF), 2013).

The study was approved by the Ethical Board of Helsinki and Uusimaa Hospital District. Written consent was obtained from parents and oral consent from the child.

### 2.2. Stimuli and paradigm

The semi-synthetic CV syllables and their acoustically matched nonspeech counterparts used as stimuli were created using the Semisynthetic Speech Generation (Alku et al., 1999) method. Vowels /i/ and /e/ were compiled by extracting a glottal excitation from a natural speech signal, obtaining the desired formant structure with a digital all-pole filter, and adding the filtering effect of the vocal tract to the model. The F0 was 101 Hz for both vowels. For /i/, the lowest four formant frequencies were 410, 2045, 2260, and 3320 Hz and for /e/, 320, 2240, 2690, and 3275 Hz. The unvoiced plosives /k/ and /p/ were extracted from syllable /ke:/ and the short word /pito/, and inserted to the beginning of the semi-synthetic vowels to create standard stimulus syllables /pi/ and /ke/. The total duration of the standard stimulus was 170 ms (12 ms consonant and 158 ms vowel sections, including 5 ms rise and fall times), and its intensity set to approximately 55 dB SPL.

The nonspeech sounds were created by mimicking the glottal flow of the semi-synthetic syllables with a carefully controlled impulse train, so that the F0 was equal with the speech stimuli. Linear predictive coding (LPC; Rabiner and Schafer, 1978) of a prediction order of 10 was used to match the spectral envelope to that of the speech sound. Then the impulse train was used as an excitation to an all-pole filter which modeled only the second formant, i.e., the all-pole vocal tract consisted of a single resonance at 2240 Hz and 2045 Hz for the nonspeech counterpart of the /i/ and /e/, respectively. LPC coding of a prediction order of 50 was used to model /p/ and /k/, exciting it with random noise. The nonspeech syllables were formed by combining the corresponding nonspeech consonant and vowel counterparts (see Fig. 1). The speech and nonspeech stimuli were thus matched in terms of duration, F0, intensity



**Fig. 1.** Sound waveforms and spectrograms of the used stimuli. The semisynthetic CV syllables/pi:/and/ke:/are on the upper, and their respective nonspeech counterparts on the lower half of the figure.

and spectral envelope behavior, with only one of them recognized as speech, and the other described as “the sound of a buzzer or a horn” (Kuuluvainen et al., 2014).

The stimuli were presented in a multi-feature MMN paradigm with five deviant stimuli (consonant/counterpart, vowel/counterpart, sound F0, duration, and intensity) derived from the standard sounds, all syllables having no meaning in Finnish. The results for deviant stimulus responses are reported in another study (Kuuluvainen et al., 2015). In addition, 46 novel 200 ms complex sounds (telephone ringing, etc.; see Sorokin et al., 2010 for further details) were included to study attention, results being reported elsewhere.

The standard sounds/pi:/, /ke:/, nonspeech ‘pi’, and ‘ke’ were presented in separate, pseudo-randomised blocks with their respective deviants, resulting in a total of 538 stimuli (10 standards excluded from the analysis, followed by 528 stimuli, of which 50% were standards, 8.33% of each deviant, 8.33% of novel sounds). Standards after novel sounds were excluded from the analysis. Each block was presented twice, resulting in eight four-minute blocks in total. The order of the blocks was counterbalanced between participants.

### 2.3. EEG data recording and experiment procedure

EEG was recorded with 28 active electrodes (BrainVision QuickAmp & ActiCap, Brain Products, Germany), with four active electrodes placed on left and right mastoids, near the outer canthus of the right eye (horizontal electro-oculogram, HEOG) and the nose. Two passive bipolar electrodes were used to record the vertical electro-oculogram (VEOG). Recording reference was the average signal of all electrodes, sampling rate 250 Hz, and frequency band from DC to 50 Hz.

During the recording, the children sat in a comfortable chair in an electrically shielded and sound-attenuated room. Stimuli were delivered to both ears via Sony Professional MDR-7506 headphones, while they watched a self-selected, silent video. The children were instructed to watch the program carefully for questions of its content after each block, and not to pay attention to the stimuli. The overall quality of the children’s reports was good, indicating they had attended to the movie. After every two blocks, the child could choose a small item (toys, pens, stickers) to keep up motivation. During a longer mid-experiment break, the child was given juice and cookies, and after finishing, could choose

**Table 1**  
Children's neurocognitive test results.

	Mean (SD)	Min-Max
Age (years)	6.5 (0.2)	6.0–7.0
Total testing time (min)	138 (33)	100–360
Reasoning skills		
PIQ	103 (13)	85–139
VIQ	107 (10)	81–128
General language skills		
BNT (correct)	39 (5.0)	25–48
Comprehension of instructions (SP <sup>a</sup> )	10.6 (2.6)	4–17
Working memory		
Sentences (SP)	11.3 (2.2)	5–15
Numbers (SP)	10.8 (2.3)	5–15
Phonological skills		
Pseudowords (SP)	9.2 (1.8)	4–13
Phonological processing (SP)	10.0 (2.3)	6–16
Rapid Alternating Naming		
Colours, time (s)	63 (14)	43–113
Numbers, time (s)	64 (19)	33–141
Letters, time (s)	60 (17)	31–98
Objects, time (s)	75 (14)	42–114
Numbers & Letters, time (s)	57 (16)	36–89
Colours, errors	1.9 (1.9)	0–8
Numbers, errors	3.1 (4.3)	0–27
Letters, errors	3.8 (4.1)	0–18
Objects, errors	3.1 (2.9)	0–13
Numbers & Letters, errors	4.1 (5.3)	0–28
Reading skills <sup>b</sup>		
Letter naming (correct)	24 (6.2)	3–29
Syllable reading (correct)	7.5 (8.1)	0–18
Word reading (correct)	12.8 (18.2)	0–74

SD = Standard deviation, SP = Standard points, BNT = Boston naming test.

<sup>a</sup> SPs are based on age norms, where the normative age mean is 10 and standard deviation 3 points.

<sup>b</sup> The maximum score for letter naming is 29, for syllable reading 18 and for word reading 90.

between movie tickets and a monetary compensation. The experiment, including preparation and breaks, lasted approximately 2 h.

