UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Fast phonetic learning occurs already in 2-to-3-month old infants. An ERP study

Wanrooij, K.; Boersma, P.; van Zuijen, T.L.

DOI
10.3389/fpsyg.2014.00077
Publication date
2014
Document Version
Final published version
Published in
Frontiers in Psychology

Link to publication

Citation for published version (APA):
Wanrooij, K., Boersma, P., & van Zuijen, T. L. (2014). Fast phonetic learning occurs already in 2-to-3-month old infants. An ERP study. *Frontiers in Psychology*, *5*, Article 77. https://doi.org/10.3389/fpsyg.2014.00077

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

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:13 Nov 2026

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 25 February 2014
doi: 10.3389/fpsyg.2014.00077

Fast phonetic learning occurs already in 2-to-3-month old infants: an ERP study

Karin Wanrooij*[1]*, Paul Boersma[1] and Titia L. van Zuijen[2]

[1] *Amsterdam Center for Language and Communication, University of Amsterdam, Amsterdam, Netherlands*
[2] *Department of Child Development and Education, University of Amsterdam, Amsterdam, Netherlands*

Edited by:
Janet F. Werker, University of British Columbia, Canada

Reviewed by:
Alejandrina Cristia, Max Planck Institute for Psycholinguistics, Netherlands
Eino Partanen, University of Helsinki, Finland

***Correspondence:**
Karin Wanrooij, Amsterdam Center for Language and Communication, University of Amsterdam, Spuistraat 210, 1012 VT Amsterdam, Netherlands
e-mail: karin.wanrooij@uva.nl

An important mechanism for learning speech sounds in the first year of life is "distributional learning," i.e., learning by simply listening to the frequency distributions of the speech sounds in the environment. In the lab, fast distributional learning has been reported for infants in the second half of the first year; the present study examined whether it can also be demonstrated at a much younger age, long before the onset of language-specific speech perception (which roughly emerges between 6 and 12 months). To investigate this, Dutch infants aged 2 to 3 months were presented with either a unimodal or a bimodal vowel distribution based on the English /æ/∼/ɛ/ contrast, for only 12 minutes. Subsequently, mismatch responses (MMRs) were measured in an oddball paradigm, where one half of the infants in each group heard a representative [æ] as the standard and a representative [ɛ] as the deviant, and the other half heard the same reversed. The results (from the combined MMRs during wakefulness and active sleep) disclosed a larger MMR, implying better discrimination of [æ] and [ɛ], for bimodally than unimodally trained infants, thus extending an effect of distributional training found in previous behavioral research to a much younger age when speech perception is still universal rather than language-specific, and to a new method (using event-related potentials). Moreover, the analysis revealed a robust interaction between the distribution (unimodal vs. bimodal) and the identity of the standard stimulus ([æ] vs. [ɛ]), which provides evidence for an interplay between a perceptual asymmetry and distributional learning. The outcomes show that distributional learning can affect vowel perception already in the first months of life.

Keywords: distributional learning, infant MMR (mismatch response), perceptual asymmetry, language acquisition, category learning, ERP, speech perception

INTRODUCTION

Distributional learning, i.e., learning by simply being exposed to the frequency distributions of stimuli in the environment, may be one of the mechanisms by which infants start to acquire the phonemes of their language (Lacerda, 1995; Guenther and Gjaja, 1996). *Fast* distributional learning of speech sounds after just a few minutes of exposure in the lab has been observed in infants in the second half of the first year (e.g., Maye et al., 2008). This study investigates whether such fast distributional learning can also take place in very young infants, i.e., 2-to-3-month olds. This is relevant if we want to establish that the distributional learning mechanism is in place early enough to be able to contribute to the transition from universal to language-specific speech perception, which becomes apparent in infants' speech sound discrimination from around 6 months of age (e.g., Werker and Tees, 1984/2002; Polka and Werker, 1994), or perhaps even from 4 months (Yeung et al., 2013).

In the first year of life, infants' speech sound perception has been observed to change from universal to language-specific. Specifically, in the course of this transition discrimination performance is enhanced for native speech sound contrasts (Cheour et al., 1998b; Kuhl et al., 2006; Tsao et al., 2006), and reduced for non-native contrasts that are irrelevant in the native language (Werker and Tees, 1984/2002; Kuhl et al., 1992; Tsushima et al., 1994; Polka and Werker, 1994; Best et al., 1995; Bosch and Sebastián-Gallés, 2003; Kuhl et al., 2006; Tsao et al., 2006). In general, language-specific speech sound discrimination emerges between 4 and 6 months for tones (i.e., in tonal languages; Cheng et al., 2013; Yeung et al., 2013), around 6 months for vowels (Kuhl et al., 1992; Polka and Werker, 1994; Bosch and Sebastián-Gallés, 2003), and between 8 and 12 months for consonants (Werker and Tees, 1984/2002; Tsushima et al., 1994; Best et al., 1995; Kuhl et al., 2006; Tsao et al., 2006), although language-specific discrimination of difficult contrasts may develop later (e.g., Cheour et al., 1998b; Polka et al., 2001; Sundara et al., 2006).

One of the mechanisms that has been hypothesized to contribute to the emergence of language-specific speech perception is distributional learning (Lacerda, 1995; Guenther and Gjaja, 1996). The existence of this mechanism has indeed been supported by observations in the lab. In particular, fast distributional learning has been demonstrated most reliably in 8-month olds by Maye et al. (2008; $p < 0.001$), and (nearly) significantly in 6-to-8-month olds by Maye et al. (2002; $p = 0.063$), in 10-to-11-month olds by Yoshida et al. (2010; $p = 0.036$ for one of the experiments), and in 11-month olds by Capel et al. (2011; $p = 0.053$), although null results were found in 10-to-11-month olds by Yoshida et al.

(2010; for two experiments) and ambiguous results were found in 5-month olds by Cristià et al. (2011; $p > 0.16$ for the main effect, but $p = 0.007$ for an interaction effect).

If distributional learning indeed contributes to the acquisition of language-specific perception, and discriminational evidence for the latter starts being observed from 4 or 6 months on, fast distributional learning can be expected to be detectable in even younger infants. This expectation is supported by neuroscientific research. Cortical layers involved in top-down processing (e.g., Kral and Eggermont, 2007) become anatomically available in humans from around 4 to 5 months of age (Moore and Guan, 2001; Moore, 2002; Moore and Linthicum, 2007), which suggests that speech perception before 4 months relies mainly on bottom-up processing. The distributional learning mechanism, which supposedly does not require top-down processing (Guenther and Gjaja, 1996), should therefore at this early age be relatively unimpeded by learning mechanisms that require top-down influence from higher-level (e.g., lexical) representations.

We therefore performed a fast distributional learning experiment with infants aged 2 to 3 months. Specifically, we presented Dutch infants of this age with speech sounds from an acoustic continuum encompassing the British-English vowel contrast /æ/∼/ɛ/; this is a contrast that does not exist in Dutch, and which Dutch adults find difficult to master (e.g., Schouten, 1975; Weber and Cutler, 2004; Broersma, 2005; Escudero et al., 2008). These vowels differ in their first formant (F1), as illustrated in **Figure 1**, where the F1 values are given in ERB (Equivalent Rectangular Bandwidth; see section Stimuli for details). In our experiment, one half of the infants were exposed to a *unimodal distribution* (**Figure 1**, gray), i.e., to a large number of different vowel tokens whose F1 values center around 11.47 ERB, which is phonetically halfway between English [ɛ] and [æ], and the other half of the infants were exposed to a *bimodal distribution* (**Figure 1**, black), i.e., to a large number of vowel tokens whose F1 values center around 10.44 and 12.50 ERB, which are F1 values typical of English [ɛ] and [æ], respectively. The bimodal distribution thus suggests the existence of a contrast between /æ/ and /ɛ/ (as would be appropriate for learners of English), while the unimodal distribution does not suggest a contrast between the two vowels (as would be appropriate for learners of Dutch). Immediately after the training we tested how well the infants discriminated an open variant of English [ɛ], i.e., a vowel with an F1 of 10.78 ERB, and a closed variant of English [æ], i.e., a vowel with an F1 of 12.16 ERB, both visible in **Figure 1**. If distributional learning occurred, bimodally trained infants should discriminate them better than unimodally trained infants.

