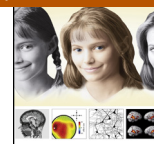




## Developmental Cognitive Neuroscience

journal homepage: <http://www.elsevier.com/locate/dcn>

## Auditory discrimination predicts linguistic outcome in Italian infants with and without familial risk for language learning impairment

Chiara Cantiani<sup>a,\*</sup>, Valentina Riva<sup>a</sup>, Caterina Piazza<sup>b,c</sup>, Roberta Bettoni<sup>a</sup>, Massimo Molteni<sup>a</sup>, Naseem Choudhury<sup>d,e</sup>, Cecilia Marino<sup>a,f,\*\*</sup>, April A. Benasich<sup>d,\*\*\*</sup><sup>a</sup> Child Psychopathology Unit, Scientific Institute, IRCCS Eugenio Medea, Bosisio Parini, Lecco, Italy<sup>b</sup> Bioengineering Lab, Scientific Institute IRCCS Eugenio Medea, Bosisio Parini, Lecco, Italy<sup>c</sup> Department of Electronics Information and Bioengineering (DEIB), Politecnico di Milano, Milano, Italy<sup>d</sup> Center for Molecular and Behavioral Neuroscience, Rutgers University, USA<sup>e</sup> Ramapo College of New Jersey, Mahwah, NJ, USA<sup>f</sup> Centre for Addiction and Mental Health (CAMH), University of Toronto

## ARTICLE INFO

## Article history:

Received 25 April 2015

Received in revised form 23 January 2016

Accepted 1 March 2016

Available online 24 March 2016

## Keywords:

Language acquisition

Infants

Auditory discrimination

Language learning impairment

EEG/ERPs

Cross-cultural

## ABSTRACT

Infants' ability to discriminate between auditory stimuli presented in rapid succession and differing in fundamental frequency (Rapid Auditory Processing [RAP] abilities) has been shown to be anomalous in infants at familial risk for Language Learning Impairment (LLI) and to predict later language outcomes. This study represents the first attempt to investigate RAP in Italian infants at risk for LLI (FH+), examining two critical acoustic features: frequency and duration, both embedded in a rapidly-presented acoustic environment. RAP skills of 24 FH+ and 32 control (FH-) Italian 6-month-old infants were characterized via EEG/ERP using a multi-feature oddball paradigm. Outcome measures of expressive vocabulary were collected at 20 months.

Group differences favoring FH- infants were identified: in FH+ infants, the latency of the N2\* peak was delayed and the mean amplitude of the positive mismatch response was reduced, primarily for frequency discrimination and within the right hemisphere. Moreover, both EEG measures were correlated with language scores at 20 months.

Results indicate that RAP abilities are atypical in Italian infants with a first-degree relative affected by LLI and that this impacts later linguistic skills. These findings provide a compelling cross-linguistic comparison with previous research on American infants, supporting the biological unity hypothesis of LLI.

© 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

Language is a fundamental and complex human attribute. During the first months of life, infants approach language with a set of neuropsychological abilities (e.g., Saffran et al., 1996). These include the ability to perform fine-grained acoustic analyses in the tens of

milliseconds range (i.e., Rapid Auditory Processing or "RAP"). Studies have shown that RAP is critical to the decoding of the speech stream and the subsequent establishment of phonemic maps (e.g., Aslin, 1989; Benasich and Tallal, 2002; Kuhl, 2004), and RAP plays a crucial role in setting up the building blocks of spoken language (Benasich et al., 2006; Choudhury and Benasich, 2011).

For some children, language acquisition is difficult and subsequent language and learning abilities can be impaired. Specific Language Impairment (SLI) and dyslexia are disorders in which language and reading, respectively, are delayed despite normal nonverbal intelligence and adequate educational opportunities. At the behavioral level, phonological deficits have consistently been shown to be a core feature of both disorders (Leonard, 2014; Ramus et al., 2003; Schulte-Körne and Bruder, 2010). However, it is still a matter of long-lasting debate whether such deficits are primarily induced by speech-specific mechanisms or by disturbances in more basic auditory perception such as RAP deficits (e.g., Tallal and Piercy, 1973; Cantiani et al., 2010; Lorusso et al., 2014; Malenfant

**Abbreviations:** RAP, Rapid Auditory Processing; LLI, language and learning impairment; FH+/-, family history positive/negative; ISI, interstimulus interval; MMR, mismatch response.

\* Corresponding author at: Child Psychopathology Unit, Scientific Institute 'Eugenio Medea', via don Luigi Monza, 20, 23842 Bosisio Parini, Lecco, Italy.

\*\* Corresponding author at: Centre for Addiction and Mental Health (CAMH), University of Toronto, 1001 Queen St W - Toronto, ON M6J 1H4, Canada.

\*\*\* Corresponding author at: Center for Molecular and Behavioral Neuroscience, Rutgers University, 197 University Avenue, Newark NJ 07102, USA.

E-mail addresses: [chiara.cantiani@lanostrafamiglia.it](mailto:chiara.cantiani@lanostrafamiglia.it) (C. Cantiani), [cecilia.marino@utoronto.ca](mailto:cecilia.marino@utoronto.ca) (C. Marino), [benasich@andromeda.rutgers.edu](mailto:benasich@andromeda.rutgers.edu) (A.A. Benasich).

<http://dx.doi.org/10.1016/j.dcn.2016.03.002>

1878-9293/© 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/>).

et al., 2012). Specific Language Impairment and dyslexia are often comorbid and aggregate in families suggesting a genetic etiology (Plomin and Kovas, 2005). Around one third of children diagnosed with SLI in preschool-age go on to develop dyslexia by elementary school (e.g., Catts et al., 2005). In addition, bivariate genetic analyses, which are based on cross-trait correlations in twins' pairs and assess the proportion of phenotypic covariance attributable to common genetic contributions, have shown that genetic effects on reading strongly correlate with genetic effects on language in twins from the general population ( $r = .67$ – $1.0$ ) (Hohnen and Stevenson, 1999) and in affected twins ( $r = .53$ – $.86$ ) (Bishop, 2001), supporting the view that language and reading (dis)abilities share common genes and are not etiologically distinct. For these reasons, the term “language-learning impairment” (LLI) has become increasingly popular and encompasses children with either or both disorders. Given the high heritability of LLI, studying infants of biological families with a greater-than-expected prevalence of LLI, and thus at higher genetic risk, is a valuable approach to identify early markers of the disorder (Luyster et al., 2011).

In order to successfully study early auditory processing in infants, age-appropriate techniques that do not require overt responses are needed, and examination of dense-array auditory evoked-response brain potentials (dEEG/ERPs) provides non-invasive functional brain measurements of these skills in infancy. In the first year of life, auditory ERPs are characterized by a positive deflection (P1) at about 150 ms from stimulus onset, followed by a negative peak (N2) at 200–250 ms (e.g., Ceponiene et al., 2002). These two early peaks are referred to as obligatory responses, and are thought to be associated with auditory detection (P1), and feature processing (N2) (e.g., Ceponiene et al., 2008). However, in older children the N2 amplitude has been shown to increase with stimulus repetition (Karhu et al., 1997) and for this reason it was suggested that the N2 indexes the build-up of a neural representation or a sensory memory trace of the repeated stimulus. Recent literature suggests that in typical “oddball” paradigms (where deviants are presented within a series of standard or frequent stimuli), the N2 in response to deviant stimuli might represent the beginning point/onset of the discrimination response (Choudhury and Benasich, 2011). For this reason, in the present manuscript we have labeled this peak on the deviant waveform as “N2\*” to emphasize that this peak could well be functionally different from the N2 on the standard wave.

Finally, in oddball paradigms the electrophysiological pattern characterized by the P1/N2 (N2\*) peaks is generally followed by a mismatch response (MMR), thought to reflect a neural change detection process. This component does not require conscious attention to the stimuli, and thus provides a measure of fine acoustic discrimination abilities even in preverbal infants. Within the infant literature, the MMR is usually characterized by a large positivity at about 300 ms from deviant stimulus onset (e.g., Kushnerenko et al., 2002). The reasons for the positive polarity of the MMR in infancy (as compared to the typical mismatch negativity elicited within the same paradigms in older children and adults) are still not clear, and researchers hypothesize that this polarity shift could be related to the level of alertness or the sleep stage of sleeping infants (e.g., Friederici et al., 2002); to the maturational level of the infant (e.g., Leppänen et al., 2004); to the magnitude of the deviant stimulus change (e.g., Morr et al., 2002); or to particular filter settings (e.g., He et al., 2007). In the present paper the mismatch response will be referred to as the P3 component—labeled to reflect its polarity and average time of onset.

ERPs have been extensively used in the first year of life in order to investigate infants' ability to discriminate changes that occur in the processing of speech, including changes in a single phonetic feature in the consonant or vowel of a syllable, (e.g., Dehaene-Lambertz and Dehaene, 1994; Rivera-Gaxiola et al., 2005) and/or

changes in the duration of the vowel (e.g., Friederici et al., 2002; Guttorm et al., 2005; Leppänen et al., 1999) or of the consonant (e.g., Kushnerenko et al., 2001a; Leppänen et al., 2002). Discrimination of changes in non-speech signals have also been investigated, including frequency changes (e.g., Fellman et al., 2004; Leppänen et al., 2010; Wunderlich et al., 2006) and changes in the temporal aspects of auditory stimulation, such as sound duration (Cheour et al., 2002; Kushnerenko et al., 2001b). Overall, these studies suggest that these acoustical features are encoded in auditory sensory memory, as reflected in the elicitation of the aforementioned ERP responses (especially N2/N2\* and P3), although with differences in the morphology, latency and amplitude of these peaks/components. These cross-stimulus feature differences might imply that speech vs. non-speech stimuli, as well as sound frequency vs. sound duration are encoded differently in the auditory cortex and thus are reflected differently in the ERPs (Ceponiene et al., 2002).