#### 2.4. Neurocognitive tests

The neurocognitive testing session was carried out on average 18 days (range 1–64) after the EEG. We obtained *performance and verbal intelligence quotients (PIQ and VIQ)* with the Finnish version of the Wechsler Preschool and Primary Scale of Intelligence–III (WPPSI-III; (Wechsler, 2009)). If the child's preschool PIQ was under 85 or VIQ under 75, they were reassessed a year later, using the Wechsler Intelligence Scale for Children–IV (WISC-IV; Wechsler, 2010) This was done to minimize shyness or lack of motivation as the reason for a below-threshold PIQ/VIQ score in preschool. After reassessment, all but one child (data discarded) had PIQ and VIQ within limits (see Table 1).

In addition to reasoning abilities, we assessed the children's *general language abilities* with the Boston Naming Test (BNT; Kaplan et al., 1997), and the Comprehension of instructions subtest from A Developmental Neuropsychological Assessment–II (NEPSY-II; Korkman et al., 2008) which tests the child's ability to point pictures following auditory instructions. In the BNT, the child names pictures, and is given a cue, e.g. the common use of the object, only if s(he) clearly misinterprets the picture. The score is the number of correctly named pictures. *Verbal short term and working memory* were assessed with the The Digit span subtest from WISC–III (Wechsler, 1999) and the Sentence repetition subtest from NEPSY-II (Korkman et al., 2008) In the Digit span subtest the child is asked to repeat numbers first in the same, and then in reverse order. Each correctly repeated sequence scores one raw point. In the Sentence

repetition subtest the child hears more and more complex sentences and is to repeat them word-for-word. Correctly repeated sentences score two raw points, and those with one or two mistakes one point.

*Phonological skills* were assessed using the Phonological processing subtest from NEPSY-II and the Repetition of nonsense words subtest from NEPSY I (Korkman et al., 1997). In the Phonological processing subtest the child hears three words with matching pictures and is asked to point the picture for a word containing the separately given target syllable. Later in the test the child is asked to first repeat a word and then either omit or change one or more phonemes to form a new word. Each correct answer scores one point. In the Repetition of nonsense words subtest the child hears nonwords, and is asked to repeat them as accurately as possible, each correctly repeated nonword scoring one point. The children's *naming speed* was tested with the Rapid Automatized Naming test (RAN; Ahonen et al., 2003), in which the child is asked to name items as fast and accurately as (s)he can. We used the first five tables (colours; numbers; letters; objects; numbers & letters; colours, numbers & letters) which contain five rows of ten items each. The child's time as well as errors made during naming are recorded. Finally, we assessed *prereading skills* with letter naming (the Finnish alphabet in random order) and syllable reading (18 non-word syllables au, is sa, eu, ki, tu, oi, pe, as, iu, ke, pi, isu, upe, ako, eso, uti, oke, given to the child one per time in a small card to be read out loud). Children mastering at least half of the syllables proceeded to the Finnish LukiLasse (Häyrynen et al., 1999) word reading task (measuring correctly read words in 120 s). For standardized tests (WPPSI-III, WISC-IV, Nepsy I and II) raw scores were converted to age-appropriate standardized scores (mean score in the normative age group being 10, and standard deviation 3 points).

We used two different testing orders, balancing auditory-only and visually guided subtests. Orders were assigned balancing for gender, and no differences emerged between scores from different orders ( $t(58-61) = -0.05-1.8$ ;  $p > 0.05$ ). The neurocognitive testing session lasted 2.5 h on average, including three breaks, with similar rewards as the EEG session. Parents wishing so received feedback of the age-typicality of their child's performance over the phone, emphasizing that single-time test results of small children can be unreliable due to attention, motivation and fatigue effects, and that children's cognitive skills often develop in different paces. Parents could discuss the results also with one of the authors (S.K., a licensed psychologist). If the parents wished to have the child assessed further, they were guided to contact the Finnish health care system.

For all investigated skills where age-appropriate norms were available, the group average was close to the age-typical standardized mean score of 10 points (100 points for VIQ and PIQ; see Table 1). For the RAN and BNT, conversion to standard points was not possible, but the average naming times and numbers of errors were comparable to those reported in the normative data for six-year-olds (Ahonen et al., 2003; Kaplan et al., 1997). No individual score was set as exclusion limit apart from those of VIQ and PIQ, and the range of performance varied therefore from poor to excellent in different tasks.

#### 2.5. EEG data analysis

BESA 6.0 (Besa GmbH, Germany) software was used to preprocess the raw data individually for each block. After interpolating missing or bad channels, blink artifacts were removed using the BESA Principal Component Analysis (PCA) Tool, the thresholds set at 75  $\mu$ V for both VEOG and HEOG channels. If blink patterns were still observed, the "Define Artifact Topography" Tool was used, the blink definition criteria being a PC with bipolar topography across the



**Table 2**  
Test scores used in rANCOVA.

	Mean (SD)	MIN	25th	50th	75th	MAX
Phonological skills (%) <sup>a</sup>	58.5 (8.2)	41.1	52.4	58.0	63.3	75.2
Prereading skills (%) <sup>a</sup>	62.5 (30.1)	5.2	38.0	48.3	95.5	100
Rapid Alternating Naming (s) <sup>b</sup>	69.2 (12.3)	44.5	62	70	77	102
Verbal Short Term Memory (%) <sup>a</sup>	53.3 (6.6)	34.5	49.5	53.9	60.0	66.4
PIQ	103 (13)	85	91	102	111	139
VIQ	107 (10)	81	100	109	115	128

SD = Standard deviation; MIN = minimum; 25th, 50th and 75th percentiles; MAX = maximum.

<sup>a</sup> Average percent correct across two subtests.

<sup>b</sup> Average naming speed in seconds across two subtests.

horizontal axis of the eyes, explaining at least 85% of the variance in the selected data.