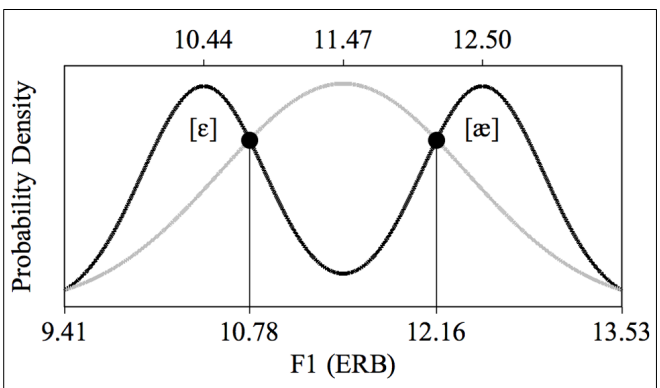

FIGURE 1 | Unimodal (gray curve) and bimodal (black curve) training distributions of the first vowel formant (F1). The values of the test stimuli lie at the intersections of the two distributions.

Discrimination ability after training had to be measured with a method appropriate for young infants. All previous research on infant or adult distributional learning employed behavioral measures, which for infants always meant looking time. Since suitable behavioral responses are difficult to obtain from 2-to-3-month olds, we instead measured an automatic brain response, namely the *mismatch response* (MMR; e.g., Näätänen et al., 1978). In contrast to behavioral measurements, which require the infant's cooperation and attention (Cheour et al., 2000, p. 6), the MMR is elicited even in the absence of voluntary attention to the stimuli (e.g., Schröger, 1997; Näätänen and Winkler, 1999), and can be measured even when the infant is asleep (Friederici et al., 2002; Martynova et al., 2003). The MMR has been shown to reflect behavioral discrimination in adults (for a review, see Näätänen et al., 2007), and has been used successfully before to demonstrate vowel discrimination in infants of 3 months and younger (e.g., Cheour-Luhtanen et al., 1995; Cheour et al., 1998a; Martynova et al., 2003; Kujala et al., 2004; Shafer et al., 2011; Partanen et al., 2013). The MMR can be elicited in an oddball paradigm (e.g., Näätänen, 1992), where a series of "standard" stimuli (e.g., [ɛ] tokens) is interspersed infrequently with "deviant" stimuli (e.g., [æ] tokens). If the auditory perception system detects that deviants differ from standards, it will process the two kinds of stimuli in different ways, which can be reflected in the event-related potentials (ERPs). The MMR can be computed as the difference between the ERP elicited by the deviants and the ERP elicited by the standards.

When measuring MMRs to speech sounds in an oddball paradigm, it can make a difference whether one or the other stimulus of a pair is chosen as a standard. Possible asymmetries in participants' perception can exist, which can make discrimination easier if one particular stimulus (e.g., [æ]) is the standard than if the contrasting stimulus (e.g., [ɛ]) is the standard. Perceptual biases have been reported for several speech sounds (Pisoni, 1977; Aslin and Pisoni, 1980; Polka and Bohn, 1996, 2003) and seem especially strong in young infants (Pons et al., 2012). For vowels one relevant perceptual bias is a peripherality-related asymmetry: when hearing a more peripheral vowel after a more central vowel (i.e., in a two-dimensional acoustic space defined by the first and second vowel formants) discrimination is easier than when hearing the same vowels in the opposite order (e.g., Polka and Bohn, 1996, 2003; Pons et al., 2012). This would predict that in our oddball paradigm discrimination may be easier if [ɛ] is the standard stimulus than if [æ] is the standard. Further, the "natural referent vowel" hypothesis (Polka and Bohn, 2011) predicts that this perceptual bias will vanish or grow fainter for native contrasts and will remain or grow stronger for non-native contrasts. This would predict that if fast distributional training already leads to some sort of vowel category formation, unimodally trained infants, for whom the contrast /æ/∼/ɛ/ is new ("non-native"), will show a perceptual asymmetry, whereas the bias will not be clear in bimodally trained

infants, for whom the contrast is experienced during training ("native"). Other perceptual biases can be expected on the basis of hypotheses involving underspecification (Lahiri and Reetz, 2010), according to which a featurally underspecified phoneme will mismatch with a preceding specified phoneme, but the reverse order will not lead to a mismatch. This would predict that if [æ] is specified for the feature [low] and [ɛ] is not, discrimination may be easier if [æ] is the standard stimulus than if [ɛ] is the standard. To accommodate the main and interaction effects of any perceptual biases, we counterbalanced the identity of the standard ([æ] or [ɛ]) across the infants and included it as a factor in the analysis.

In sum, the aim of the current study was to investigate whether 2-to-3-month old infants already show fast distributional learning, by training Dutch infants of this age on either a unimodal or a bimodal distribution of the English vowel contrast /æ/∼/ɛ/, and then testing in an ERP oddball paradigm how well they discriminate [æ] from [ɛ]. If the distributional learning mechanism exists, it is expected that bimodally trained infants, who hear a distribution that suggests the existence of a contrast between /æ/ and /ɛ/, discriminate [æ] and [ɛ] better, and thus have a *larger MMR amplitude*, than unimodally trained infants, who hear a distribution that does not suggest a contrast between /æ/ and /ɛ/.

MATERIALS AND METHODS

PARTICIPANTS

The 32 infants (11 girls) accepted for the study met the following criteria. The language spoken at home had to be Dutch only. The infant had to be healthy and had to have passed the Dutch otoacoustic emissions test for newborns. Birth weight had to be normal (each infant weighed over 2500 g). The Apgar score had to be 8 or higher 10 minutes after birth. The gestational age at birth had to be between 37 and 42 weeks, and the post-natal age from birth to time of testing between 8 and 12 weeks. Finally, we excluded infants born with complications, but accepted infants delivered by Caesarean section. The study protocol was approved by the Ethical Committee of the Faculty of Social and Behavioral Sciences at the University of Amsterdam. Parents signed informed consent forms.

DESIGN

All infants listened to a training distribution and performed a subsequent discrimination test. During the training, half of the infants heard a bimodal distribution, with peaks around [æ] and around [ɛ], and the other half a unimodal distribution, with a single broad peak between [æ] and [ɛ]. During the test, half of the infants in each distributional training group listened to standard [æ] and deviant [ɛ], and the other half to standard [ɛ] and deviant [æ]. Thus, based on Distribution Type (unimodal vs. bimodal) and Standard Vowel ([æ] vs. [ɛ]) the 32 infants were assigned to four "groups," namely Unimodal [æ], Unimodal [ɛ], Bimodal [æ], and Bimodal [ɛ], each consisting of eight infants. Apart from balancing the sexes, assignment to the groups was random.

After separating the data into non-quiet sleep (non-QS) and quiet sleep (QS) data (section Coding Sleep Stages) and applying a criterion for a sufficient number of valid responses (section ERP Recording and Analysis), we could include the non-QS data of 22 infants in the non-QS dataset, and the QS data of 21 infants in the QS dataset (12 infants contributed to both datasets, 19 to one dataset, and one to no dataset). In the non-QS dataset the number of contributing infants was five in Unimodal [æ], six in Unimodal [ɛ], six in Bimodal [æ], and five in Bimodal [ɛ]. In the QS dataset the number of contributing infants was six in Unimodal [æ], four in Unimodal [ɛ], five in Bimodal [æ], and six in Bimodal [ɛ].

To sum up, the experimental design for measuring the effect of distributional training had Distribution Type (unimodal vs. bimodal) and Standard Vowel ([æ] vs. [ɛ]) as between-subject factors, and the MMR amplitude as the dependent variable, to be determined separately for the QS and the non-QS dataset.