Studies using these techniques, including samples of infants with a positive family history (FH+) for LLI (Benasich et al., 2006; Choudhury and Benasich, 2011; Friedrich et al., 2004) or for dyslexia (Leppänen et al., 2011; van der Leij et al., 2013), demonstrate that early auditory processing of both speech and non-speech stimuli is impaired in FH+ newborns and/or infants from 2 to 6 months of age, suggesting that basic auditory perception and RAP deficits could be causally linked to LLI. Despite the heterogeneity of criteria (i.e. definition of risk, age of evaluation, nature of the stimuli, experimental paradigms and features of the native language), experimental results are robust. Group differences in response to non-speech stimuli are consistently reported in favor of control infants (FH-) for latency (e.g., Choudhury and Benasich, 2011) and/or amplitude (e.g., Choudhury and Benasich, 2011; van Zuijlen et al., 2012) of the main emerging ERP peaks (e.g., N2 and MMR) and these differences are more pronounced for faster (i.e. 70 ms) as compared to slower (i.e. 300 ms) interstimulus intervals (ISI) (Benasich et al., 2006; Choudhury and Benasich, 2011). Similar results have been reported in response to speech stimuli (e.g., Friedrich et al., 2004; Leppänen et al., 2002).

A second aspect that has been extensively investigated in these populations is the predictive value of ERP correlates of early auditory processing for later linguistic and cognitive development. Longitudinal studies have shown that ERP correlates of non-speech auditory processing in infants and toddlers impact later linguistic (e.g., Benasich et al., 2006; Choudhury and Benasich, 2011; van Zuijlen et al., 2012) and/or reading skills (e.g., Leppänen et al., 2010; Molfese, 2000; van Zuijlen et al., 2012). The underlying neural mechanisms contributing to these findings are still to be elucidated, particularly when the specificity of languages is considered.

This study represents the first attempt to investigate RAP as a marker of risk for LLI in Italian FH+ infants. Based on previous research showing consistently poorer processing in FH+ infants at faster ISIs (Benasich et al., 2006; Choudhury and Benasich, 2011), we employed an acoustic paradigm using rapidly-presented (i.e. 70 ms ISI) non-speech stimuli. By replicating the methodology used in previous American studies, we provide an opportunity for cross-linguistic comparison, which can contribute to understanding the issues surrounding the “biological unity hypothesis” of developmental disorders (Paulesu et al., 2001). According to this hypothesis, these disorders are characterized by the same neurocognitive deficit with the same universal basis in the brain, despite the variability and the culture-specific manifestations at the clinical level.

In order to specifically address the role of differing acoustic attributes on early RAP, we investigated two critical acoustic features: changes in fundamental frequency and variation in sound duration. We hypothesized that changes in *fundamental frequency* would be closely related to fine-grained acoustic analyses and changes in *sound duration* more related to slowly-varying enve-

lope changes, (i.e., to the analyses of the rhythmic timing). Previous studies on American FH+ infants, investigated RAP using frequency changes (Benasich et al., 2006; Choudhury and Benasich, 2011). Inclusion of the duration deviant was based on the observation that discrimination of duration changes may be critical for phonemic discrimination in Italian, whereas it is not a phonemic cue in many other languages, including English. The Italian language allows clustering of the same consonant in vowel contexts, a phenomenon known as “consonant germination”, where the duration of the intervocalic consonant is a discriminative feature indicating differences in meaning (e.g., the word “capa” [ka:pa] boss vs. the word “cappa” [kap:a] mantle). Since duration is an important acoustic cue in Italian, we hypothesized that early discrimination of this acoustical feature would be specifically predictive of language acquisition in Italian infants.

This hypothesis was based on studies that suggest that linguistic experience with “difficult to discriminate” durational cues alters ERP response to this feature (e.g., Nenonen et al., 2003; Tervaniemi et al., 2006) and that this enhanced processing extends to the non-speech domain (Kirmse et al., 2008). In the Kirmse et al. (2008) study, adult native-speakers of two languages were tested: Finnish, where vowel (and consonant) duration is crucial for phoneme distinctions, and German, where vowel duration is still important but of lower relevance as a phonetically distinctive cue with respect to Finnish. The results showed that Finnish ERP responses were characterized by shorter MMN latency for duration changes in both vowels and non-speech tones. Although this study demonstrates that populations where durational cues are relevant to speech discrimination are faster in detecting variations in duration cues in both speech and non-speech stimuli, it is important to note that these abilities are set up very early in life and are based on pre-linguistic shaping of language representations (Benasich et al., 2014; Kuhl et al., 1992). Thus, within populations where duration cues are important for speech segmentation, any variability in setting up the auditory system that disadvantages the processing of such cues could lead to atypical language acquisition. For this reason, we expected that language experience with durational cues in Italian, specifically the ability to process non-speech durational cues, could be a marker of familial risk and predictive of language acquisition in Italian infants.

In the present study, a multi-feature paradigm has been introduced to elicit comparable responses for two different auditory attributes, specifically changes in frequency and duration, within the same paradigm in the same sample of infants at familial risk for LLI. This novel multi-deviant paradigm is employed to achieve two distinct but related aims: (1) to test RAP as a risk marker of LLI in 6-month-old Italian infants at-risk versus not-at-risk for LLI, by using age-appropriate techniques that do not require overt responses; (2) to identify developmental trajectories from early RAP to later language acquisition, by assessing correlations between RAP and expressive language skills at 20 months of age. Overall, we expected discrimination of both frequency and dura-

tion changes to be impaired in Italian infants at familial risk for LLI, and for these deficits to impact later language acquisition. Since a similar paradigm using frequency change has already been applied to American infants, this study provides a cross-linguistic comparison, allowing us to examine the role of non-speech auditory discrimination in infants with differing native language experience.

## 2. Method

### 2.1. Participants

Two groups of infants are included in this study: typically developing control infants and infants at high familial risk for LLI. Infants are tested at 6 months of age using a test battery including recording of EEG/ERPs and behavioral assessments. In addition, follow-up information on language development is collected at 20 months of age by means of a parental questionnaire filled out by parents.

Families were recruited by local advertisement and/or physician referral from three hospitals within the Lecco and Monza-Brianza area (Northern Italy). After full disclosure of the methodology, parents could declare their availability to take part in the study. Written informed consent was obtained from all parents prior to testing. Ethical and Scientific Committees of all institutes involved in the study approved the study protocol.

Socio-demographic, prenatal and perinatal information were collected using an ad hoc form filled out by parents. Infants were tested on the cognitive subscale of the Bayley Scales of Infant Development (Bayley, 1993). Infants were included in the study if they had: (1) both parents native-Italian speakers, (2) gestational age  $\geq 37$  weeks, (3) birth-weight  $\geq 2500$  grams, (4) APGAR scores at birth at 1' and 5'  $\geq 9$  and (5) a Bayley Cognitive Score  $\geq 7$ , (6) absence of certified diagnosis of intellectual deficiency, attention-deficit disorder, sensorial and neurological disorders or autism within first-degree relatives.

Infants were assigned to either a family-history negative (FH-) control group or a family-history positive (FH+) group based on a two-step procedure. First, an intake interview was used to determine if there was a clinical diagnosis of LLI, sensorial, neurological and/or intellectual disorders, attention-deficit disorder, and/or autism in any first-degree relatives. Second, all parents were tested for reading performance by a trained psychologist, using a standardized text-reading test (Judica and De Luca, 2005), a single-word test, and a pseudoword-reading test (Sartori et al., 1995). All measures are age-standardized on the Italian population.

Infants were assigned to the FH+ group if at least one first-degree relative (1) had a certified (clinical) diagnosis of LLI and/or (2) performed at least two standard deviations (SD) below the population mean on at least one out of the three administered reading tests. This yielded 24 infants (13 males and 11 females), of which 7 had a positive family history for language impairment, 14 for dyslexia and 3 for both. A composite score was generated from the z-scores each parent received on the three reading tests. The parent's mean

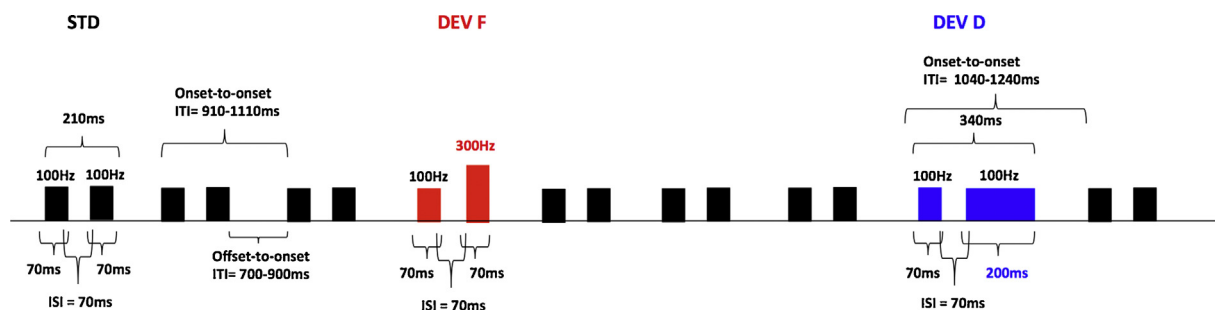


Fig 1. Schematic representation of the non-speech multi-feature oddball paradigm.