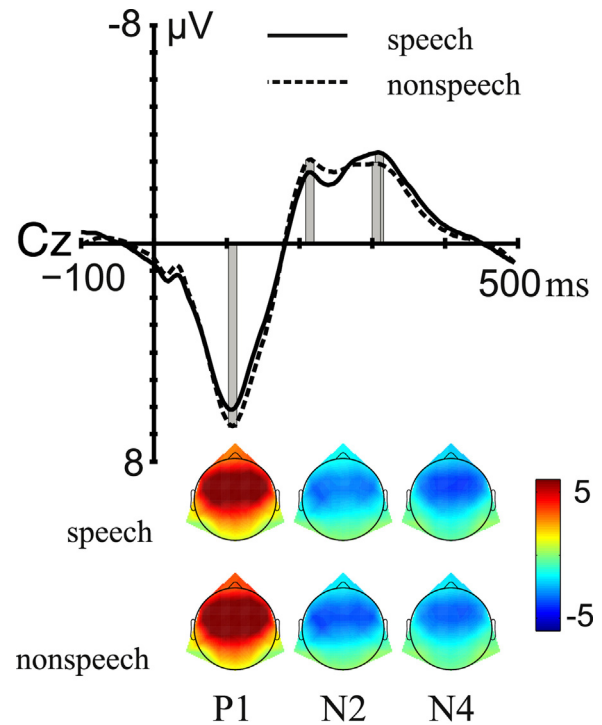
The data were then re-referenced to mastoids (see Supplement A for nose-referenced data) and epoched (–100–500 ms), filtered (1–30 Hz), and corrected for baseline (–100–0 ms). The data were averaged rejecting epochs with (a) peak-to-peak amplitude (s) greater than 75  $\mu\text{V}$ . These blockwise averaged ERPs were then exported as ASCII multiplexed files to EEGLAB (Delorme and Makeig, 2004), where averages over blocks were formed. The average number of accepted trials was 667 (range 437–786) for the speech and 663 (range 482–767) for the nonspeech standard stimuli. P1 was defined as the positive, and N2 and N4 as the negative maxima, at 50–150 ms, 150–300 ms, and 300–450 ms, respectively, from the Cz electrode. Amplitudes were quantified by calculating the individual mean response amplitude in a  $\pm 5$  ms time window centered at the peak of the group grand-mean response.

## 2.6. Statistical analyses

The statistical significance of the ERPs was determined at the Cz electrode, using a *t*-test comparing the amplitude values to zero. For each of the three speech–nonspeech response pairs, we conducted a three-way rANOVA of condition (speech versus nonspeech), anterior-posterior, and lateral distribution, using nine electrodes (F3, Fz, F4, C3, Cz, C4, P3, Pz, and P4).

Averaged scores across tests and electrodes were used to investigate the relationship between ERPs and neurocognitive tests. For the ERPs, the mean amplitudes of three electrodes in left (F3, C3, P3) and three in right (F4, C4, P4) scalp locations were calculated. PIQ and VIQ scores were used unaltered. We used percentage correct raw scores from phonological processing and repetition of nonsense word for phonological skills, from digit span forward and sentence repetition for vSTM, and from letter naming and syllable reading for prereading skills. For the RAN score, naming speeds (in s) of colours and objects were averaged, since reliable data for these two was available also for those children having problems with numbers or letters (see Table 2 for details of all scores).

For each of the six scores, a two-way repeated measures analysis of covariance (rANCOVA) of condition (speech versus nonspeech) and laterality (left versus right) was conducted for P1, N2, and N4 separately, using the test score as covariate. To interpret the results, Pearson correlation coefficients were calculated between all test scores and between each test score and: (1) ERP measure; (2) amplitude differences between left and right, and speech and nonspeech conditions, always subtracting the amplitude of right or nonspeech from left or speech ERP amplitude, respectively; (3) the difference of the left and right speech–nonspeech amplitude difference. This last measure was at its greatest when the amplitude of the left speech ERP was larger than the corresponding nonspeech ERP, with an opposite pattern on the right. The amplitudes of the negative ERPs (N2 and N4) were reversed to keep the correlation directions constant across all three ERPs.



**Fig. 2.** ERPs to standard stimuli. Grey areas highlight the time windows of amplitude quantification for each response. Respective topographic maps are presented below the waveforms.

To correct for observed violations of sphericity, we applied Greenhouse-Geisser corrections (original degrees of freedom reported with  $\epsilon$  and corrected significance). Statistically significant interactions were further examined with Bonferroni-corrected post hoc tests. All analyses were conducted with SPSS Statistics 22 (IBM, U.S.A.).