STIMULI

Test and training stimuli were made with the Klatt synthesizer in the computer program Praat (Boersma and Weenink, 2010) and varied only in the values for the first and second formants, F1 and F2 (see sections In the Training and In the Test). The duration of each stimulus was kept at 100 ms (e.g., Cheour-Luhtanen et al., 1995; Cheour et al., 1998a, 2002b) including rise and fall times of 5 ms. The fundamental frequency contour fell from 150 to 112.5 Hz, which represents a male voice (e.g., Cheour-Luhtanen et al., 1995; Cheour et al., 1998a, 2002b; Martynova et al., 2003). The source signal was filtered with eight additional formants (F3 through F10). The values for F3, F4, and F5, which were 2400, 3400, and 4050 Hz respectively, were extracted from American-English vowels representing /æ/ and /ɛ/ in the TIMIT database (Lamel et al., 1986), while those for F6 through F10 were calculated as the previous formant plus 1000 Hz (e.g., F6 = F5 + 1000 Hz). Similarly, the bandwidth values for the first four bandwidths, which were 80, 160, 360, and 530 Hz, respectively, were based on the TIMIT database, while an additional six bandwidths were calculated as the corresponding formant divided by 8.5 (e.g., bandwidth 5 = F5/8.5). Each stimulus was made equally loud, to avoid possible confounds in the ERPs based on intensity differences (Näätänen et al., 1989; Sokolov et al., 2002). The stimuli were played (during training and test) at around 70 dB SPL, measured at about one meter from the two loudspeakers, where the infant was lying.

In the training

The unimodal and bimodal training distributions were created in the manner reported by Wanrooij and Boersma (2013). In contrast with previous research, which typically employed only eight different stimulus values, each of which was repeated multiple times during training, this method uses more ecologically valid *continuous* training distributions, where all presented stimuli are acoustically different. Each of the two distributions thus consisted of 900 unique vowels and had an identical range of F1 and F2 values: 9.41 to 13.53 ERB for F1 and 21.05 to 18.31 ERB for F2 (see also **Figure 1**). These ranges were based on values for F1 and F2 as reported by Hawkins and Midgley (2005). Specifically, we took the reported F1 and F2 values of /æ/ and /ɛ/, each pronounced four times by five male speakers of British English in the age group 35–40 years, and converted the hertz values to ERB.

Hawkins and Midgley's mean F1 and F2 were 12.51 ERB and 18.94 ERB, respectively for /æ/, and 10.43 ERB and 20.42 ERB for /ɛ/. Because in the current study the stimuli were produced by one synthetic speaker, a single-speaker standard deviation, for F1 and F2 separately, was calculated as the mean of the five speakers' standard deviations for the vowel /ɛ/. The standard deviations were 0.51 ERB for F1 and 0.32 ERB for F2. The edges of the F1 and F2 ranges, mentioned above, were determined to lie two standard deviations from the mean F1 and F2 values of the vowels; for instance, the lower edge of the F1 continuum lay at $10.43 - 2 \times 0.51 = 9.41$ ERB. Note that in going from /ɛ/ to /æ/ the F1 rises, while F2 declines.

The shape of the distributions was defined in accordance with earlier distributional learning studies in that the ratio of the least to most frequent stimuli was about 1 to 4 (e.g., Maye et al., 2002, 2008). As illustrated in **Figure 1**, the unimodal mean lay exactly in the middle of the range of F1 (or F2) values and precisely in between the two bimodal means, which lay at 25 and 75% of the range, for both F1 and F2. This led to the mean F1 and F2 values listed in **Table 1**, which are quite close to those reported for /æ/ and /ɛ/ by Hawkins and Midgley (2005; see above in this section). The unimodal and bimodal distributions consisted of one and two Gaussian peaks, respectively, with standard deviations equal to 22 and 11% of the range, respectively. On the basis of these distributions, the F1 and F2 values for the 1800 training vowels were determined by a procedure described by Wanrooij and Boersma (2013), which approximates the intended probability densities of **Figure 1** optimally. The order of presentation of the 900 stimuli in the training was randomized separately for each infant. The inter-stimulus interval (the silent interval between the end of a stimulus token and the start of the next token) was 707 ms.

In the test

In the test phase, infants were presented with two different stimuli, i.e., a standard and a deviant, repeated at most 2200 and 300 times respectively, depending on the infant (see section Procedure). Thus, deviants were presented at a rate of 12%. The F1 and F2 values of the test stimuli (**Table 1**) were determined by computing the intersections (circles in **Figure 1**) between the unimodal and bimodal distributions. In this way, the two groups of listeners came to the test phase with equal prior exposure (during training) to sounds in the region of the test stimuli, so that any difference between the groups observed in the test could not be attributed to differences in familiarity with the test stimuli. As during training the inter-stimulus interval in the test was 707 ms. In the test, minimally three standards (10 at the start of the test) appeared before each deviant. Apart from this constraint, the presentation of standards and deviants was randomized separately for each infant.

Table 1 | F1 and F2 values (in ERB): means in the unimodal and bimodal training distributions, and values of the two test stimuli.

	Bimodal /ɛ/	Test stimulus 1	Unimodal	Test stimulus 2	Bimodal /æ/
F1	10.44	10.78	11.47	12.16	12.50
F2	20.37	20.14	19.68	19.22	18.99

PROCEDURE

Before training, the EEG cap with electrodes was placed on the infant's head. During training and testing, infants were lying on the caregiver's lap or in an infant seat beside the caregiver, in a sound-shielded room. Caregivers could watch a silent movie. Researchers in the adjacent room could hear caregiver and infant via loudspeakers, and observe them through a window. Researcher and caregiver did not know and could not consciously detect whether the distribution that was played during the training was unimodal or bimodal. The infant's behavior was monitored and documented. Notes on behavior included the documentation of open or closed eyes, movement, fussiness, and pauses. Caregivers were asked not to interact with the infant, unless necessary to keep the infant quiet. In this case, recording was paused or (if it happened in the last minutes of the test) stopped. Excluding pauses, the training always lasted 12.1 minutes (900 training stimuli) and the test lasted between 29.7 and 33.6 minutes (between 2208 and 2500 test stimuli).

CODING SLEEP STAGES

A factor that has to be considered when measuring MMRs is that during the relatively long experimental duration (viz., in the current experiment over 30 minutes, as compared to less than 10 minutes in behavioral distributional learning experiments) young infants tend to fall asleep (see also e.g., Friederici et al., 2002; He et al., 2009). It was therefore important to take a possible influence of sleep stages on MMR measurements into account. Infant sleep stages are usually divided into quiet sleep (QS), active sleep (AS), and wakefulness. Although some studies have not found any differences in neonates' MMR amplitudes between different sleep stages (e.g., Martynova et al., 2003), there are two arguments to analyze data obtained in QS separately from data obtained during wakefulness for 2-to-3-month olds. First, for 2-month olds Friedrich et al. (2004) report a significantly larger positive MMR in QS than during wakefulness, as well as a preceding small negative MMR in wakefulness that was absent in QS. Second, sleep stages and the related EEG-patterns develop quickly into adult-like patterns already in the first 3 months of life (e.g., Crowell et al., 1982; Kahn et al., 1996; Graven and Browne, 2008), and the adult MMR during wakefulness differs from that during sleep, particularly during the successor of QS, non-rapid eye movement (NREM) sleep, where the response tends to disappear (e.g., Loewy et al., 1996; Loewy et al., 2000). In sum, there is at least some evidence that for 2-to-3-month olds the MMR in QS is different from that during wakefulness.

Sleep stages for each infant were determined on the basis of the infant's behavior and the EEG. Stages in the EEG were coded in accordance with the AASM manual (Iber et al., 2007) and, because the manual's age granularity is not precise enough to deduct recommendations for 2-to-3-month olds specifically, specifications for approximately the same age group from Crowell et al. (1982) and Niedermeyer (2005). Specifically, the stage was coded as "QS" when the infant's eyes were closed and the EEG contained frequent

spindles (i.e., more or less sinusoidal waves of 12 to 14 Hz, clearly distinguishable from background activity, and lasting at least 0.5 s; see also Rodenbeck et al., 2007) or apparent slow waves (with or without spindles) coming after parts with abundant spindling. The stage was coded as "AS" when the infant's eyes were closed and the EEG featured transient muscle movements and low-amplitude mixed frequency activity. Finally, the stage was coded as "awake" when the eyes were open. When unequivocal identification was not possible (i.e., the eyes were closed but the EEG did not suggest QS or AS), the state was coded as "indeterminate sleep" (IS). A change of stage was not coded if the relevant changes in EEG and behavior lasted for less than 30 s (Iber et al., 2007).