**Table 1**

Descriptive statistics: mean (standard deviation) and group comparisons on individual, demographic and clinical characteristics. (a) Mother's and father's educational level were scored on a 9-point ordinal scale created ad hoc and based on the Italian school system. Scores ranged between 10, corresponding to the fifth-grade of elementary school; 50, equivalent to a high-school diploma; and 90, corresponding to post-doctoral degree; (b) socio-economic status was scored according to the Hollingshead 9-point scale, whereby a score ranging 10–90 was assigned to each parental job and the higher of two scores was used when both parents were employed (Hollingshead, 1975). Scores ranged between 10, corresponding to unskilled workers; 50, corresponding to sales workers; and 90, corresponding to major professional.

	FH+ (N = 24)	FH- (N = 32)	t(df)	p
Birth weight (grams)	3223.7 (325.4)	3265.8 (485.4)	0.37 (54)	.699
Gestational age (weeks)	39.2 (1.3)	39.6 (1.3)	1.11 (54)	.273
Mother's age (years)	34.4 (3.7)	34.5 (4.0)	0.15 (54)	.883
Father's age (years)	37.1 (3.7)	37.2 (5.2)	0.05 (54)	.960
Mothers' educational level <sup>a</sup>	52.9 (16.0)	59.1 (14.4)	1.50 (54)	.138
Fathers' educational level <sup>a</sup>	40.8 (20.0)	49.4 (17.6)	1.69 (54)	.097
Socioeconomic status <sup>b</sup>	56.0 (22.0)	64.7 (13.7)	1.80 (54)	.099
Bayley cognitive subscale	12.5 (1.7)	12.2 (1.8)	-0.35 (54)	.728

composite z-score for the FH+ group were: mothers,  $M = -0.32$ ,  $SD = 1.07$ ; fathers,  $M = -1.58$ ,  $SD = 1.55$ .

Infants were assigned to the FH- group if all first-degree relatives (1) had an absence of certified diagnosis of LLI and (2) a performance above 0.5 SD below the population mean on all three reading tests. This yielded 32 infants (17 males and 15 females). The parent's mean composite reading scores for the FH- group were: mothers,  $M = 0.31$ ,  $SD = 0.64$ ; fathers,  $M = 0.23$ ,  $SD = 0.69$ . Demographic and clinical characteristics of the two groups are reported in Table 1.

## 2.2. General testing procedure

Families were contacted two weeks prior to the child's 6-month birthday: a first visit to the laboratory was scheduled at 6 months, 15 days  $\pm$  two weeks (FH-,  $M = 6.49$ ,  $SD = 0.48$ ; FH+,  $M = 6.36$ ,  $SD = 0.33$ ). During the visit, the EEG/ERP was recorded (Section 2.3) and the standardized tests administered. Language skills were assessed at 20 months of age ( $M = 20.56$ ,  $SD = 0.32$ ). Before the child's 20-month birthday, caregivers were mailed packets containing the Language Development Survey (Section 2.4) to complete at home. Parents were asked to bring the filled-in forms to a scheduled laboratory visit at 20 months, 15 days  $\pm$  two weeks (during this visit children participated in a different ERP experiment that is not reported in the present manuscript). Data at 20 months were available for 39 subjects (70% of the original sample).

## 2.3. Experimental task on multi-feature auditory discrimination

### 2.3.1. Stimuli

RAP was assessed by means of an electrophysiological task. The paradigm and stimuli were based on previous infant research (Benasich et al., 2006; Choudhury and Benasich, 2011). In the non-speech multi-feature paradigm, pairs of complex tones with an ISI of 70 ms were presented. Fig. 1 schematically represents the experimental paradigm. The first tone in the pair always had a fundamental frequency of 100 Hz with 15 harmonics (6 dB roll-off per octave) and duration of 70 ms (5 ms rise time and 5 ms fall time). For standard tone-pairs (STD) the same 100 Hz tone was repeated twice (i.e., 100–100 Hz). Two deviant tone-pairs differing with respect to the second tone were presented: for the frequency deviant (DEVF) the second tone had a fundamental frequency of 300 Hz; for the duration deviant (DEV) the second tone had a duration of 200 ms.

The stimuli were presented in a passive oddball paradigm where 1200 stimuli (80% STD, 10% DEVF, 10% DEV) were pseudo-randomized, so that at least three standard tone-pairs were

presented before each deviant pair. A temporal jitter of  $\pm 100$  ms was added to the intertrial interval (offset-to-onset, ITI) and thus ITI randomly varied from 700 to 900 ms while the onset-to-onset ITI varied from 910 to 1110 and from 1040 to 1240, for DEVF and DEV) stimulus types, respectively.

All stimuli were presented free field at an intensity of 75 dB via speakers located on either side of and equidistant (95 cm) from the subject.

### 2.3.2. EEG data acquisition and preprocessing

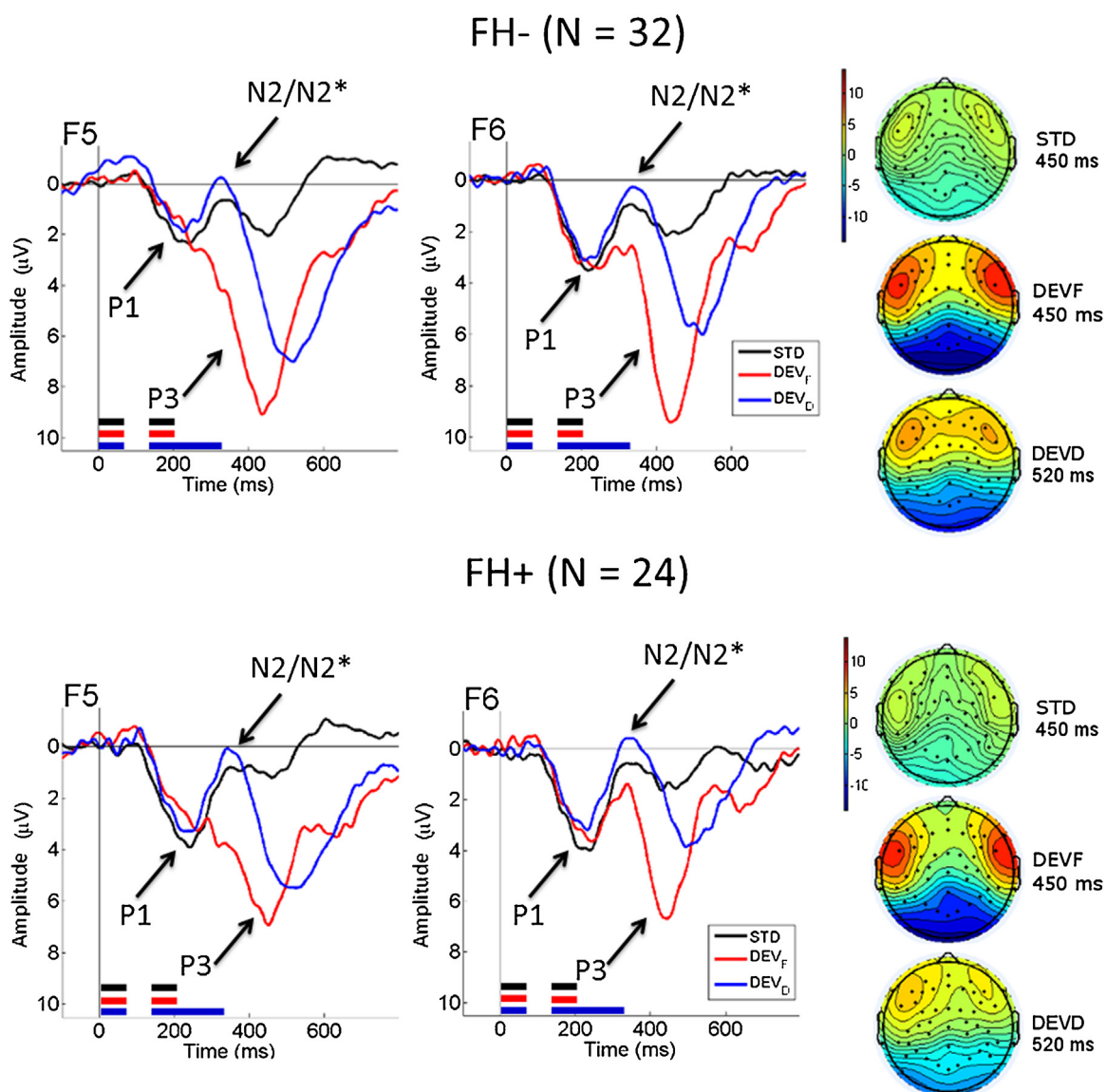
During EEG recording, children were seated on their caregiver's lap in a sound-attenuated and electrically shielded room. Age-appropriate silent movies/cartoons were played on a video monitor in front of the infants. An experimenter was also present in the room to entertain the child with age-appropriate silent toys.

Auditory ERPs were recorded from 60 scalp sites using a dense-array EGI recording system (Electrical Geodesics, Inc., Eugene, Oregon). Vertex was used as an online reference. EEG was sampled at 250 Hz and bandpass filtered (0.1–100 Hz) online. After recording, data were exported to a MATLAB (Mathworks, Natick, MA) compatible format and processed using EEGLAB (Delorme and Makeig, 2004) and ERPLAB (Lopez-Calderon and Luck, 2014).