## 3. Results

### 3.1. ERP amplitudes

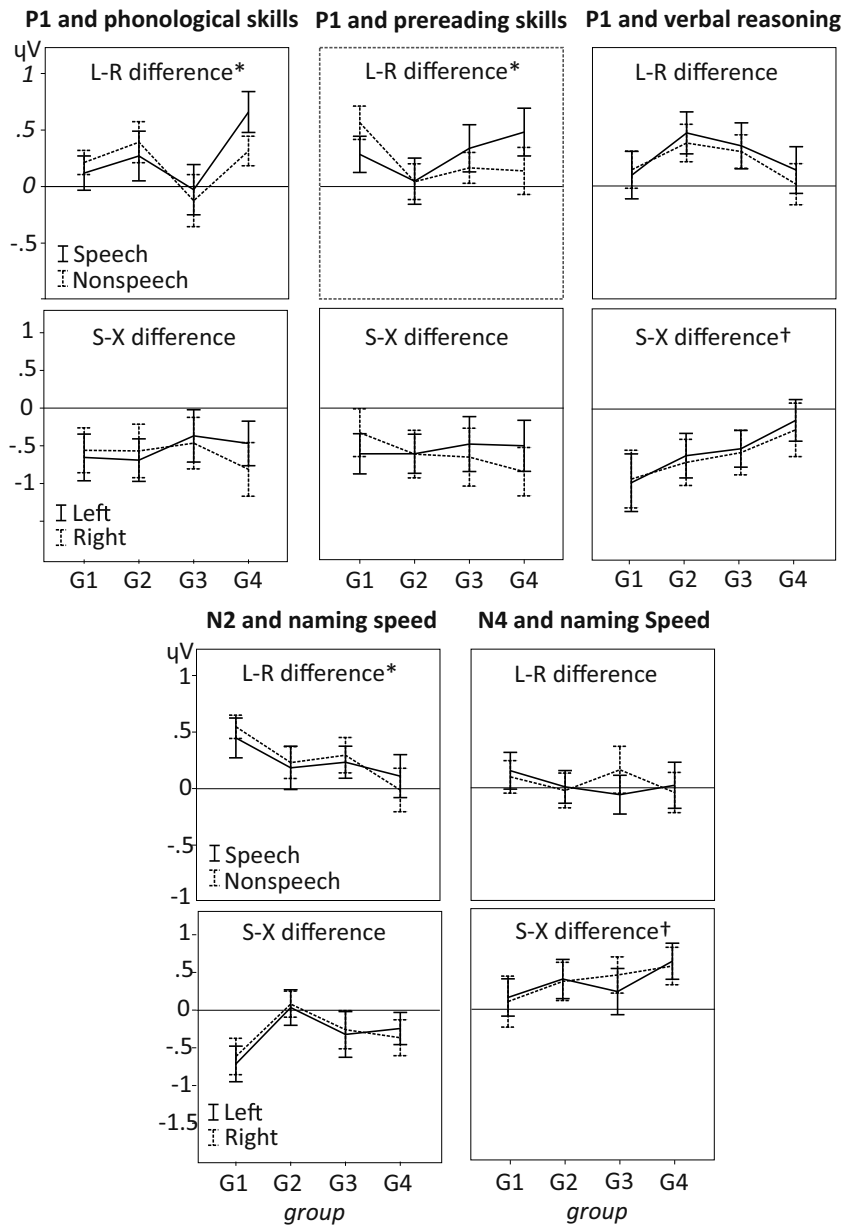
All standard stimulus responses significantly differed from zero ( $t(62) = -15.5$ – $19.9$ ,  $p < 0.001$ ); see Table 3). There was a main effect of condition for all three responses (P1:  $F(1,62) = 13.6$ ,  $p < 0.001$ , partial  $\eta^2 = 0.18$ , observed power  $P_o = 0.95$ ; N2:  $F(1,62) = 9.1$ ,  $p = 0.004$ , partial  $\eta^2 = 0.13$ ,  $P_o = 0.84$ ; N4:  $F(1,62) = 7.8$ ,  $p = 0.007$ , partial  $\eta^2 = 0.11$ ,  $P_o = 0.79$ ). For P1 and N2 this resulted from larger responses to nonspeech than speech stimuli, whereas for N4 the effect was the opposite (see Fig. 2). Further, there was a condition  $\times$  anterior-posterior  $\times$  laterality interaction for the P1 response ( $F(4,248) = 2.9$ ,  $\epsilon = 0.78$ ,  $p = 0.036$ , partial  $\eta^2 = 0.044$ ,  $P_o = 0.70$ ), which resulted from larger responses for nonspeech than speech stimuli in the middle and right than left and anterior electrode sites. Also, there was a statistically significant condition  $\times$  anterior-posterior interaction for the N2 response ( $F(2,124) = 11.6$ ,  $\epsilon = 0.70$ ,  $p < 0.001$ , partial  $\eta^2 = 0.16$ ,  $P_o = 0.99$ ), which resulted from frontocentrally larger responses for nonspeech than speech stimuli (see Fig. 2).

### 3.2. Covariance analysis for test scores and ERPs

Better performance in phonological and prereading skills were associated with larger P1 amplitudes at the left than the right scalp locations, especially for the speech stimuli (condition  $\times$  laterality  $\times$  phonological skills interaction:

**Table 3**  
ERP latencies, amplitudes and statistical significance.

Speech	Speech				Nonspeech			
	Latency (Cz; ms)	Amplitude <sup>a</sup> (Cz; $\mu$ V)	t(62)	Sign. <sup>b</sup> (p)	Latency (Cz; ms)	Amplitude (Cz; $\mu$ V)	t(62)	Sign. (p)
P1	108	6.0 (2.4)	19.9	<0.000	108	6.6 (2.7)	19.6	<0.000
N2	216	-2.5 (1.8)	-11.4	<0.000	216	-3.0 (1.7)	-13.4	<0.000
N4	308	-3.3 (1.7)	-15.5	<0.000	304	-2.9 (1.9)	-12.2	<0.000

<sup>a</sup> Mean, standard deviation in brackets.<sup>b</sup> Statistical significance of the response compared to zero.

**Fig. 3.** The relationship between test performance and cortical responses. For illustrative purposes, children were divided into quartile groups (G1-G4) according to test performance, so that children in G1 were the lowest 25% and children in G4 the highest 25% of scorers for the skill in question (see Table 2 for details; note that groups are reversed for naming speed, as good performers are faster). The upper images show amplitude difference between left and right scalp locations, positive values mean larger responses on the left than right. The lower images show amplitude differences between speech and nonspeech responses, positive values mean larger responses to speech than nonspeech stimuli. Differences were calculated for reversed values for negative responses. Error bars represent  $\pm 1$  standard error of the mean. L = left, R = right, S = speech, X = nonspeech.

$F(1,61)=6.2$ ,  $p=0.016$ , partial  $\eta^2=0.092$ ,  $P_o=0.69$ , and condition  $\times$  laterality  $\times$  prereading skills interaction:  $F(1,61)=5.8$ ,  $p=0.019$ , partial  $\eta^2=0.087$ ,  $P_o=0.66$ ). Furthermore, there was a marginally significant condition  $\times$  VIQ interaction ( $F(1,61)=3.0$ ,

$p=0.086$ , partial  $\eta^2=0.048$ ,  $P_o=0.40$ ), which resulted from an increased difference in nonspeech versus speech P1 amplitudes associating with poorer VIQ scores (see Fig. 3). Furthermore, slower performance in RAN was associated with larger N2 amplitudes in

the left than right scalp locations (laterality  $\times$  RAN speed interaction:  $F(1,61)=7.0$ ,  $p=0.010$ , partial  $\eta^2=0.10$ ,  $P_o=0.72$ ). Finally, there was a marginally significant overall effect of RAN speed and N4 amplitude ( $F(1,61)=3.8$ ,  $p=0.057$ , partial  $\eta^2=0.058$ ,  $P_o=0.48$ ), faster naming performance being associated with larger speech than nonspeech N4 amplitudes (Fig. 3). There were no statistically significant results for the rANCOVAs with PIQ or verbal short term memory as a covariate.