It turned out that none of the infants stayed awake during recording. On average, they spent 13% of test time awake, 47% in QS, 1% in AS and 39% in IS. There were no significant differences in the time spent in each sleep stage between the four groups (four independent-samples Kruskal–Wallis tests, one for each sleep state, all p-values > 0.74).

For all subsequent analysis, we combined the three non-QS sleep stages (AS, IS, wakefulness) and labeled them together as "non-QS" (cf. Weber et al., 2004). As for AS, only three infants were in this stage for a short while (accounting for less than 2% of test time in any group), which is not surprising in the light of the rare AS onsets at 3 months of age and the relatively late expected start of AS after sleep onset as compared to the total test duration (Ellingson and Peters, 1980; Crowell et al., 1982); moreover, no reliable differences have been reported between MMRs during AS and MMRs during wakefulness in newborns (e.g., Cheour et al., 1998a; Kushnerenko, 2003). As for IS, we suspected that the infant was either well awake or drowsy, even though the eyes were closed, because the EEG in IS looked similar to that during wakefulness and did not contain any visual sign of QS. After combining the three non-QS variants, the sleep stages ended up being nearly equally divided between QS (47% of the time) and non-QS (53%).

ERP RECORDING AND ANALYSIS

The EEG was recorded with a 32-channel Biosemi Active Two system (Biosemi Instrumentation BV, Amsterdam, The Netherlands) at a sampling rate of 8 kHz. Beside the 32 electrodes in the cap, two external electrodes were placed on the mastoids. After recording, the EEG was downsampled to 512 Hz (with Biosemi Decimator 86). Subsequent analysis was done in the computer program Praat (Boersma and Weenink, 2010). First, the EEG was tagged for sleep stages (see section Coding Sleep Stages). Then the EEG in each of the 32 channels was referenced to the mastoids (i.e., the average of the two mastoid channels was subtracted from each channel), "detrended" (i.e., a line was subtracted so that beginning and end of the channel signal were zero) and filtered (Hann-shaped frequency-domain, i.e., zero-phase, filter: pass-band 1–25 Hz, low width 0.5 – high width 12.5 Hz).

The subsequent analysis was done for QS and non-QS data separately, as follows. The EEG was segmented into epochs (32-channel ERP waveforms) of 760 ms duration (from 110 ms before to 650 ms after stimulus onset), for standard and deviant stimuli separately. For each epoch, a baseline correction was performed in each channel by subtracting from each (1-channel) ERP waveform the mean of the waveform in the 110 ms before stimulus onset. If after this an epoch (i.e., a 32-channel ERP waveform) still contained a peak below –150 μV or above +150 μV in one or more channels, the whole epoch was deemed invalid and rejected from further analysis. If after this fewer than 75 deviant epochs remained, the infant was rejected from the dataset for the relevant sleep stage. For each remaining infant, the standard and deviant responses were averaged separately, so as to obtain a mean standard ERP and a mean deviant ERP for each electrode. The infant's 32-channel MMR waveform was obtained by subtracting the standard ERP from the deviant ERP.

MMR ANALYSIS

In order to be able to submit the MMR measurements to statistical analysis, each infant's MMR waveform was reduced to a small set of MMR amplitude values (see below in this section). To achieve this reduction, it was necessary to decide what electrodes and what time window(s) to include in the analysis. The literature that uses infant MMR analysis varies in these decisions and, relatedly, also in the reported results on where on the scalp the MMR was found and when the response occurred (see below in this section). In addition, the literature reports different polarities for the infant MMR (see below in this section). Thus, whereas the adult MMR is invariably a negative deflection (hence usually called a mismatch negativity, or MMN) that usually occurs between 150 and 250 ms after change onset, and is strongest at frontocentral electrodes (when the mastoids or the nose is used as a reference; for a review, see Näätänen et al., 2007), the infant MMR is much less defined in terms of *what* its polarity is, and *when* it occurs *where* on the scalp. We now explain our decisions on how these three aspects of the MMR waveform enter in our analysis.

As for the polarity of the infant MMR, it is sometimes reported as negative (e.g., Cheour-Luhtanen et al., 1995; Cheour et al., 1998b), sometimes as positive (e.g., Dehaene-Lambertz and Baillet, 1998; Dehaene-Lambertz, 2000; Carral et al., 2005), and sometimes as both negative and positive (e.g., Morr et al., 2002; Friederici et al., 2002; Friedrich et al., 2004). Regarding the variation in observed MMR polarities for infants across studies, we include both negative and positive values of individual infant's MMR amplitudes in our analysis.

As for the location of interest on the scalp, some previous research selected only frontal electrodes (e.g., Morr et al., 2002) or frontal and central electrodes (e.g., Cheour et al., 1998b; Morr et al., 2002). When more posterior electrodes were included a significant infant MMR was sometimes reported only at frontal or frontocentral electrodes (Cheour-Luhtanen et al., 1995; Friederici et al., 2002; Friedrich et al., 2004), and sometimes also in more posterior areas (Cheour et al., 2002a; Van Leeuwen et al., 2008; He et al., 2009). As there is therefore some evidence that the infant MMR can be measured beyond frontocentral electrodes, our analysis includes not only six frontocentral electrodes (Fz, F3, F4, Cz, C3, C4), but also two temporal electrodes (T7 and T8); parietal and occipital electrodes were not included, because some infants had been lying on these electrodes. Following Cheour et al. (1998b), Morr et al. (2002) and Friedrich et al. (2004) we include the eight electrodes in the main analysis as a within-subject factor.

As for the chosen time window, the previous literature on infant MMR used various windows for vowels (e.g., 0–500 ms after stimulus onset in Cheour-Luhtanen et al., 1995; 200–500 ms in Cheour et al., 1998a) and various windows for 2- or 3-month olds (e.g., 0–1000 ms in Friederici et al., 2002; 200–600 ms in Friedrich et al., 2004; 100–450 ms and 550–900 ms in He et al., 2009). The only publication on vowels with infants in our age range (3-month olds: Cheour et al., 2002b) used a window from 150 to 400 ms. Regarding the reported variation, and because control of the Type I error rate dictates that analysis windows be chosen before the ERP results are seen, we had to choose in advance a window that includes at least the possible times at which the MMR can occur, namely a window running from 100 to 500 ms. In order to submit this window to an analysis of variance (ANOVA), we divide it into eight consecutive time bins of 50 ms each (Cheour-Luhtanen et al., 1995; Morr et al., 2002; Friedrich et al., 2004; He et al., 2009), and compute the average amplitude of the difference waveform in each bin as our measurement variable. To conclude, each infant's MMR waveform is reduced to only 64 (8 time bins × 8 channels) MMR amplitude values.

STATISTICAL ANALYSIS

To test whether there is a difference between unimodally and bimodally trained infants, while controlling for differences in the presented standard, we subjected the QS and non-QS datasets separately to an ANOVA with a mixed design (between-subject factors and repeated measures). The MMR amplitude was the dependent variable, Time Bin (100–150, 150–200, 200–250, 250–300, 300–350, 350–400, 400–450, and 450–500 ms) and Electrode (Fz, F3, F4, Cz, C3, C4, T7, and T8) were within-subject factors, and Distribution Type (unimodal vs. bimodal) and Standard Vowel ([æ] vs. [ɛ]) were between-subject factors. The design also included all possible interactions between the factors, up to the fourth order. To compensate for the double chance of finding results (separate QS and non-QS analyses) all tests employ a conservative α level of 0.025.

RESULTS

The grand average waveforms for each Distribution Type (unimodal vs. bimodal) pooled over the two levels of the factor Standard Vowel are presented in **Figure 2**, for 10 electrodes. In line with previous research on 2-to-3-month olds, the standard and deviant ERPs contained prominent slow positive waves (e.g., Friederici et al., 2002; Morr et al., 2002; Carral et al., 2005; Shafer et al., 2011), and the ERPs in the QS data appeared large compared to those in the non-QS data (e.g., for 2-month olds: Friederici et al., 2002; for newborns: Pihko et al., 2004; Sambeth et al., 2009; but see Cheour et al., 2002a, for conflicting results).