Continuous EEG data were bandpass filtered offline at 0.5–30 Hz. After filtering, as a first step channels with high impedance ( $>50$  K $\Omega$ ), or visually evident noise, were interpolated with a spherical spline. No more than 12 of the 60 channels, which correspond to 20%, were interpolated ( $M = 3.24$ ,  $SD = 2.35$ , range = 0–12). The signals were then re-referenced to an average reference (based on the 60 channels) and the vertex channel was reincluded within the channels (resulting in 61 channels). After re-reference, due to significant movement-related artifacts and a high rate of interpolation, the 13 outermost channels (Supplementary material, Fig. S2) were removed from analysis for all children. The remaining 48 channels were considered for analyses. For the standard stimuli, only the responses to the immediate pre-deviant standard stimuli were included in the average in order to achieve a more comparable signal-to-noise ratio to the averaged deviants. The continuous EEG was segmented according to stimulus type (pre-deviant STD, DEVF and DEV) with 900 ms epoch lengths (100 ms before stimulus presentation and 800 ms after). The 100 ms pre-stimulus segment was used for baseline correction. Bad EEG epochs were identified and rejected using two automatic criteria and visual inspection. First, a moving window (200 ms width, 50 ms step) was used to identify segments containing signals with voltage differences  $>150$   $\mu$ V. Then, an additional artifact-rejection tool was used to specifically detect and remove any muscle activity present in the infants' data. In particular, since motor artifacts are characterized by relatively strong high frequency activity ( $>20$  Hz), which is easily identifiable in the frequency domain, we used spectral parameters to detect epochs corrupted with muscle artifact (Delorme et al., 2007). Specifically, trials whose spectrum (in one or more channels) deviates from baseline by  $+25/-100$  dB in the frequencies  $>20$  Hz were removed. A minimum of 60 artifact-free trials was used for averaging ERPs. 130 pre-deviant STD ( $SD = 25.9$ , range = 96–183), 70 DEVF ( $SD = 11.0$ , range = 60–104) and 72 DEV) ( $SD = 11.9$ , range = 60–100) trials remained for the FH- group, whereas a mean of 130 pre-deviant STD ( $SD = 25.2$ , range = 99–190), 74 DEVF ( $SD = 11.6$ , range = 60–99) and 71 DEV) ( $SD = 8.9$ , range = 60–95) trials remained for the FH+ group. The number of accepted standard and deviant trials included for each group did not differ by group, as revealed by one-way ANOVAs ( $p$ -values ranged from 0.2–0.9).

### 2.4. Standardized measures of linguistic outcome

The Language Development Survey (LDS) is a 310-word parental-report screening tool for expressive language delay in



**Fig. 2.** Grand average waveforms for FH– (upper panel) and FH+ (lower panel B) infants. Channels F5 and F6 located respectively on left and right fronto-central regions are shown. The standard waveform (STD, black line) is plotted against the waveforms for the frequency deviant (DEVF, red line) and the duration deviant (DEVd, blue line). The black, red and blue lines beneath each waveform indicate the temporal sequence of stimulus presentation in each stimulus type (black lines=STD; red lines=DEVF; blue lines=DEVd). Negative voltage is plotted upward. On the right, the topographical maps of the distribution of P3 amplitude for the three stimulus types (STD, DEVF, DEVd) are shown in the middle of the time-window of interest (450 ms for STD and DEVF and 520 ms for DEVd). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

toddlers (18–35 months of age). The inventory provides a total vocabulary score and has been shown to be significantly correlated with the MacArthur Communicative Development Inventories: Words and Sentences ( $r = .95$ ) (Fenson et al., 1993; Rescorla et al., 2005). Norms are available from ages 18 to 35 months (Rescorla and Alley, 2001) and yield percentile ranks and age-equivalent scores. The LDS has recently been standardized on an Italian population (Rescorla et al., 2014).

## 2.5. Analytic strategy

### 2.5.1. Analyses of ERP data

Time windows and electrode sites to be submitted to statistical analyses were selected based on mass univariate analyses applied to the FH– ERP data (for the full description of this procedure see the Supplementary material). This procedure allows the identification of channel clusters and time windows where differences between

stimulus types are significant, taking into account application of the appropriate corrections for multiple comparisons (Groppe et al., 2011). See Supplementary material, Fig. S1 for a graphical representation of the permutation test results that drove the selection of the time windows and the electrode sites to be submitted to statistical analyses.

For each participant and each component of interest, ERPs were extracted from a subset of 18 electrodes localized in the left and right fronto-central areas, with the selection based on the results of the mass univariate analyses. Data were then averaged in two clusters corresponding to left and right fronto-central areas, each including 9 channels (see Supplementary material, Fig. S2).

P1 and N2/N2\* peaks were extracted in the range of 100–300 ms and 250–450 ms, respectively. Peak amplitudes were measured as the largest positive (for P1) or negative (N2) peak amplitudes relative to the corrected baseline; latencies of the peaks were measured from the onset of the first tone in the pair. The peaks were

identified through an automatic algorithm, detecting the maximum/minimum peak in the aforementioned time windows for each individual. For the later positive response, corresponding to the mismatch response (P3 peak), mean amplitudes were calculated for different time-windows selected on the basis of the previous mass-univariate analyses (see Supplementary material and specifically Fig. S1): 350–550 ms time window for the STD and DEVF waveforms and a 420–620 ms for DEVD waveforms. Mean amplitude was extracted within these time windows (instead of at peak amplitude) because the components were wide and without clearly identifiable peaks.

### 2.5.2. Statistical analyses

Statistical analyses were conducted separately for each ERP component of interest. Different statistical approaches were used for the different peaks/components based on our a priori hypotheses. To assess differences on the P1 peak amplitude and latency, where no differences between stimulus types were expected, overall repeated-measures ANOVA models with Stimulus type (3 levels: STD vs. DEVF vs. DEVD), Group (2 levels: FH+ vs. FH–), and Hemisphere (2 levels: Left vs. Right) were used. For the N2 peak, functional differences between stimulus types were expected, since the N2 in response to deviant stimuli is thought to represent the onset of the discrimination response. Based on the previous literature (Choudhury and Benasich, 2011), we have labeled this peak on the deviant waveform as “N2\*” so as to not imply that this peak is functionally similar to the N2 on the standard wave. In order to treat N2 and N2\* separately, three separate  $2 \times 2$  ANOVAs (Group  $\times$  Hemisphere) were conducted for each stimulus type. Finally, differences in the mean amplitude of the P3 peak were investigated using two separate 3-way ANOVA models, in order to directly contrast each deviant stimulus type with the standard. This particular analytic strategy allowed us to investigate the amplitude of the mismatch response as a measure of auditory discrimination. Model 1 compared P3 mean amplitude to Stimulus Type (STD versus DEVF) by Group by Hemisphere ( $2 \times 2 \times 2$ ). Model 2 was a parallel analysis with STD and DEVD in Stimulus Type. Additionally, independent-sample *t*-tests were used to assess group differences by each stimulus type within each hemisphere.

Finally, Pearson’s product moment correlation analyses were conducted to assess associations between infant ERP components (peak amplitude and latency of P1 and N2/N2\* and mean amplitude of P3, separated by stimulus type and hemisphere) and 20-month language abilities (expressive vocabulary percentile score). Given the high number of correlation analyses performed (30), critical alpha level was set at .01.

## 3. Results

### 3.1. Morphology of the pre-deviant STD, DEVF and DEVD waveforms

In both groups, there was a clear positive peak (P1) appearing at around 220 ms followed by a negative peak (N2/N2\*), appearing at around 350 ms. The N2\* for the two deviant stimulus types was followed by a large positive peak (P3) for both groups, occurring with different timing for each stimulus type. The P3 appeared at approximately 450 ms for DEVF and at about 520 ms for DEVD. On the DEVF waveform, the N2\* peak was truncated by the following large positivity, mainly on the right hemisphere. In Fig. 2, topoplots and grand average waveforms are shown for the two groups (upper panel: FH– group; lower panel: FH+ group) for two channels, 13 (corresponding to F5 in the 10/10 international system) and 59 (corresponding to F6), each representative of the two channel clusters, located respectively in the left and right fronto-central areas.

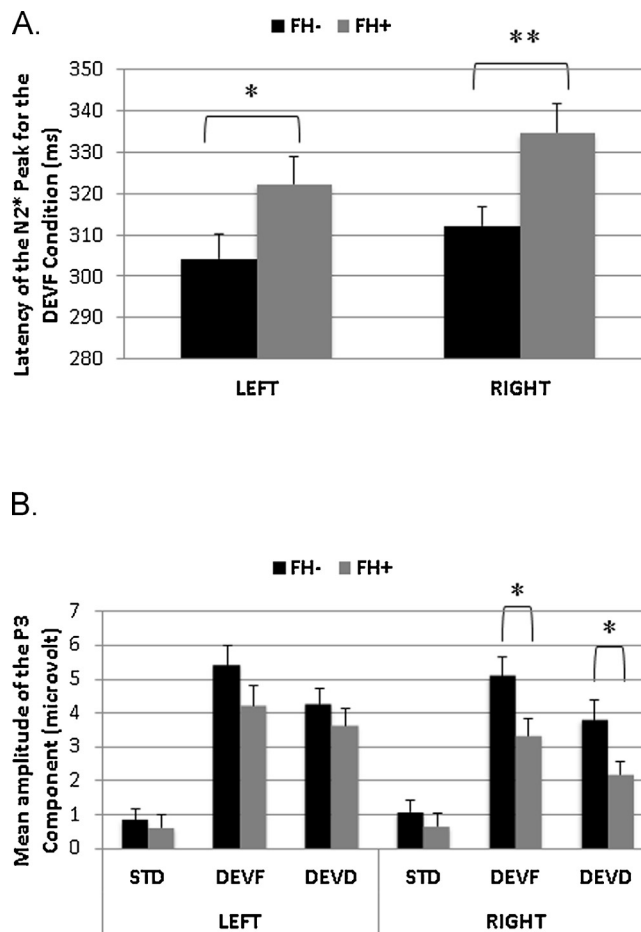


Fig. 3. Bar graphs representing: (A) latency of the N2\* (ms) for the DEVF Stimulus type by Group (FH– in black and FH+ in gray) and Hemisphere (Left and Right), and (B) mean amplitude of the P3 component ( $\mu$ V) by Group, Hemisphere and Stimulus type (STD vs. DEVF vs. DEVD). Error bars represent standard error of the mean. Significant differences between groups are reported (\* $<.05$ ; \*\* $<.01$ ).