From the neurocognitive skills associated with ERP measures, significant intercorrelations were found between verbal reasoning, phonological, and prereading skills, whereas none of these three correlated with RAN speed ( $r=0.340$ – $0.640$ , see Table 4). The correlations between neurocognitive skills and ERP measures were in line with the results of the rANCOVA: Phonological and prereading skills correlated with P1 left-speech right-nonspeech bias ( $r=0.304$  and  $0.205$ ,  $p=0.016$  and  $0.019$ , respectively), and verbal reasoning was marginally significantly correlated with P1 left speech-nonspeech difference ( $r=0.248$ ,  $p=0.050$ ). RAN speed correlated significantly with nonspeech N2 left-right amplitude difference ( $r=0.375$ ,  $p=0.002$ ) and inversely with the amplitude of right speech N4 ( $r=-0.290$ ,  $p=0.021$ ) as well as marginally with left speech N4 ( $r=-0.229$ ,  $p=0.071$ ; see Table 5).

#### 4. Discussion

We investigated cortical pre-attentive processing of speech and nonspeech sounds in typically developed six-year-old children, using semi-synthetic syllables and their acoustically matched nonspeech counterparts. Our results suggest that already at preschool age, the neural substrates responsible for sound detection and content analysis are distinct for speech and equally complex nonspeech sounds. We found that P1 and N2 were smaller for speech than nonspeech sounds, whereas the N4 showed an opposite pattern, indicating that sound “speechness” quality, and/or sound familiarity, is detected and processed at very early stages of auditory analysis. Furthermore, individual differences between responses to speech and nonspeech stimuli were associated with several aspects of verbal cognitive functioning.

##### 4.1. The relationship of P1, phonological and verbal reasoning skills

The “speechness” quality of a sound, or alternatively sound familiarity, was cortically detected already after the first 100 ms of sound processing, as reflected by the smaller P1 amplitude to syllables than nonspeech counterparts. This is consistent with previous studies in school-aged children, showing smaller P1s to syllables than complex harmonic sounds, which was explained by the nonspeech stimuli being perceptually more salient and thus more easily detectable (Čeponienė et al., 2005, 2008). However, since we carefully matched the speech and nonspeech stimuli, differences in stimulus complexity or acoustic properties, including stimulus energy, cannot explain the observed differences. However, the child P1 has been suggested to be merged with the equivalent of adult P2 (Čeponienė et al., 2005, 2008), which in turn has been linked to memory trace formation for unfamiliar contrasts (Tremblay et al., 2001). It is thus possible that the larger P1 responses to nonspeech counterparts than speech sounds reflect the build-up or maintenance of a memory trace for these equally complex but unfamiliar sounds. This is also supported by the findings of Bruder et al. (2011), who showed that more accurate vowel prototype memory traces, established prior to the experiment, associated with smaller P1s.

Differences between P1s to speech and nonspeech sounds appeared at the right and midline, but not at left and anterior sites, suggesting a right-hemisphere bias in the processing of nonspeech

compared to speech sounds. This is consistent with previous studies of speech and nonspeech processing in adults, indicating a rightward bias for nonspeech processing (for reviews, see McGettigan and Scott, 2012; Tervaniemi and Hugdahl, 2003). However, P1 lateralization effects were not observed in the previous studies of speech and nonspeech processing in school-aged children (Bruder et al., 2011; Čeponienė et al., 2001, 2005, 2008). One possible explanation is that the smaller group sizes of these studies ( $N=11$ – $14$ ) did not allow for lateralization effects to emerge.

In the current study, larger P1s to speech sounds on the left than right, and larger P1s to nonspeech sounds in the right than left scalp locations, were related to better phonological and prereading skills. The children performing poorest in these tasks showed the opposite pattern, with larger nonspeech P1s at left than right scalp locations. Note, that the similarity of the associations of phonological and prereading skills with P1 is probably explained by the relatively strong inter-correlation of these measures. Furthermore, larger P1 differences between nonspeech and speech sounds, especially left to midline, were associated with poorer verbal reasoning skills. In contrast, children performing best had equal-sized P1s to both stimuli. Hence, we propose that P1 reflects processes related to successful establishment of long-term phonological representations especially in the left hemisphere, which in turn is associated with better phonological skills and with more mature verbal functioning. In contrast, if the processes behind P1 elicitation are biased towards the right hemisphere for speech, or are engaged in extensive processing of nonspeech sounds in the left hemisphere, verbal and especially phonological skills fall below age expectations.

The physiological explanation behind this might be in neuronal over-representation of native language speech sounds, resulting in erroneous assignment of relevance. This would lead to neural engagement in the left cortical areas to the nonspeech counterparts, which have a speech-like acoustic structure but are novel to the children. This interpretation is also consistent with studies showing smaller P1s to associate with increased maturation of the cortex. P1 amplitude elicited to pure tones showed steady decrease from age seven onwards (Bishop et al., 2011; Sussman et al., 2008), and smaller P1s to vowel prototypes reflected more precision in neural tuning curves (Bruder et al., 2011). These possibilities could be further investigated by using also non-native speech sounds as in Bruder et al. (2011), or with a learning paradigm, in which children would be taught to recognize and discriminate the nonspeech sounds as communicative.

##### 4.2. The relationship of N2 and N4 with speed of lexical access

Unlike 8–10 year-old children (Čeponienė et al., 2005, 2008), the preschoolers of the current study had larger N2 responses to the nonspeech analogues than to the CV syllables. Our result is, however, similar to that obtained by Čeponienė et al. (2001), who reported larger N2s to complex tones than vowels in schoolchildren. They then speculated that N2 reflects memory-trace formation, since Karhu et al. (1997), had previously shown that N2 amplitude increases with stimulus repetition. In this case the larger nonspeech than speech N2 amplitudes in our study would be explained by a process similar to that explaining the P1 results.

A second possible explanation to the different results between preschoolers and schoolchildren is the maturation of speech- and nonspeech-elicited N2. Earlier longitudinal studies suggest that N2 elicited by vowels is relatively stable in amplitude between ages five and eight years (Shafer et al., 2015), whereas N2 elicited by harmonic tones decreases in amplitude from age four to nine years (Čeponienė et al., 2002). It is thus possible that during the transition to school-age, the cognitive processes underlying N2 mature,

**Table 4**  
Intercorrelations of neurocognitive measures.

Neurocognitive measure	PIQ	VIQ	PHON	PRER	VSTM	RAN
Performance reasoning (PIQ)	1					
Verbal reasoning (VIQ)	.421 <sup>b</sup>	1				
Phonological skills (PHON)	0.011	.340 <sup>b</sup>	1			
Prereading skills (PRER)	0.10	.349 <sup>b</sup>	.640 <sup>b</sup>	1		
Verbal short term memory (VSTM)	.272 <sup>a</sup>	.447 <sup>b</sup>	.426 <sup>b</sup>	.342 <sup>b</sup>	1	
Naming speed (RAN)	0.073	−0.15	−0.072	−0.21	−0.19	1

<sup>a</sup> Correlation significant at  $p < 0.05$  level (2-tailed).