For the QS data, the ANOVA on the MMR amplitude yielded significant results neither for the research question (main effect of Distribution Type: $p = 0.88$), nor for any other main effect (Standard Vowel: $p = 0.23$; Electrode: $F < 1$; Time Bin: $F < 1$), nor for any of the 11 interactions (all p-values > 0.07).

For the non-QS data, the ANOVA revealed a positive grand mean (+0.84 μV), with a 97.5% confidence interval (CI) that does not include zero (+0.35 ∼ +1.33 μV), implying that on average Dutch 2-to-3-month old infants can discriminate the test vowels, and that vowel discrimination in these infants is reflected in a *positive* MMR. Regarding our specific research question, the analysis showed a main effect of Distribution Type (mean difference = +1.06 μV, CI = +0.08 ∼ +2.04 μV, $F[1,18] = 7.03$, $p = 0.016$, $\eta_p^2 = 0.28$): across electrodes and time windows the bimodally trained infants had a higher positive MMR (+1.37 μV, CI = +0.68 ∼ +2.06 μV) than the unimodally trained infants (+0.31 μV, CI = –0.38 ∼ +1.00 μV), indicating that Dutch 2-to-3-month olds' neural discrimination of [æ] and [ɛ] is better after bimodal than after unimodal training.

As for factors not directly pertaining to our research question, there was no effect of Standard Vowel ($p = 0.98$), so that we cannot state with confidence that one of the two combinations of standard and deviant vowel yields a higher MMR amplitude (and thus better neural discrimination) than the other combination. Further, the analysis showed no main effects of Time Bin ($F[7\varepsilon,126\varepsilon$, $\varepsilon = 0.334] = 1.37$, Greenhouse–Geisser corrected $p = 0.27$) or Electrode ($F < 1$). Thus, there was no support for a more positive or more negative MMR in any specific time window as compared to other ones within 100 and 500 ms, and at any specific electrode as compared to other ones among the frontocentral and temporal electrodes. Interestingly, we found a highly significant interaction effect between Distribution Type and Standard Vowel [$F(1,18) = 20.22$, $p = 0.0003$, $\eta_p^2 = 0.53$], which shows that the attested difference between unimodally and bimodally trained Dutch 2-to-3-month olds differs depending on the standard that they hear in the oddball test (see section Exploratory Results for the Four Groups).

EXPLORATORY RESULTS FOR THE FOUR GROUPS

To examine the responses of the four non-QS groups separately, we pooled the MMR amplitudes across electrodes and time bins in view of the lack of significant differences herein (see section Results). **Figure 3** shows the pooled MMR waveforms per group, and **Table 2** lists the corresponding averaged MMR amplitudes. The amplitude differed from zero significantly only for the Bimodal [ɛ] group ($p = 0.004$, uncorrected for multiple comparisons) implying that bimodally trained Dutch 2-to-3-month olds who are tested with standard [ɛ] and deviant [æ] can hear the difference between the two vowels.

The individual group's MMR amplitudes presented in **Table 2** are visualized in **Figure 4**. The interaction between Distribution Type and Standard Vowel, which was found in the main ANOVA for the non-QS data (see section Results), is clearly visible. We did the four relevant group comparisons, assuming equal variances for all groups (as in the ANOVA): Bimodal [ɛ] vs. Unimodal [ɛ], Bimodal [æ] vs. Unimodal [æ], Bimodal [ɛ] vs. Bimodal [æ] and Unimodal [æ] vs. Unimodal [ɛ] (technically, this was done via *post hoc* comparisons using Fisher's Least Significant Difference in SPSS). The Bimodal [ɛ] group's response was reliably more positive than that of the Unimodal [ɛ] group (see the arc numbered 1 and the black line in **Figure 4**; uncorrected $p = 0.00008$); this indicates that when the standard in the oddball paradigm is [ɛ] and the deviant is [æ], bimodally trained Dutch 2-to-3-month olds show better neural discrimination than unimodally trained infants. The difference between Bimodal [æ] and Unimodal [æ] was not significant ($p = 0.21$); thus, when the standard is [æ]

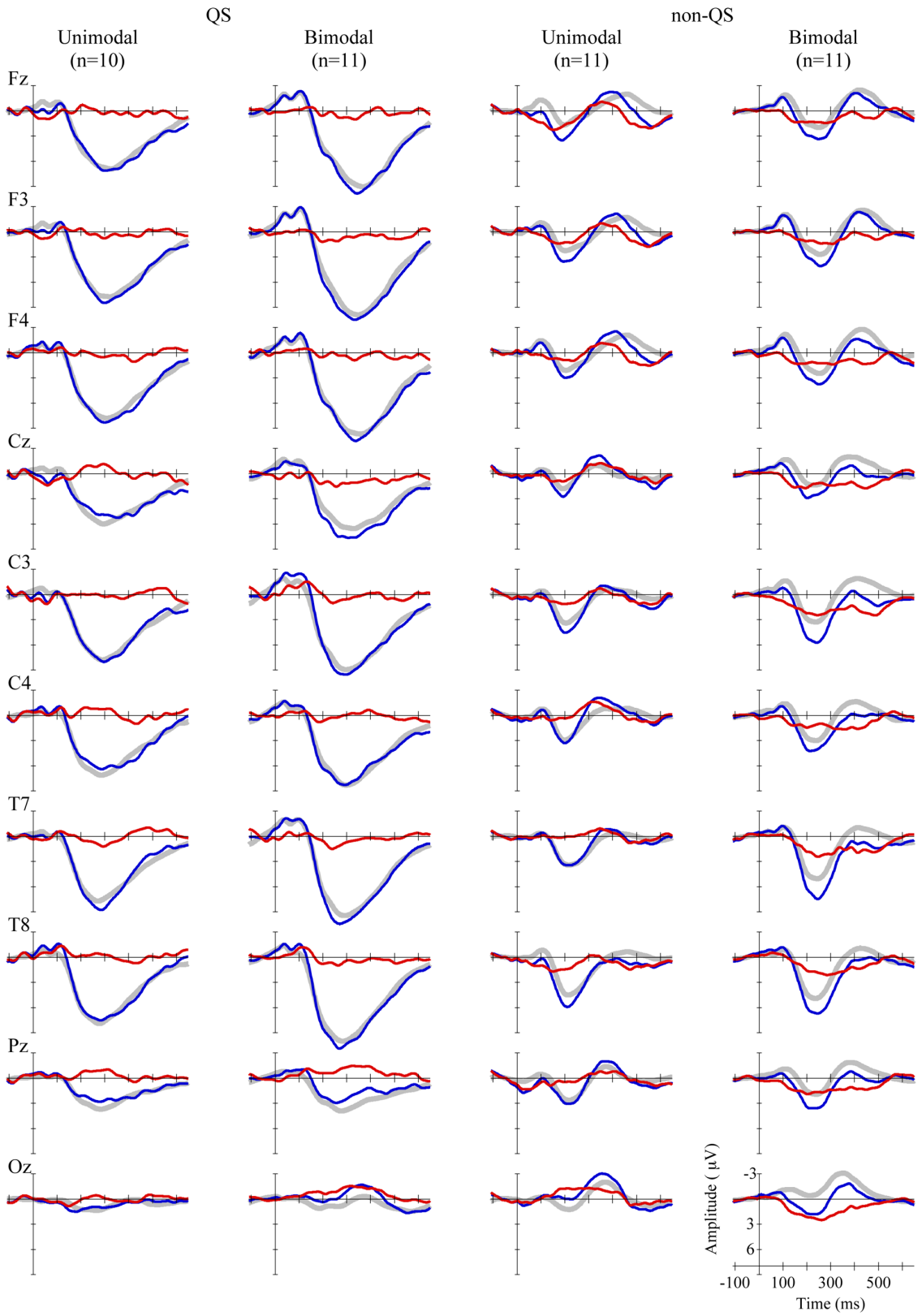

FIGURE 2 | Grand average standard (gray, thick curves), deviant (blue, thin curves), and MMR (red, thin curves) waveforms, at 10 electrodes (see rows), for unimodally and bimodally trained infants in QS (left two columns), and non-QS (right two columns).