### 3.2. Analyses of the ERP waveforms

The overall three-way ANOVA models revealed a significant main effect of Hemisphere for both the P1 peak amplitude  $F(1,54)=11.81$ ,  $p=.001$ , and the P1 peak latency  $F(1,54)=8.47$ ,  $p=.005$ . Overall, for all three stimulus types and in both groups the P1 peak was larger,  $M=3.83$ ,  $SD=1.62$  and faster,  $M=218.7$ ,  $SD=18.1$  on the right compared to the left hemisphere ( $M=3.29$ ,  $SD=1.47$  and  $M=225.6$ ,  $SD=20.9$  for amplitude and latency, respectively, Table 2). No significant main effect or interaction was found.

A main effect of Hemisphere was found for the N2\* peak amplitude for the DEVF stimulus,  $F(1,54)=10.00$ ,  $p=.003$ , with smaller amplitude in the left hemisphere ( $M=0.73$ ,  $SD=2.36$ ) compared to the right ( $M=-0.25$ ,  $SD=2.39$ ). This result is coherent with the evidence that in both groups the N2\* was truncated by the following large positivity for DEVF, mainly in the left hemisphere. Again for the DEVF, a significant main effect of Group was found for N2\* peak latency,  $F(1,54)=8.14$ ,  $p=.006$ . Overall, the N2\* peak was delayed for the FH+ group ( $M=328.6$ ,  $SD=26.6$ ) compared to the FH– group ( $M=308.3$ ,  $SD=26.1$ ) (Fig. 3, Panel A). A main effect of Hemisphere was also found for this stimulus type,  $F(1,54)=4.90$ ,  $p=.031$ , with overall longer latency seen in the right hemisphere ( $M=321.9$ ,  $SD=32.0$ ) compared to the left ( $M=312.1$ ,  $SD=33.2$ ). This result also suggests that the N2\* had been truncated by the following large positivity in the left hemisphere. No significant effect

**Table 2**

Mean (standard deviation) of the amplitude (expressed in microvolt) and the latency (expressed in milliseconds) of the three auditory peaks/components (P1, N2, P3) separate for Group (FH– and FH+), Hemisphere (Left and Right) and Condition (STD, DEVF, DEVD). (a) Values corresponding to N2 are reported for the STD condition whereas values corresponding to N2\* are reported for the deviant conditions (DEVF and DEVD).

			Peak Amplitude ( $\mu\text{V}$ )		Peak Latency (ms)	
			FH–	FH+	FH–	FH+
P1 peak	Left	STD	2.77 (1.43)	3.89 (1.77)	222.9 (30.7)	224.8 (25.3)
		DEVF	3.27 (1.94)	3.58 (2.33)	233.3 (39.3)	229.1 (33.1)
		DEVD	2.83 (1.97)	3.72 (1.66)	220.0 (33.5)	223.4 (23.9)
	Right	STD	3.47 (1.54)	4.12 (1.74)	216.6 (24.5)	215.6 (19.0)
		DEVF	4.08 (2.04)	4.18 (2.13)	217.7 (28.0)	226.7 (31.0)
		DEVD	3.53 (2.34)	3.77 (2.13)	220.3 (21.5)	215.5 (27.2)
N2/N2* peak <sup>a</sup>	Left	STD	–0.51 (1.50)	–0.39 (1.93)	351.5 (34.5)	376.2 (35.0)
		DEVF	0.89 (2.07)	0.52 (2.72)	304.4 (32.3)	322.4 (32.2)
		DEVD	–1.24 (1.78)	–0.91 (1.94)	338.7 (30.5)	350.2 (24.6)
	Right	STD	–0.31 (1.68)	–0.64 (1.87)	358.2 (36.7)	364.0 (37.4)
		DEVF	0.10 (2.25)	–0.72 (2.54)	312.2 (26.2)	334.7 (35.0)
		DEVD	–0.87 (2.56)	–1.49 (1.71)	343.6 (33.8)	356.5 (39.6)
			Mean Amplitude ( $\mu\text{V}$ )			
			FH–	FH+		
P3 component	Left	STD	0.85 (1.90)	0.61 (1.87)		
		DEVF	5.41 (3.34)	4.20 (3.10)		
		DEVD	4.25 (2.80)	3.62 (2.45)		
	Right	STD	1.07 (1.98)	0.65 (1.96)		
		DEVF	5.08 (3.27)	3.34 (2.52)		
		DEVD	3.80 (3.38)	2.17 (2.02)		

was found for STD and DEVD for the N2/N2\* peak amplitude or latency (Table 2).

The group difference for the N2\* latency response for the DEVD stimulus type was then examined in more detail. All the data were inspected manually in order to detect the proportion of children in each group clearly showing a N2\* peak at the individual level (defined as having a detectable peak in more than half of the channels for each cluster). A relatively high proportion of children clearly showed a N2\* peak at the individual level in both groups: in the FH– group, 29/32 (91%) in the left hemisphere and 30/32 (94%) in the right hemisphere; in the FH+ group, 20/24 (83%) in both the left and right hemisphere. Statistical analyses were repeated including only the children with a clearly detectable peak, and the results reported above were confirmed (main effect of Group,  $F(1,42) = 4.42$ ,  $p = .042$ , with N2\* peak delayed for the FH+ group,  $M = 328.9$ ,  $SD = 28.7$ , as compared to the FH– group,  $M = 310.6$ ,  $SD = 27.6$ . Group differences were significant in both the left,  $t(47) = -2.04$ ,  $p = .047$  and the right hemispheres,  $t(48) = -2.87$ ,  $p = .006$  (left: FH+,  $M = 325.2$ ,  $SD = 33.9$ ; FH–,  $M = 305.1$ ,  $SD = 33.9$ ; right: FH+,  $M = 336.8$ ,  $SD = 29.1$ ; FH–,  $M = 314.4$ ,  $SD = 25.5$ ).

When contrasting the P3 mean amplitude for the STD and DEVD stimuli by Group and by Hemisphere ( $2 \times 2 \times 2$  ANOVA, Model 1), a significant Stimulus type  $\times$  Group interaction,  $F(1,54) = 4.41$ ,  $p = .040$  and a main effect of Stimulus type,  $F(1,54) = 182.43$ ,  $p < .001$  were found. As expected, mean amplitude was significantly higher for DEVD ( $M = 4.61$ ,  $SD = 2.85$ ) than STD ( $M = 0.82$ ,  $SD = 1.70$ ), with this difference being more prominent in the FH– group ( $M = 5.25$ ,  $SD = 2.99$  and  $M = 0.96$ ,  $SD = 1.72$ , for DEVD and STD, respectively) than in the FH+ group ( $M = 3.77$ ,  $SD = 2.47$  and  $M = 0.64$ ,  $SD = 1.69$ , for DEVD and STD, respectively). Simple effects analyses to evaluate group differences by stimulus type within each hemisphere showed significantly higher mean amplitude for the FH– group in the DEVD stimulus for the right hemisphere only (Table 2,  $t(54) = 2.17$ ,  $p = .034$ ; see Fig. 3, panel B).