<sup>b</sup> Correlation significant at  $p < 0.01$  level (2-tailed).

**Table 5**  
Correlations between neurocognitive skills and ERP measures.

Neurocognitive skill	ERP measure	$r^a$	Sig. <sup>b</sup>
Verbal reasoning	P1 Left Speech-Nonspeech Difference <sup>c</sup>	0.248	0.050
Phonological processing	P1 Left-Speech Right-Nonspeech Bias <sup>d</sup>	0.304	0.016
Prereading skills	P1 Left-Speech Right-Nonspeech Bias	0.295	0.019
Naming speed	N2 Nonspeech Left-Right Difference <sup>e</sup>	0.375	0.002
Naming speed	N4 Speech Left <sup>f</sup>	−0.229	0.071
Naming speed	N4 Speech Right	−0.290	0.021

Note that the amplitudes of N2 and N4 were reversed to keep correlation directions consistent.

<sup>a</sup> Pearson's correlation coefficient; note that negative ERPs were reversed for analysis.

<sup>b</sup> Two-tailed significance.

<sup>c</sup> Amplitude difference of left speech and nonspeech P1s.

<sup>d</sup> Difference of left and right amplitude differences of speech and nonspeech P1s (maximal when speech P1 is largest at left and smallest at right, and nonspeech P1 vice versa).

<sup>e</sup> Difference of left and right nonspeech N2s.

<sup>f</sup> The amplitude of left speech N4.

changing the response pattern to favor syllables over nonspeech analogues.

A third explanation of the different N2 results in our study and the previous studies of Čeponienė et al. (2005, 2008) are the stimuli used. Compared to Čeponienė et al. (2005, 2008), who used CV syllables and nonspeech counterparts with only 60–80 ms vowel/counterpart sections, the vowels of our CV syllables/counterparts were relatively long (158 ms). It is possible that the longer vowel length of 158 ms in the current study led to smaller speech than nonspeech N2s. This interpretation is supported by the results of Čeponienė et al. (2001), who reported larger N2s to 260 ms long complex tones than vowels in schoolchildren. Furthermore, the studies of Shafer et al. (2005) and Datta et al. (2010) in 8–10 year-olds used short (50 ms) and long (250 ms) versions of the same vowels, and shared six of their nine control children. Comparing across these two studies, N2 to the short vowel/i/ was approximately twice the size as to the long version (−4 versus −2  $\mu$ V), whereas P1 sizes were more or less equal (−5  $\mu$ V) and N4 was elicited only to the long but not the short vowel. As, to our knowledge, there are no longitudinal within-participant studies of speech versus acoustically matched nonspeech processing covering both preschool and school-aged children, nor studies comparing vowels of different lengths directly, these questions have to be addressed by further research.

Interestingly, larger N2s in the left than the right scalp areas, especially for nonspeech sounds, were associated with slower naming speed, which is a strong predictor of later reading ability (for a review, see Kirby et al., 2010; see also Dandache et al., 2014; Torppa et al., 2010). This effect emerged for both speech and nonspeech sounds, suggesting that it reflects a domain-general rather than speech-specific neural process. The result is consistent with previous research by Hämäläinen et al. (2013) reporting larger N2 amplitudes in preschool children at risk for dyslexia than controls using short (300 ms), intermediate (360 ms) and long (460 ms) versions of the disyllabic pseudoword/a:ta/, varying the length of the gap between the two syllables. ERPs were also recorded to respective nonspeech sounds, where the/a/vowels were replaced

with harmonic tones comprised of seven sinusoidals, consonant portion staying the same. The difference between at-risk and control children appeared at left scalp locations for the short 300 ms speech stimulus and for all nonspeech stimuli, and at the right scalp locations only for the intermediate-length nonspeech stimulus. In Hämäläinen et al. (2013), larger speech and nonspeech N2 amplitudes at preschool were associated with slower and less accurate reading at school in both groups, similarly to those reported by Espy et al. (2004). The results of our study, showing larger N2s to associate with slower naming speed, suggest that these links between preschool N2 amplitude and later reading performance might be mediated via lexical access. They imply that smaller N2s reflect a more mature cognitive process, or alternatively, larger responses reflect a maturational delay.

Similarly as in the school-aged children of Čeponienė et al. (2001, 2005, 2008), preschooler's N4 amplitudes in our study were larger for speech than nonspeech sounds. Furthermore, larger N4s to speech than nonspeech sounds were associated with faster RAN performance, the amplitudes of N4s to speech sounds correlating inversely with naming speed, so that children with shorter naming times had larger N4s to speech sounds. These results are consistent with the earlier interpretations of the processes behind the child N4 response. It seems to be associated with further processing of sounds identified as speech, and is likely to reflect automatic scan for access or success of retrieval of lexical representations linked to the perceived sounds.

#### 4.3. Study limitations and future directions

The current study holds a few limitations. First, the direction of attention was controlled only by giving the children a task of attending to the videos and answering questions, and there was no control condition in which the children would have been asked to attend to the stimuli. Consequently, the effects of possible covert attention to the stimulus stream cannot be excluded as an explanation of the results. Shafer et al. (2007) studied 8–10-year-old children with SLI and their controls, and found that directing atten-



tion to speech stimuli reduced P1 amplitude, presumably via the influence of elicited processing negativity Nd. Furthermore, when the children were watching a silent video and asked not to attend to stimuli, P1 amplitudes were smaller for typically developing children than those with SLI, suggesting the former still allocated some attention to the speech stimuli even when requested not to. In the current study, smaller P1s to speech than nonspeech stimuli were seen at the children having lowest verbal reasoning scores, whereas best performers had equal-sized P1s to both conditions. If the result was to be interpreted as in Shafer et al. (2007), it would imply that the verbally less mature children would allocate more attention to speech than their more mature peers, which is in contrast to the original finding. The role of attention, nevertheless, remains unresolved, and should be controlled in future studies.