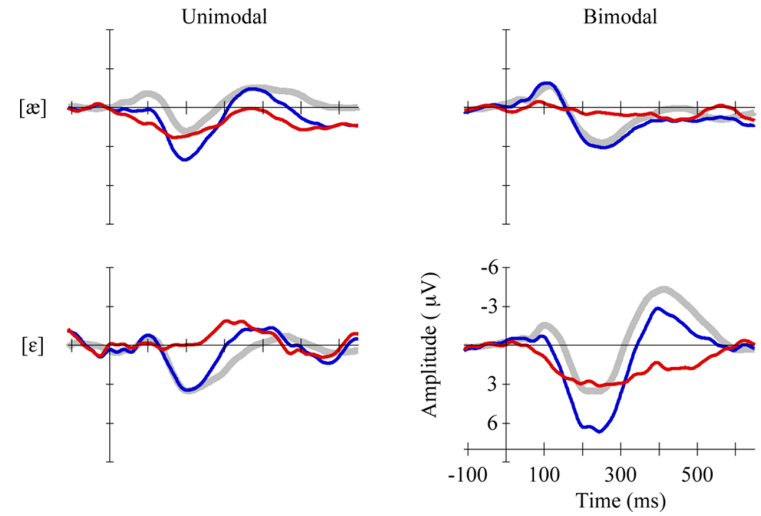

FIGURE 3 | Standards (gray, thick curves), deviants (blue, thin curves) and MMRs (red, thin curves) in non-QS, pooled across eight electrodes, per group (Unimodal [æ] top left vs. Bimodal [æ] top right, Unimodal [ε] bottom left vs. Bimodal [ε] bottom right).

Table 2 | Mean MMR amplitudes (in μV) between 100 and 500 ms across eight electrodes per subgroup for non-QS data, with within-group standard deviations (SD; between parentheses), and 97.5% confidence intervals.

Distribution Type	Standard Vowel	*N*	Mean (SD)	Confidence interval	*t*	*p*
Unimodal	[ε]	6	–0.59 (0.86)	–1.71 to +0.52	–1.69	0.153
Unimodal	[æ]	5	+1.21 (1.23)	–0.71 to +3.14	+2.20	0.092
Bimodal	[ε]	5	+2.26 (0.83)	+0.97 to +3.55	+6.12	0.004
Bimodal	[æ]	6	+0.48 (0.80)	–0.55 to +1.50	+1.46	0.203

Significance is tested against zero in four one-sample t-tests (without correction for multiple testing).

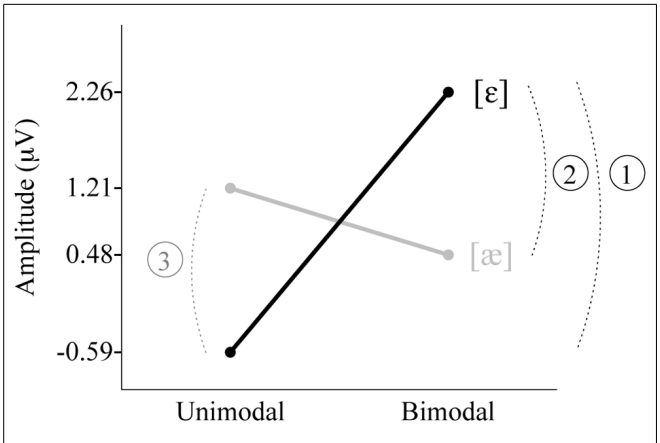

FIGURE 4 | Three *post hoc* significant differences in MMR amplitude between the four subgroups. Unimodal [ε] (left black), Unimodal [æ] (left gray), Bimodal [ε] (right black), Bimodal [æ] (right gray). Note: Among the four amplitudes, only the one for the Bimodal [ε] group differed from zero significantly.

and the deviant [ε], unimodally trained infants do not necessarily have higher response amplitudes. The Bimodal [ε] group's response was greater than that of the Bimodal [æ] group (the arc numbered 2 in **Figure 4**; $p = 0.005$), suggesting that neural discrimination is easier for bimodally trained Dutch 2-to-3-month olds when the standard is [ε] and the deviant is [æ] than when standard and deviant are reversed. Conversely, the Unimodal [æ] group's response was more positive than that of the Unimodal [ε] group's response (the arc numbered 3 in **Figure 4**; $p = 0.005$), which suggests that neural discrimination is easier for unimodally trained Dutch 2-to-3-month olds when the standard is [æ] and the deviant is [ε] than when standard and deviant are reversed.

DISCUSSION

The present study provides the first evidence for fast distributional learning in very young infants. The specific research question was whether Dutch 2-to-3-month old infants show larger mismatch responses, hence presumably better discrimination, of English [ε] and [æ] after bimodal than after unimodal training. This was answered in the affirmative with a *p*-value of 0.016. The age of 2 to 3 months is early enough for the distributional learning mechanism to be able to play a role in the transition from universal to language-specific speech perception, which has been observed to take place from 4 to 12 months.

This outcome extends previous research in two ways. First, fast distributional learning has now been attested at widely different ages, namely at 2 to 3 months (the present study), between 6 and 11 months (Maye et al., 2002, 2008; Yoshida et al., 2010; Capel et al., 2011), and in adults (Maye and Gerken, 2000, 2001; Gulian et al., 2007; Hayes-Harb, 2007; Escudero et al., 2011; Wanrooij et al., 2013; Wanrooij and Boersma, 2013). One can now hypothesize that the mechanism is available throughout life and can contribute to first and second language acquisition. Second, the ERP method has now been added to the set of methods by which distributional learning can be demonstrated. We needed the ERP method because of the young age of our participants, but this technique might have the general advantage over behavioral methods that it does not require the participant's attention and that it taps the response process at a time when the response is still little influenced by the myriads of factors that contribute to the behavioral part of the response. An assessment of the general usefulness of the ERP technique, especially in comparison with behavioral techniques, has to await replication with more age groups, larger sample sizes and more phonological contrasts.

The ERP method potentially yields information on the scalp distribution and the timing of the responses. Our results, however, do not allow us to determine any precise scalp location or timing. This indeterminacy is not uncommon in studies on infant MMRs (see section MMR Analysis), and may be due to the more pronounced shapes of sulci and gyri in adults than in infants (Hill et al., 2010) and to the larger variability in MMR timing among infants than among adults (e.g., Kushnerenko, 2003). More location- or time-specific results can be expected at later ages.

This study detected an interaction between the type of distribution (bimodal vs. unimodal) in the training and the identity of the standard vowel ([ɛ] vs. [æ]) in the test ($p < 0.001$); *post hoc* exploration suggested that bimodally trained infants discriminated better if the standard was [ɛ] and unimodally trained infants discriminated better if the standard was [æ]. This confirms none of the three predictions that we derived from previous literature in the Introduction: the peripherality-related asymmetry predicted on the basis of Polka and Bohn (1996), namely that the MMR should be larger if the standard is [ɛ], was not found (main effect of Standard Vowel: $p = 0.98$); a prediction indirectly derived from the "natural referent vowel" hypothesis (Polka and Bohn, 2011), namely that the peripherality-related asymmetry should occur only in the unimodal group, was contradicted by our detection of the asymmetry in the bimodal group and the opposite asymmetry in the unimodal group; a prediction derived from the "featural underspecified lexicon" model (Lahiri and Reetz, 2010), namely that the MMR should be larger if the standard is [æ], was not confirmed (main effect of Standard Vowel: $p = 0.98$). None of the hypotheses in the literature predicted the asymmetry that we did find, and we cannot speculate on it before many more ERP results on asymmetries have been collected.

Given the effect of distributional training in the young infants tested, the question arises what the mechanism is: is there an enhanced discrimination in the bimodally trained infants (acquired distinctiveness), or is there a reduced discrimination in the unimodally trained infants (acquired similarity), or both? We cannot answer this question on the basis of our results, because time constraints prevented us from testing the infants' perception before training. Also, a pre-test would have been an additional distributional training and could therefore have distorted the intended training distributions. Although to our knowledge MMRs for 2-to-3-month olds in response to similar small differences between vowels as between our test vowels (i.e., 1.38 ERB in F1 and 0.92 ERB in F2) have not been examined before, the acoustic difference between the test vowels was well above the discrimination threshold reported for 8-week old infants as measured behaviorally by high-amplitude sucking (Swoboda et al., 1976, 1978). On the other hand, the vowels in those studies were different, had different durations and were presented with different inter-stimulus intervals than in the current study, so that we cannot be certain that our 2-to-3-month olds discriminated the test vowels before training. Similarly, we cannot say if a potential perceptual ease of listening to the order [ɛ] – [æ] strengthened the effect of distributional learning for the bimodally trained infants and/or if a potential perceptual difficulty of listening to the opposite order weakened this effect.