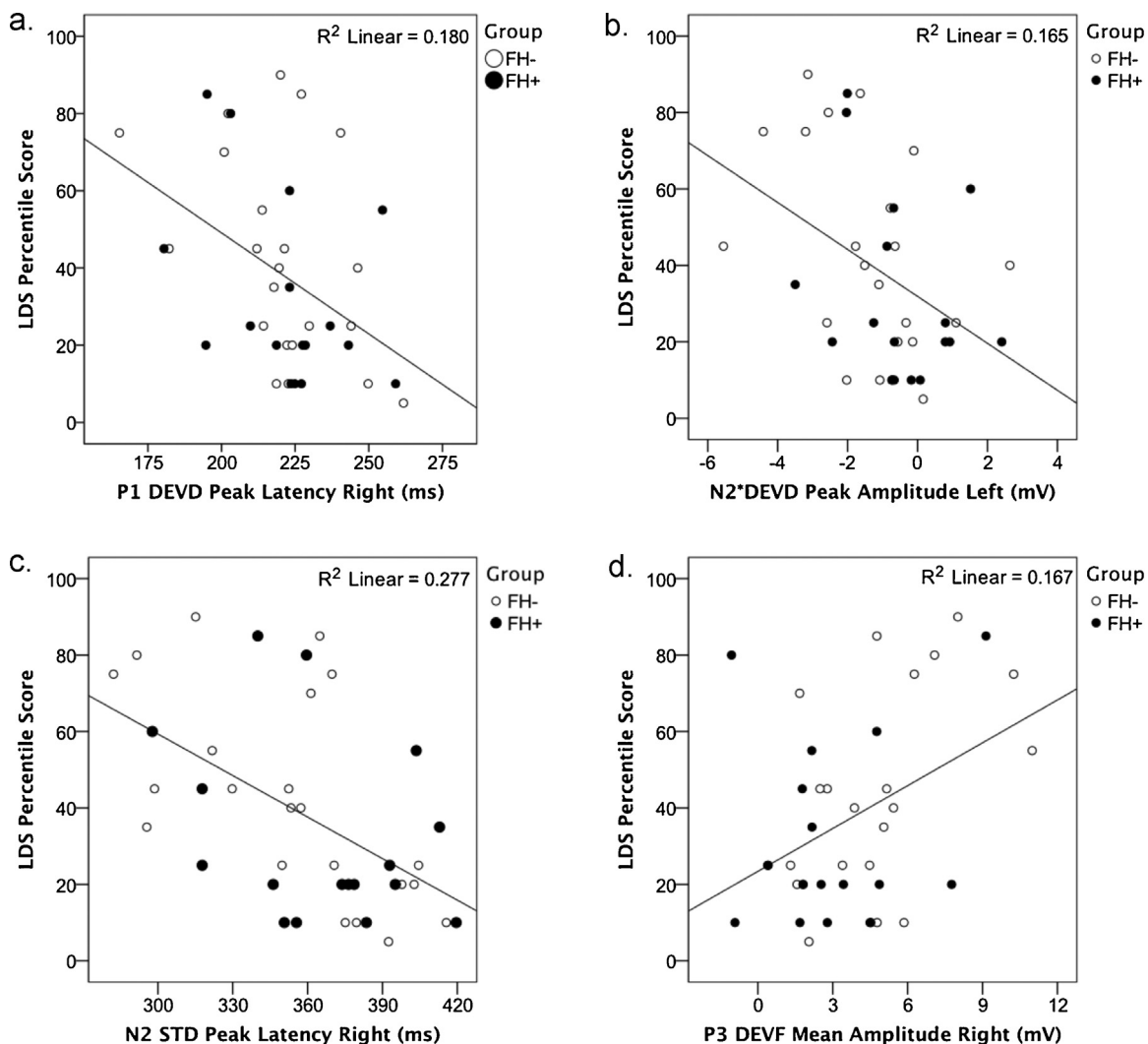
Similar analyses were conducted to compare the P3 mean amplitude for the STD and DEVD stimuli by Group and by Hemisphere ( $2 \times 2 \times 2$  ANOVA, Model 2), revealing a main effect of Stimulus type,  $F(1,54) = 81.52$ ,  $p < .001$ , and a significant Stimulus

type  $\times$  Hemisphere interaction,  $F(1,54) = 6.90$ ,  $p = .011$ . As expected, mean amplitude was higher for DEVD ( $M = 3.54$ ,  $SD = 2.52$ ) than STD, with this difference being more prominent in the left ( $M = 0.75$ ,  $SD = 1.88$  and  $M = 3.98$ ,  $SD = 2.65$ , for STD and DEVD, respectively) than in the right hemisphere ( $M = 0.89$ ,  $SD = 1.96$  and  $M = 3.10$ ,  $SD = 2.97$ , for STD and DEVD, respectively). Finally, simple effects analyses to evaluate group differences for each stimulus type within each hemisphere showed significantly higher mean amplitude for the FH– group in the DEVD stimulus type for the right hemisphere only (Table 2,  $t(54) = 2.09$ ,  $p = .041$ ; see Fig. 3, panel B).

### 3.3. Associations between ERPs and language abilities at 20 months of age

Predictions from the 6-month ERPs to performance on 20-month language measures (Language Development Survey–LDS) (Rescorla et al., 2014) were assessed using Pearson's product moment correlations. For ERP measures, peak amplitude and latency of P1 and N2/N2\*, and mean amplitude of P3 for each stimulus type (STD, DEVD, DEVD) and each hemisphere (left and right) were entered in the correlation analyses. Pearson correlations were thus computed between the resulting 30 ERP variables and the measure of expressive language at 20 months of age. Follow-up data were available for 39 subjects, 70% of the larger sample. The percentile score based on gender-specific norms was used. Since no differences emerged between groups in the percentile scores (FH–,  $n = 22$ ,  $M = 42.27$ ,  $SD = 26.80$ ; FH+,  $n = 17$ ,  $M = 32.35$ ,  $SD = 28.18$ ), correlations were computed using the combined groups. Correlations revealed 4 (out of 30) statistically significant associations between ERP peaks at 6 months and expressive vocabulary at 20 months (Table 3 and Fig. 4). Infants with faster obligatory responses (P1 and N2), more negative N2\* amplitude and more positive P3 amplitude produced more words at 20 months of age.

To deepen our understanding of these results, we dichotomized both the linguistic outcome and the ERP measures. The cut-offs were set at the mean for each value. For each of the correlations significant at the critical alpha level of .01 (see Fig. 4), we estimated the proportion of children within each group that fell within the lower



**Fig. 4.** Person product moment correlations between 6-month ERP measures and 20-month expressive language (percentile score in the Language Development Survey) for FH– (white dots) and FH+ (black dots) infants ( $N = 39$ ). In particular, correlations include (a) latency of the P1 peak for DEVD, (b) amplitude of the N2\* peak for DEVD (c) latency of the N2 peak for STD and (d) mean amplitude of P3 component for DEVD.  $R^2$  scores are reported.

**Table 3**  
Summary of the Person product moment correlations significant at the critical alpha level of .01 between 6-month ERP measures and 20-month expressive language (percentile score in the Language Development Survey) for the overall sample ( $N = 39$ ).

Peak	Condition	Hemisphere	$R$	$p$
P1 latency	DEVD	Right	-.42	.007
N2 amplitude	DEVD	Left	-.41	.010
N2 latency	STD	Right	-.53	.001
P3 amplitude	DEVD	Right	.41	.010

quadrant (poor talkers with less robust ERP components—long peak latencies and/or low amplitudes). The proportion of poor talkers with less robust ERP components is higher for the FH+ group with respect to the FH– group for each of the significant correlations (P1 latency: FH–, 7/22, 32%; FH+, 9/17, 53%; N2 amplitude: FH–, 6/22, 27%; FH+, 9/17, 53%; N2 latency: FH–, 8/22, 36%; FH+, 8/17, 47%; P3 amplitude: FH–, 5/22, 23%; FH+, 9/17, 53%) (odds ratios ranging from 1.29 to 2.32). Overall, the risk of having lower expressive vocabulary scores at 20 months of age doubles when an infant is at familial risk for LLI and in addition presents abnormal ERP measures (longer peak latencies and lower amplitudes).

#### 4. Discussion

The main aim of this study was to investigate RAP skills in an Italian sample of 6-month-old infants with and without family history of LLI. This is the first study in such a sample and provides evidence for the importance of fine-grained auditory processing skills, such as RAP, in a language different from English. Specifically, two critical acoustic features that have been shown to be systematically impaired in individuals with dyslexia (for review, [Hämäläinen et al., 2013](#)) were assessed using a multi-feature paradigm: fundamental frequency and sound duration. Both acoustic features were embedded in a rapidly-presented acoustic environment (i.e., 70 ms ISI) in order to maximize differences between groups and specifically test Rapid Auditory Processing (RAP) abilities. The changes in fundamental frequency closely replicate previous paradigms ([Benasich et al., 2006](#); [Choudhury and Benasich, 2011](#)) and thus provide a compelling cross-linguistic comparison. The results for discrimination of frequency change will be discussed and compared below to analogous results from the earlier American studies. This direct comparison allows us to examine the influence of differing native language experience on the discrimination of non-speech frequency changes. The change in sound duration included here represents a novel addition to the paradigm and have been included



in order to assess particular characteristics of Italian, where sound duration is a critical phonemic difference and thus must recruit attention.

Overall, our results show that processing and discrimination of auditory changes for both fine-grained and slowly-varying envelope information already appear to be compromised in 6-month-old Italian infants with a family history of LLI. In particular, differences between groups of infants with and without a family history of LLI emerge at several levels, including ERP measures reflecting the building of memory traces and onset of the discrimination response (i.e., latency of the N2\* peak) as well as ERP components reflecting the MMR, which indexes the discrimination between standard and deviants (i.e., mean amplitude of the P3 component). In addition, these early ERP measures appear to predict later language acquisition (expressive vocabulary at 20 months of age). In the following sections, after describing the typical ERP patterns observed, we will discuss the above-mentioned results in more detail.

#### 4.1. Results on typically developing infants

Our results with a cohort of typically developing 6-month-old Italian infants show the expected electrophysiological pattern constituted by the obligatory responses (P1-N2/N2\*) followed by a large positive response (P3). Visual inspection of the waveforms indicates that both acoustic features are already detectable to the individuals at this early age. However, some differences emerged between the DEVF and DEVD stimulus types in the latency of the P3 peak (which occurred at least 70 msec later for the DEVD). These differences are clearly related to the point in time when these acoustic differences can first be perceived (at the onset of the second tone for DEVF stimuli and at the point where the second tone offset begins to exceed the duration of the standard stimulus for the DEVD stimuli). Similarly, an earlier study of newborns aged 1–3 days (Ceponiene et al., 2002) reported that the latency of the MMR for changes in duration in non-speech stimuli (STD = 200 ms tones and DEV = 100 ms tones) was significantly longer than that for frequency changes in non-speech stimuli (100 ms tones differing in fundamental frequency). The authors interpreted these latency differences in the same way—the processes generating the MMR cannot be initiated until the offset of the duration deviant. In our case, this would be at the point where the deviant stimulus offset begins to exceed the duration of the standard stimulus.