Secondly, the stimuli included only semi-synthetic speech and their nonspeech counterparts, based on Finnish language. The underlying functions should thus be investigated further using both natural and semi-synthetic speech sounds from different languages, and expanding to other auditory domains such as music. This would further illuminate the speech and language specificity of the current findings. Furthermore, learning paradigms could reveal if memory trace formation is indeed reflected in the obligatory responses. Longitudinal studies looking at different age groups could show if the functional significance of the responses changes in typical development, and the inclusion of clinical groups would show if the associations also hold for atypical speech processing. Finally, methods allowing for source localization, such as magnetoencephalography (MEG) or high-density EEG with anatomical images, could elucidate more precisely the neural substrates involved in the generation of the ERPs.

## 5. Conclusion

Based on the current results, we propose that syllables and corresponding nonspeech sounds are processed by at least partially different neural substrates in preschoolers' auditory cortices. P1 and N2 responses were enhanced for nonspeech compared to speech sounds, suggesting that the early stages of neural detection at this age are more responsive to unfamiliar than familiar complex auditory material. Enhancement of speech compared to nonspeech sounds was observed only for N4. The investigated ERPs seem to reflect different cognitive processes: the P1 appears to be associated with phonological and prereading skills as well as with overall verbal functioning, whereas N2 and N4s seem to be associated with the speed of access to lexical representations. Overall, our results suggest that speech and nonspeech-elicited ERPs can be useful measures of children's cortical functioning. Since these responses can be recorded without any active role of the participant, they could illuminate speech vs. nonspeech functions and development even in smaller children.

## Conflict of interest

The authors declare no conflicts of interest.

## Acknowledgments

We thank the children and families for participating, and Pirita Uuskoski, Piia Turunen, Lilli Kimppa, Saira Seppänen, Henna Markkanen, and Roope Heikkilä for their help in data collection, as well as Prof. Paavo Alku for his comments on the stimuli. This work was supported by the Academy of Finland (grants number 128840 and 1276414), the Finnish Cultural Foundation, Ella and Georg Ehrnrooth Foundation, Emil Aaltonen Foundation, Kone Foundation, Lundbeck Foundation, and Jane and Aatos Erkko Foundation.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.dcn.2016.04.001>.

## References

- Ahonen, T., Tuovinen, S., Leppäsaari, T., 2003. Nopeat Sarjallisen Nimeämisen Testi. Niilo Mäki Instituutti & Haukkarannan koulu, Jyväskylä (Rapid Alternating Naming Test).
- Alku, P., Tiitinen, H., Näätänen, R., 1999. A method for generating natural-sounding speech stimuli for cognitive brain research. *Clin. Neurophysiol.* 110 (8), 1329–1333.
- Bishop, D.V.M., Anderson, M., Reid, C., Fox, A.M., 2011. Auditory development between 7 and 11 years: an event-related potential (ERP) study. *PLoS One* 6 (5), e18993. <http://dx.doi.org/10.1371/journal.pone.0018993>.
- Bruder, J., Leppänen, P.H.T., Bartling, J., Csépe, V., Démonet, J.-F., Schulte-Körne, G., 2011. An investigation of prototypical and atypical within-category vowels and non-speech analogues on cortical auditory evoked related potentials (AERPs) in 9 year old children. *Int. J. Psychophysiol.* 79 (2), 106–117.
- Čeponienė, R., Shestakova, A., Balan, P., Alku, P., Yaguchi, K., Naatanen, R., 2001. Children's auditory event-related potentials index sound complexity and speechness. *Int. J. Neurosci.* 109 (3), 245.
- Čeponienė, R., Rinne, T., Näätänen, R., 2002. Maturation of cortical sound processing as indexed by event-related potentials. *Clin. Neurophysiol.* 113 (6), 870–882.
- Čeponienė, R., Alku, P., Westerfield, M., Torki, M., Townsend, J., 2005. ERPs differentiate syllable and nonphonetic sound processing in children and adults. *Psychophysiology* 42 (4), 391–406.
- Čeponienė, R., Torki, M., Alku, P., Koyama, A., Townsend, J., 2008. Event-related potentials reflect spectral differences in speech and non-speech stimuli in children and adults. *Clin. Neurophysiol.* 119 (7), 1560–1577.
- Dandache, S., Wouters, J., Ghesquière, P., 2014. Development of reading and phonological skills of children at family risk for dyslexia: a longitudinal analysis from kindergarten to sixth grade. *Dyslexia* 20 (4), 305–329. <http://dx.doi.org/10.1002/dys.1482>.
- Datta, H., Shafer, V.L., Morr, M.L., Kurtzberg, D., Schwartz, R.G., 2010. Electrophysiological indices of discrimination of long-Duration, phonetically similar vowels in children with typical and atypical language development. *J. Speech Lang. Hear. Res.* 53 (3), 757–777.
- Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134 (1), 9–21. <http://dx.doi.org/10.1016/j.jneumeth.2003.10.009>.
- Espy, K.A., Molfese, D.L., Molfese, V.J., Modglin, A., 2004. Development of auditory event-related potentials in young children and relations to word-level reading abilities at age 8 years. *Ann. Dyslexia* 54 (1), 9–38.
- Hämäläinen, J.A., Guttorm, T.K., Richardson, U., Alku, P., Lyytinen, H., Leppänen, P.H.T., 2013. Auditory event-related potentials measured in kindergarten predict later reading problems at school age. *Dev. Neuropsychol.* 38 (8), 550–566.
- Häyrynen, T., Serenius-Sirve, S., Korkman, M., 1999. Lukilasse. Psykologien Kustannus Oy, Helsinki, Finland.
- Kaplan, E., Goodglass, H., Weintraub, S., Segal, O., Laine, M., Koivuselkä-Sallinen, P., Niemi, J., 1997. Bostonin Nimentäesti. Waverly, Inc./Psykologien Kustannus Oy., Helsinki (Boston Naming Test).
- Karhu, J., Herrgård, E., Pääkkönen, A., Luoma, L., Airaksinen, E., Partanen, J., 1997. Dual cerebral processing of elementary auditory input in children. *Neuroreport* 8 (6), 1327–1330.
- Kirby, J.R., Georgiou, G.K., Martinussen, R., Parrila, R., 2010. Naming speed and reading: from prediction to instruction. *Read. Res. Q.* 45 (3), 341–362. <http://dx.doi.org/10.1598/RRQ.45.3.4>.
- Korkman, M., Kirk, U., Kemp, S.L., 1997. NEPSY: Lasten Neuropsykologinen Tutkimus. Psykologien Kustannus Oy, Helsinki (NEPSY: A developmental neuropsychological assessment).
- Korkman, M., Kirk, U., Kemp, S.L., 2008. Nepsy II—Lasten Neuropsykologinen Tutkimus. Psykologien Kustannus Oy, Helsinki (NEPSY II—A developmental neuropsychological assessment).
- Kuhl, P.K., Conboy, B.T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., Nelson, T., 2008. Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philos. Trans. R. Soc. London B: Biol. Sci.* 363 (1493), 979–1000.
- Kuhl, P.K., 2004. Early language acquisition: cracking the speech code. *Nat. Rev. Neurosci.* 5 (11), 831–843.
- Kushnerenko, E., Čeponienė, R., Balan, P., Fellman, V., Huotilaine, M., Näätänen, R., 2002. Maturation of the auditory event-related potentials during the first year of life. *Neuroreport* 13 (1), 47–51.
- Kuuluvainen, S., Nevalainen, P., Sorokin, A., Mittag, M., Partanen, E., Putkinen, V., ... Kujala, T., 2014. The neural basis of sublexical speech and corresponding nonspeech processing: a combined EEG-MEG study. *Brain & Language* 130, 19–32.
- Kuuluvainen, S., Alku, P., Makkonen, T., Lipsanen, J., Kujala, T., 2015. Cortical speech and nonspeech discrimination in relation to cognitive measures in preschool children. *Eur. J. Neurosci.* <http://dx.doi.org/10.1111/ejn.13141>, n/a-n/a.