One may wonder why the training–test paradigm works at all. After all, the test phase presents a (shrunk) bimodal distribution to the infants, and it can be expected that they continue to learn during the test, which lasts quite a bit longer (30 minutes) than what we call the "training" (12 minutes). The persistent influence of the training is possibly related to the much larger variability during training (900 different stimuli) than during the test (2 different stimuli). From other training paradigms it is known that a large variability in training stimuli can facilitate learning and could be instrumental in category formation (e.g., Lively et al., 1993). Future research is necessary to examine the persistence of short-term distributional learning over time.

With regard to the methodology of testing 2-to-3-month olds, the results highlight the importance of documenting sleep stages and analyzing QS data separately from non-QS data. In QS the MMR did not emerge, which is in line with the disappearance of the MMN in adult NREM sleep, and with the development of infant QS into an adult-like NREM in the first 3 months of life (see section Coding Sleep Stages), but in contrast to the lack of differences in the MMR between sleep stages in newborns (Martynova et al., 2003), and, for 2-month olds, to the *larger* MMR in QS than during wakefulness in Friedrich et al. (2004) and to the robust MMR in QS in Van Leeuwen et al. (2008). The many differences between these infant studies and the current study (if not simply due to chance) make it difficult to pinpoint the cause of this discrepancy. One difference from the studies mentioned is that the current study tested perception *after short-term training*. Thus, it may be that training effects were not yet sufficiently encoded in neural activation patterns to surface in QS. Alternatively, if infants who were in QS during the test, had already been in QS during the training, learning may have been hampered in QS as compared to non-QS.

We conclude that 2-to-3-month olds are sensitive to distributions of speech sounds in the environment. This is earlier than what has been shown in previous experiments with fast distributional learning, and earlier than the onset of language-specific speech perception. A linguistic interpretation of these results is

that at 2 months of age infants already have a mechanism in place that can support the acquisition of phonological categories.

ACKNOWLEDGMENTS

This research was supported by grant 277.70.008 from the Netherlands Organization for Scientific Research (NWO) awarded to the second author. We thank Dirk Jan Vet for his indispensable support with all technical matters, Anna Plakas, Johanna de Vos, Marieke van den Heuvel, Marja Caverlé, and Gisela Govaart for assisting with infant testing, and the parents and infants for their kind cooperation.

REFERENCES

Aslin, R. N., and Pisoni, D. B. (1980). "Some developmental processes in speech perception," in *Child Phonology,* Vol. 2, *Perception*, eds G. H. Yeni-Komshian, J. F. Kavanagh, and C. A. Ferguson (New York/London: Academic Press), 67–96.

Best, C. T., McRoberts, G. W., LaFleur, R., and Silver-Isenstadt, J. (1995). Divergent developmental patterns for infants' perception of two nonnative consonant contrasts. *Infant Behav. Dev.* 18, 339–350. doi: 10.1016/0163-6383(95)90022-5

Boersma, P., and Weenink, D. (2010). *Praat: Doing Phonetics by Computer*. Available at: http://www.praat.org (accessed 2010–2013).

Bosch, L., and Sebastián-Gallés, N. (2003). Simultaneous bilingualism and the perception of a language-specific vowel contrast in the first year of life. *Lang. Speech* 46, 217–243. doi: 10.1177/00238309030460020801

Broersma, M. (2005). Perception of familiar contrasts in unfamiliar positions. *J. Acoust. Soc. Am.* 117, 3890–3901. doi: 10.1121/1.1906060

Capel, D. J. H., De Bree, E. H., De Klerk, M. A., Kerkhoff, A. O., and Wijnen, F. N. K. (2011). "Distributional cues affect phonetic discrimination in Dutch infants," in *Sound and Sounds. Studies Presented to M.E.H (Bert) Schouten on the Occasion of his 65th Birthday*, eds W. Zonneveld, H. Quené, and W. Heeren (Utrecht: UiL-OTS), 33–43.

Carral, V., Huotilainen, M., Ruusuvirta, T., Fellman, V., Näätänen, R., and Escera, C. (2005). A kind of auditory "primitive intelligence" already present at birth. *Eur. J. Neurosci.* 21, 3201–3204. doi: 10.1111/j.1460-9568.2005.04144.x

Cheng, Y., Wu, H., Tzeng, Y., Yang, M., Zhao, L., and Lee, C. (2013). The development of mismatch responses to Mandarin lexical tones in early infancy. *Dev. Neuropsychol.* 38, 281–300. doi: 10.1080/87565641.2013.799672

Cheour, M., Alho, K., Čeponiené, R., Reinikainen, K., Sainio, K., Pohjavuori, M., et al. (1998a). Maturation of mismatch negativity in infants. *Int. J. Psychophysiol.* 29, 217–226. doi: 10.1016/S0167-8760(98)00017-8

Cheour, M., Čeponiené, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K., et al. (1998b). Development of language-specific phoneme representations in the infant brain. *Nat. Neurosci.* 1, 351–353. doi: 10.1038/1561

Cheour, M., Čeponiené, R., Leppänen, P., Alho, K., Kujala, T., Renlund, M., et al. (2002a). The auditory sensory memory trace decays rapidly in newborns. *Scand. J. Psychol.* 43, 33–39. doi: 10.1111/1467-9450.00266

Cheour, M., Martynova, O., Näätänen, R., Erkkola, R., Sillanpää, M., Kero, P., et al. (2002b). Speech sounds learned by sleeping newborns. *Nature* 415, 599–600. doi: 10.1038/415599b

Cheour, M., Leppänen, P., and Kraus, N. (2000). Mismatch negativity (MMN) as a tool for investigating auditory discrimination and sensory memory in infants and children. *Clin. Neurophysiol.* 111, 4–16. doi: 10.1016/S1388-2457(99)00191-1

Cheour-Luhtanen, M., Alho, K., Kujala, T., Sainio, K., Reinikainen, K., Renlund, M., et al. (1995). Mismatch negativity indicates vowel discrimination in newborns. *Hear. Res.* 82, 53–58. doi: 10.1016/0378-5955(94)00164-L

Cristià, A., McGuire, G. L., Seidl, A., and Francis, A. L. (2011). Effects of the distribution of acoustic cues on infants' perception of sibilants. *J. Phonet.* 39, 388–402. doi: 10.1016/j.wocn.2011.02.004

Crowell, D. H., Kapuniai, L. E., Boychuk, R. B., Light, M. J., and Hodgman, J. E. (1982). Daytime sleep stage organization in three-month-old infants. *Electroencephalogr. Clin. Neurophysiol.* 53, 36–47. doi: 10.1016/0013-4694(82)90104-3

Dehaene-Lambertz, G. (2000). Cerebral specialization for speech and non-speech stimuli in infants. *J. Cogn. Neurosci.* 12 , 449–460.