#### 4.2. Comparison between FH+ and FH–

When comparing infants with and without family history of LLI a number of differences emerged. First, FH+ infants showed delayed N2\* peaks for frequency discrimination. These differences reflect slower processing in FH+ infants, specifically for fine-grained acoustic features, as significant between group differences are only seen in the short-ISI condition (i.e., 70 ms ISI) in the American studies and in the DEVF stimulus type in the present study. According to the literature, this peak indexes the building-up of neural representations or sensory memory traces of the repeated stimulus (Näätänen et al., 2007). In addition, the N2\* in response to deviant stimuli appears to represent the onset of the discrimination response (Choudhury and Benasich, 2011). The delayed N2\* peak latency in FH+ infants for the frequency deviant may thus reflect slower processing of these features and a delay in the recognition of the frequency change in the majority of FH+ infants. Notably, it does not seem to be associated with general auditory processing abilities, as the latency difference between groups is specific for the frequency change, here interpreted as a fine-grained acoustic feature.

Second, Italian FH+ infants showed reduced amplitude of the positive mismatch response (P3 component). Here, analyses were conducted on the deviant waveforms and significant group differences, reaching statistical significance in the right hemisphere, were shown for both deviant stimuli. Although the reduced amplitude of this positive response replicates the American results as well as other studies conducted on samples of infants at risk for dyslexia, both with non-speech and speech stimuli (e.g., Leppänen et al., 2002; van Zuijen et al., 2012), the hemispheric effect found in our Italian sample is unique. Previous studies reported group differences in the laterality of the MMR. Whereas comparable amplitudes were found for left and right hemispheres in controls, in at-risk groups some studies reported attenuated ERP responses in left rather than right hemisphere for speech (e.g., Leppänen et al., 2002; Leppänen et al., 2011; van Leeuwen et al., 2007; van Herten et al., 2008) and/or fine-grained non-speech stimuli (e.g., Choudhury and Benasich, 2011). Other studies reported atypically enhanced responses in the right hemisphere for speech stimuli involving duration changes (Friedrich et al., 2009; Guttorm et al., 2005). Although the Italian results also show equal left and right amplitudes for the control group, reduced right as compared to left hemisphere amplitude is seen in the at-risk group. Prior studies have shown that cortical activity is either lateralized to the left or recruits both hemispheres when acoustic information is rapidly modulated over time, whereas slow spectral modulations predominantly activate the right hemisphere (e.g., Poeppel et al., 2008; see Telkemeyer et al., 2011 for a NIRS confirmation in 6-month-old infants). A speculative explanation for our results may take into account the specificity of this novel paradigm, where changes in sound duration produce an irregularity of the rhythm of sound presentation. It might be hypothesized that this irregularity in the rhythm of the presentation within our multi-feature oddball paradigm requires a more substantial activation of the right hemisphere, and that is significantly more challenging for the higher-risk FH+ group. However, this novel ERP lateralization result requires further investigation, possibly using analytic strategies that allow age-appropriate source localization and fine-grained time frequency analyses (e.g., Hämäläinen et al., 2011).

We would like to discuss in more detail the specificity of the findings for the two stimulus types that were included in the experiment: fundamental frequency and sound duration. Group differences in the latency of the N2\* peak are specific for the change in fundamental frequency, interpreted as a fine-grained acoustic feature. However, the difference in the amplitude of the P3 is evident for both the frequency and duration deviants, although it is restricted to the right hemisphere for the duration change, and it is broader for the frequency change (as shown by the significant Group × Stimulus Type interaction). In other languages such as German and Finnish, where both consonant and vowel duration represent critical phonemic differences, previous studies have shown clear discrimination of sound duration changes in newborns and 2-month-old infants (Friederici et al., 2002; Leppänen et al., 1999). Sound duration has also been shown to be a critical feature for infants at familial risk for dyslexia and/or language impairment (Friedrich et al., 2004; Leppänen et al., 1999, 2002) in the same language contexts. In all these studies, however, duration has been investigated within a speech context, specifically for changes in vowel duration, as in/ba/ vs. /ba:/ (Friederici et al., 2002; Friedrich et al., 2004; Leppänen et al., 1999) or generated by adding a prolonged silent gap creating the perception of consonant duration change, as in/ata/ vs. /atta/ (Leppänen et al., 2002). In the present study, the change in duration has been investigated in a non-speech context, and this might be why less robust differences are seen between groups for this stimulus type (see for example Jaramillo et al., 2001 for evidence that speech stimuli are more efficiently processed than tones). An alternative explanation concerns

the “functional meaning” that changes in duration represent in Italian with respect to other languages. In Italian, sound duration is an essential cue for a specific phenomenon known as “consonant germination” and it might be hypothesized that discrimination of this acoustic feature will become increasingly more important at later phases of development, whereas in German and Finnish duration changes are more frequent and thus this acoustic feature might be more important even in the earliest phases of development. The strength of this particular hypothesis will become apparent as we follow our current cohort out to later ages and continue data collection using the same paradigm.

#### 4.3. Prediction of the linguistic outcome

Our second goal was to identify the developmental trajectory from early RAP to later language acquisition in an Italian cohort and evaluate whether assessment of early RAP abilities have similar utility in predicting later language and cognitive measures to that demonstrated in previous American and Finnish cohorts (e.g. Choudhury and Benasich, 2011; Guttorm et al., 2005). Our results to date confirm the utility of RAP as a predictor of later measures of expressive language (number of words via parental report) at 20 months, using a reliable measure of linguistic outcome (Lee et al., 2011). In particular, consistent correlational results were found for the latency of the obligatory responses (both P1 and N2) and amplitude of N2\* and P3. These results are in line with previous cross-linguistic findings that found early brain responses to both non-speech (Benasich et al., 2006; Choudhury and Benasich, 2011; van Zuijen et al., 2012) and speech stimuli (Guttorm et al., 2005) to be predictive of linguistic skills at pre-school ages. In concordance with results obtained in earlier group comparisons, most of the correlations involved brain responses in the right hemisphere (see Guttorm et al., 2005 and Friedrich et al., 2009 for atypical right-hemispheric involvement in response to speech stimuli related to early linguistic skills). This result also suggests strong right hemisphere contributions to language acquisition, at least in Italian populations.

When linking together results from group comparison and correlation analyses, an interesting pattern emerged. On one hand, obligatory responses (i.e., latency of P1 and latency of N2) seem to predict the linguistic outcome irrespective of the infants' status of risk. This can be considered a more general effect of processing speed on language ability (see similar results reported in Leppänen et al., 2010). On the other hand, when the familial risk is taken into account, the ERP components related to the discrimination response seem more relevant (N2\* and P3 are disturbed in FH+ infants), suggesting that the aggregation in families of LLI is driven specifically by disturbances in the ability to discriminate sounds rather than in more basic auditory processes. In particular, the latency of the N2\* for DEVF stimuli seems to specifically differentiate between groups, but does not predict later linguistic outcome. For this reason, this specific ERP signature might be considered a neural biomarker that is affected in infants having a first-degree relative with LLI (similar results are reported in Plakas et al., 2013). However, this hypothesis remains speculative until confirmed by studies with genetic sensitive design, as twin studies, where the relative contribution of heritability and shared environment can be precisely estimated for each ERP component. Finally, it is the case that in this study, the amplitude of the P3 is the only ERP signature that both differentiates between groups and predicts to later linguistic outcome. Based on these findings, screening for ERP signatures and in particular the P3, may facilitate earlier identification and remediation of those children who will eventually develop LLI. Our qualitative analyses showed that the risk of having lower expressive vocabulary scores at 20 months of age is more than doubled when an infant is at familial risk for LLI and in addition

presents with lower P3 amplitudes (23% for FH– infants vs. 53% for FH+ infants). These findings resemble those of Guttorm et al. (2010), where the combined information from the familial status and early ERPs measures (enhanced and prolonged responses in the right hemisphere) provided more accurate information about later performance (pre-reading phonological skills) than information based on familial risk status only. Here a note of caution is needed, since our longitudinal data collection is ongoing and at the present time we only show prediction of linguistic outcome at 20 months of age in the subset (70%) examined to date. Thus, clinical diagnostic assessments must wait until the children are older. However, we are prospectively following our current cohort in order to delineate linguistic trajectories at later stages of development using more comprehensive measures of linguistic performance.

Limitations of this study include the fact that language outcomes were assessed solely by parental report. Although the measure of language expression at 20 months (LDS) is highly correlated with concurrent and later measures of language development (Rescorla and Alley, 2001), and considered to be more reliable than direct assessment at this early age, assessment is dependent on the state of the infant/toddler.

## 5. Conclusion

Overall, our results support the general hypothesis that early RAP skills are impaired in infants at familial risk for LLI regardless of native language or previous experience with frequency and duration cues embedded within the language. These results are important in light of the biological unity hypothesis of LLI. Despite the laterality differences observed when comparing Italian and American at-risk infants, deficits in early pre-linguistic auditory processing skills seem to be a universal risk marker for LLI. Moreover, the pattern of cross-linguistic findings will be particularly important as we begin to investigate possible genetic mechanisms underlying the clinical manifestations of LLI.