- Lovio, R., Nääätänen, R., Kujala, T., 2010. **Abnormal pattern of cortical speech feature discrimination in 6-year-old children at risk for dyslexia.** *Brain Res.* 1335, 53–62.
- McGettigan, C., Scott, S.K., 2012. Cortical asymmetries in speech perception: what's wrong, what's right, and what's left? *Trends Cogn. Sci.* 16 (5), 269–276, <http://dx.doi.org/10.1016/j.tics.2012.04.006>.
- (OSF) Official Statistics of Finland, 2013. **Educational structure of population.** Statistics Finland (e-publication).
- Pihko, E., Kujala, T., Mickos, A., Antell, H., Alku, P., Byring, R., Korkman, M., 2005. **Magnetic fields evoked by speech sounds in preschool children.** *Clin. Neurophysiol.* 116 (1), 112–119.
- Ponton, C.W., Eggermont, J.J., Kwong, B., Don, M., 2000. **Maturation of human central auditory system activity: evidence from multi-channel evoked potentials.** *Clin. Neurophysiol.* 111 (2), 220–236.
- Rabiner, L.R., Schafer, R.W., 1978. *Digital Processing of Speech Signals*, vol. 100. Prentice Hall, Englewood Cliffs.
- Shafer, V.L., Morr, M.L., Datta, H., Kurtzberg, D., Schwartz, R.G., 2005. **Neurophysiological indexes of speech processing deficits in children with specific language impairment.** *J. Cogn. Neurosci.* 17 (7), 1168–1180.
- Shafer, V.L., Ponton, C., Datta, H., Morr, M.L., Schwartz, R.G., 2007. **Neurophysiological indices of attention to speech in children with specific language impairment.** *Clin. Neurophysiol.: Off. J. Int. Fed. Clin. Neurophysiol.* 118 (6), 1230–1243, <http://dx.doi.org/10.1016/j.clinph.2007.02.023>.
- Shafer, V.L., Yu, Y.H., Wagner, M., 2015. **Maturation of cortical auditory evoked potentials (CAEPs) to speech recorded from frontocentral and temporal sites: three months to eight years of age.** *Int. J. Psychophysiol.* 95 (2), 77–93, <http://dx.doi.org/10.1016/j.ijpsycho.2014.08.1390>.
- Sorokin, A., Alku, P., Kujala, T., 2010. **Change and novelty detection in speech and non-speech sound streams.** *Brain Res.* 1327, 77–90.
- Sussman, E., Steinschneider, M., Gumenyuk, V., Grushko, J., Lawson, K., 2008. **The maturation of human evoked brain potentials to sounds presented at different stimulus rates.** *Hear. Res.* 236 (1–2), 61–79, <http://dx.doi.org/10.1016/j.heares.2007.12.001>.
- Tervaniemi, M., Hugdahl, K., 2003. **Lateralization of auditory-cortex functions.** *brain research. Brain Res. Rev.* 43 (3), 231–246.
- Torppa, M., Lyytinen, P., Erskine, J., Eklund, K., Lyytinen, H., 2010. **Language development, literacy skills, and predictive connections to reading in Finnish children with and without familial risk for dyslexia.** *J. Learn. Disabil.* 43 (4), 308–321, <http://dx.doi.org/10.1177/0022219410369096>.
- Tremblay, K., Kraus, N., McGee, T., Ponton, C., Otis, B., 2001. **Central auditory plasticity: changes in the N1-P2 complex after speech-sound training.** *Ear Hear.* 22 (2), 79–90.
- Vouloumanos, A., Werker, J.F., 2004. **Tuned to the signal: the privileged status of speech for young infants.** *Dev. Sci.* 7 (3), 270–276.
- Vouloumanos, A., Werker, J.F., 2007. **Listening to language at birth: evidence for a bias for speech in neonates.** *Dev. Sci.* 10 (2), 159–164.
- Wechsler, D., 1999. *WISC-III—Wechsler Intelligence Scale for Children—III.* The Psychological Corporation, San Antonio, TX, U.S.A.: Psykologien kustannus, Helsinki, Finland.
- Wechsler, D., 2010. *WISC-IV—Wechsler Intelligence Scale for Children—IV.* NCS Pearson Ltd., London, U.K.: Psykologien kustannus, Helsinki, Finland.
- Wechsler, D., 2009. *WPPSI-III—Wechsler Preschool and Primary Scale of Intelligence—Third Edition.* The Psychological Corporation, San Antonio, TX, U.S.A.: Psykologien kustannus, Helsinki, Finland.