Such insights also suggest that investigation of core underlying deficits or “intermediate phenotypes” is likely to be a more successful strategy than relying on complex diagnostic phenotypes. In particular, it highlights the advantage of examining early patterns of processing in infants at familial risk of LLI in order to take advantage of their shorter exposure to cultural and environmental confounders. These bellwether results have strong implications for the future implementation of early intervention programs (for a recent example of early intervention see Benasich et al., 2014), which may need to be both specific to Italian populations and early enough in life, to take advantage of the period when infants are still constructing cortical sensory maps, thus facilitating maximal neural plasticity before clinical traits stabilize.

## Acknowledgements

The authors wish to thank the nursing and clinical staff of the Department of Gynecology & Obstetrics of the Manzoni Hospital of Lecco (Prof. Antonio Pellegrino and Dr. Roberta Tironi) and of the Hospital of Desio and Vimercate (branch of Carate Brianza; Prof. Anna Locatelli and Mrs. Rossella Fumagalli). Many thanks to Lara Lanzoni, Giulia Melesi, Chiara Miotti, Giulia Mornati and Vittoria Trezzi, for their help in data collection. Finally, special thanks go to all infants and their parents participating in this study and to two anonymous reviewers for their thoughtful suggestions. The research was supported by Grant RC2013 from the Italian Ministry of Health, by Fondazione della Provincia di Lecco Onlus/Rotary Club (Rotary Club Lecco, Rotary Club Le Grigne, Rotary Club Manzoni) and the l'Oreal-UNESCO Italy for Women in Science Fellowship (CC).



- Front. Hum. Neurosci. 8 (313), 1–16, <http://dx.doi.org/10.3389/fnhum.2014.00313>.
- Luyster, R.J., Seery, A., Talbott, M.R., Tager-Flusberg, H., 2011. Identifying early-risk markers and developmental trajectories for language impairment in neurodevelopmental disorders. *Dev. Disabil. Res. Rev.* 17 (2), 151–159, <http://dx.doi.org/10.1002/ddrr.1109>.
- Malenfant, N., Grondin, S., Boivin, M., Forget-Dubois, N., Robaey, P., Dionne, G., 2012. Contribution of temporal processing skills to reading comprehension in 8-year-olds: evidence for a mediation effect of phonological awareness. *Child Dev.* 83 (4), 1332–1346, <http://dx.doi.org/10.1111/j.1467-8624.2012.01777.x>.
- Molfese, D.L., 2000. Predicting dyslexia at 8 years of age using neonatal brain responses. *Brain. Lang.* 72 (3), 238–245, <http://dx.doi.org/10.1006/brln.2000.2287>.
- Morr, M.L., Shafer, V.L., Kreuzer, J.A., Kurtzberg, D., 2002. Maturation of mismatch negativity in typically developing infants and preschool children. *Ear. Hear.* 23 (2), 118–136, <http://dx.doi.org/10.1097/00003446-200204000-00005>.
- Näätänen, R., Paavilainen, P., Rinne, T., Alho, K., 2007. The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clin. Neurophysiol.* 118 (12), 2544–2590, <http://dx.doi.org/10.1016/j.clinph.2007.04.026>.
- Nenonen, S., Shestakova, A., Huotilainen, M., Näätänen, R., 2003. Linguistic relevance of duration within the native language determines the accuracy of speech-sound duration processing. *Cogn. Brain Res.* 16 (3), 492–495.
- Paulesu, E., Demonet, J.F., Fazio, F., McCrory, E., Chanoine, V., Brunswick, N., Cappa, S.F., Cossu, G., Habib, M., Frith, C.D., Frith, U., 2001. Dyslexia: cultural diversity and biological unity. *Science* 291 (5511), 2165–2167, <http://dx.doi.org/10.1126/science.1057179>.
- Plakas, A., van Zuijlen, T., van Leeuwen, T., Thomson, J.M., van der Leij, A., 2013. Impaired non-speech auditory processing at a pre-reading age is a risk-factor for dyslexia but not a predictor: an ERP study. *Cortex* 49, 1034–1045, <http://dx.doi.org/10.1016/j.cortex.2012.02.013>.
- Plomin, R., Kovas, Y., 2005. Generalist genes and learning disabilities. *Psychol. Bull.* 131 (4), 592–617, <http://dx.doi.org/10.1037/0033-2909.131.4.592>.
- Poeppel, D., Idsardi, W.J., van Wassenhove, V., 2008. Speech perception at the interface of neurobiology and linguistics. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363 (1493), 1071–1086, <http://dx.doi.org/10.1098/rstb.2007.2160>.
- Ramus, F., Rosen, S., Dakin, S.C., Day, B.L., Castellote, J.M., White, S., Frith, U., 2003. Theories of developmental dyslexia: insights from a multiple case study of dyslexic adults. *Brain* 126, 841–865, <http://dx.doi.org/10.1093/brain/awg076>.
- Rescorla, L., Alley, A., 2001. Validation of the language development survey (LDS): a parent report tool for identifying language delay in toddlers. *J. Speech Lang. Hear. Res.* 44 (2), 434–445, [http://dx.doi.org/10.1044/1092-4388\(2001\)035](http://dx.doi.org/10.1044/1092-4388(2001)035).
- Rescorla, L., Frigerio, A., Sali, M.E., Spataro, P., Longobardi, E., 2014. Typical and delayed lexical development in Italian. *J. Speech Lang. Hear. Res.* 57 (5), 1792–1803, <http://dx.doi.org/10.1044/2014.JSLHR-L-13-0242>.
- Rescorla, L., Ratner, N.B., Jusczyk, P., Jusczyk, A.M., 2005. Concurrent validity of the language development survey: associations with the MacArthur-Bates communicative development inventories: words and sentences. *Am. J. Speech Lang. Pathol.* 14 (2), 156–163, [http://dx.doi.org/10.1044/1058-0360\(2005\)016](http://dx.doi.org/10.1044/1058-0360(2005)016).
- Rivera-Gaxiola, M., Silva-Pereyra, J., Kuhl, P.K., 2005. Brain potentials to native and non-native speech contrasts in 7- and 11-month-old American infants. *Dev. Sci.* 8 (2), 162–172, <http://dx.doi.org/10.1111/j.1467-7687.2005.00403.x>.
- Saffran, J.R., Aslin, R.N., Newport, E.L., 1996. Statistical learning by 8-month-old infants. *Science* 274 (5294), 1926–1928, <http://dx.doi.org/10.1126/science.274.5294.1926>.
- Sartori, G., Job, R., Tressoldi, P. E., 1995. Batteria per la valutazione della dislessia e della disortografia evolutiva in età evolutiva. OS, Firenze.
- Schulte-Körne, G., Bruder, J., 2010. Clinical neurophysiology of visual and auditory processing in dyslexia: a review. *Clin. Neurophysiol.* 121 (11), 1794–1809, <http://dx.doi.org/10.1016/j.clinph.2010.04.028>.
- Tallal, P., Piercy, M., 1973. Defects of non-verbal auditory perception in children with developmental aphasia. *Nature* 241 (5390), 468–469, <http://dx.doi.org/10.1038/241468a0>.
- Telkemeyer, S., Rossi, S., Nierhaus, T., Steinbrink, J., Obrig, H., Wartenburger, I., 2011. Acoustic processing of temporally modulated sounds in infants: evidence from a combined near-infrared spectroscopy and EEG study. *Front. Psychol.* 2 (62), 1–14, <http://dx.doi.org/10.3389/fpsyg.2011.00062>.
- Tervaniemi, M., Jacobsen, T., Röttger, S., Kujala, T., Widmann, A., Vainio, M., Näätänen, R., Schröger, E., 2006. Selective tuning of cortical sound-feature processing by language experience. *Eur. J. Neurosci.* 23, 2538–2541, <http://dx.doi.org/10.1111/j.1460-9568.2006.04752.x>.
- van der Leij, A., van Bergen, E., van Zuijlen, T., de Jong, P., Maurits, N., Maassen, B., 2013. Precursors of developmental dyslexia: an overview of the longitudinal Dutch dyslexia programme study. *Dyslexia* 19 (4), 191–213, <http://dx.doi.org/10.1002/dys.1463>.
- van Herten, M., Pasman, J., van Leeuwen, T.H., Been, P.H., van der Leij, A., Zwarts, F., Maassen, B., 2008. Differences in AERP responses and atypical hemispheric specialization in 17-month-old children at risk of dyslexia. *Brain. Res.* 1201, 100–105, <http://dx.doi.org/10.1016/j.brainres.2008.01.060>.
- van Leeuwen, T., Been, P., van Herten, M., Zwarts, F., Maassen, B., van der Leij, A., 2007. Cortical categorization failure in 2-month-old infants at risk for dyslexia. *Neuroreport* 18 (9), 857–861, <http://dx.doi.org/10.1097/WNR.0b013e3280c1e2bf>.
- van Zuijlen, T.L., Plakas, A., Maassen, B.A.M., Been, P., Maurits, N.M., Krikhaar, E., van Driel, J., van der Leij, A., 2012. Temporal auditory processing at 17 months of age is associated with preliterate language comprehension and later word reading fluency: an ERP study. *Neurosci. Lett.* 528 (1), 31–35, <http://dx.doi.org/10.1016/j.neulet.2012.08.058>.
- Wunderlich, J.L., Cone-Wesson, B.K., Shepherd, R., 2006. Maturation of the cortical auditory evoked potential in infants and young children. *Hear. Res.* 212 (1–2), 185–202, <http://dx.doi.org/10.1016/j.heares.2005.11.010>.