Dehaene-Lambertz, G., and Baillet, S. (1998). A phonological representation in the infant brain. *Neuroreport* 9, 1885–1888. doi: 10.1097/00001756-199806010-00040

Ellingson, R. J., and Peters, J. F. (1980). Development of EEG and daytime sleep patterns in normal full-term infants during the first 3 months of life: longitudinal observations. *Electroencephalogr. Clin. Neurophysiol.* 49, 112–124. doi: 10.1016/0013-4694(80)90357-0

Escudero, P., Benders, T., and Wanrooij, K. (2011). Enhanced bimodal distributions facilitate the learning of second language vowels. *J. Acoust. Soc. Am.* 130, EL206–EL212. doi: 10.1121/1.3629144

Escudero, P., Hayes-Harb, R., and Mitterer, H. (2008). Novel second-language words and asymmetric lexical access. *J. Phonet.* 36, 345–360. doi: 10.1016/j.wocn.2007.11.002

Friederici, A. D., Friedrich, M., and Weber, C. (2002). Neural manifestation of cognitive and precognitive mismatch detection in early infancy. *Neuroreport* 13, 1251–1254. doi: 10.1097/00001756-200207190-00006

Friedrich, M., Weber, C., and Friederici, A. D. (2004). Electrophysiological evidence for delayed mismatch response in infants at-risk for specific language impairment. *Psychophysiology* 41, 772–782. doi: 10.1111/j.1469-8986.2004.00202.x

Graven, S. N., and Browne, J. V. (2008). Sleep and brain development. The critical role of sleep in fetal and early neonatal brain development. *Newborn Infant Nurs. Rev.* 8, 173–179. doi: 10.1053/j.nainr.2008.10.008

Guenther, F. H., and Gjaja, M. N. (1996). The perceptual magnet effect as an emergent property of neural map formation. *J. Acoust. Soc. Am.* 100, 1111–1121. doi: 10.1121/1.416296

Gulian, M., Escudero, P., and Boersma, P. (2007). "Supervision hampers distributional learning of vowel contrasts," in *Proceedings of the International Congress of Phonetic Sciences* (Saarbrücken: University of Saarbrucken), 1893–1896.

Hawkins, S., and Midgley, J. (2005). Formant frequencies of RP monophthongs in four age groups of speakers. *J. Int. Phonet. Assoc.* 35, 183–199. doi: 10.1017/S0025100305002124

Hayes-Harb, R. (2007). Lexical and statistical evidence in the acquisition of second language phonemes. *Second Lang. Res.* 23, 65–94. doi: 10.1177/0267658307071601

He, C., Hotson, L., and Trainor, L. J. (2009). Development of infant mismatch responses to auditory pattern changes between 2 and 4 months old. *Eur. J. Neurosci.* 29, 861–867. doi: 10.1111/j.1460-9568.2009.06625.x

Hill, J., Dierker, D., Neil, J., Inder, T., Knutsen, A., Harwell, J., et al. (2010). A surface-based analysis of hemispheric asymmetries and folding of cerebral cortex in term-born human infants. *J. Neurosci.* 30, 2268–2276. doi: 10.1523/JNEUROSCI.4682-09.2010

Iber, C., Ancoli-Israel, S., Chesson, A. L., and Quan, S. F. (2007). *The AASM Manual for the Scoring of Sleep and Associated Events. Rules, Terminology and Technical Specifications.* Westchester IL: American Academy of Sleep Medicine.

Kahn, A., Dan, B., Grosswasser, J., Franco, P., and Sottiaux, M. (1996). Normal sleep architecture in infants and children. *J. Clin. Neurophysiol.* 13 , 184–197. doi: 10.1097/00004691-199605000-00002

Kral., A., and Eggermont, J. J. (2007). What's to lose and what's to learn: development under auditory deprivation, cochlear implants and limits of cortical plasticity. *Brain Res. Rev.* 56, 259–269. doi: 10.1016/j.brainresrev.2007.07.021

Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., and Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Dev. Sci.* 9, F13–F21. doi: 10.1111/j.1467-7687.2006.00468.x

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., and Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science* 255, 606–608. doi: 10.1126/science.1736364

Kujala, A., Huotilainen, M., Hotakainen, M., Lennes, M., Parkkonen, L., Fellman, V., et al. (2004). Speech-sound discrimination in neonates as measured with MEG. *Neuroreport* 15, 2089–2092. doi: 10.1097/00001756-200409150-00018

Kushnerenko, E. (2003). *Maturation of the Cortical Auditory Event-Related Brain Potentials in Infancy*. Doctoral dissertation, University of Helsinki.

Lacerda, F. (1995). "The perceptual-magnet effect: an emergent consequence of exemplar-based phonetic memory," in *Proceedings of the XIIIth International Congress of Phonetic Sciences,* Vol. 2, eds K. Elenius and P. Branderyd (Stockholm: KTH and Stockholm University), 140–147.

Lahiri, A., and Reetz, H. (2010). Distinctive features: phonological underspecification in representation and processing. *J. Phonet.* 38, 44–59. doi: 10.1016/j.wocn.2010.01.002

Lamel, L. F., Kassel, R. H., and Seneff, S. (1986). *Speech Database Development: Design and Analysis of the Acoustic-Phonetic Corpus.* Proc. DARPA Speech Recognition Workshop. Report No. SAIC-86/1546, 100–109.

Lively, S. E., Logan, J. S., and Pisoni, D. B. (1993). Training Japanese listeners to identify English /r/ and /l/. II: the role of phonetic environment and talker variability in learning new perceptual categories. *J. Acoust. Soc. Am.*94, 1242–1255.

Loewy, D., Campbell, K., and Bastien, C. (1996). The mismatch negativity to frequency deviant stimuli during natural sleep. *Electroencephalogr. Clin. Neurophysiol.* 98, 493–501. doi: 10.1016/0013-4694(96)95553-4

Loewy, D., Campbell, K., De Lugt, D., Elton, M., and Kok, A. (2000). The mismatch negativity during natural sleep: intensity deviants. *Clin. Neurophysiol.* 111, 863–872. doi: 10.1016/S1388-2457(00)00256-X

Martynova, O., Kirjavainen, J., and Cheour, M. (2003). Mismatch negativity and late discriminative negativity in sleeping human newborns. *Neurosci. Lett.* 340, 75–78. doi: 10.1016/S0304-3940(02)01401-5

Maye, J., and Gerken, LA. (2000). "Learning phonemes without minimal pairs," in *BUCLD 24 Proceedings*, ed. C. Howell (Somerville, MA: Cascadilla Press), 522–533.

Maye, J., and Gerken, LA. (2001). "Learning phonemes: how far can the input take us?" in *BUCLD 25 Proceedings*, ed. A. H.-J. Do (Somerville, MA: Cascadilla Press), 480–490.

Maye, J., Weiss, D., and Aslin, R. (2008). Statistical phonetic learning in infants: facilitation and feature generalization. *Dev. Sci.* 11, 122–134. doi: 10.1111/j.1467-7687.2007.00653.x

Maye, J., Werker, J. F., and Gerken, L. A. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* 82, B101–B111. doi: 10.1016/S0010-0277(01)00157-3

Moore, J. K. (2002). Maturation of human auditory cortex: implications for speech perception. *Ann. Otol. Rhinol. Laryngol.* 111, 7–10.

Moore, J. K., and Guan, Y.-L. (2001). Cytoarchitectural and axonal maturation in human auditory cortex. *J. Assoc. Res. Otolaryngol.* 2, 297–311. doi: 10.1007/s101620010052

Moore, J. K., and Linthicum, F. H. (2007). The human auditory system: a timeline of development. *Int. J. Audiol.* 46 , 460–478. doi: 10.1080/14992020701383019

Morr, M. L., Shafer, V. L., Kreuzer, J. A., and Kurtzberg, D. (2002). Maturation of mismatch negativity in typically developing infants and preschool children. *Ear Hear* 23, 118–136. doi: 10.1097/00003446-200204000-00005

Näätänen, R. (1992). *Attention and Brain Function.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Näätänen, R., and Winkler, I. (1999). The concept of auditory stimulus representation in cognitive neuroscience. *Psychol. Bull.* 125, 826–859. doi: 10.1037/0033-2909.125.6.826

Näätänen, R., Gaillard, A. W. K., and Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychol.* 42, 313–329. doi: 10.1016/0001-6918(78)90006-9

Näätänen, R., Paavilainen, P., Alho, K., Reinikainen, K., and Sams, M. (1989). Do event-related potentials reveal the mechanisms of auditory sensory memory in the human brain? *Neurosci. Lett.* 98, 217–221. doi: 10.1016/0304-3940(89)90513-2

Näätänen, R., Paavilainen, P., Rinne, T., and Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clin. Neurophysiol.* 118, 2544–2590. doi: 10.1016/j.clinph.2007.04.026

Niedermeyer, E. (2005). "Maturation of the EEG: development of waking and sleep patterns," in *Electroencephalography: Basic Principles, Clinical Applications and Related Fields*, eds E. Niedermeyer and F. L. Da Silva (Philadelphia: Lippincott Williams & Wilkins), 209–233.

Partanen, E., Pakarinen, S., Kujala, T., and Huotilainen, M. (2013). Infants' brain responses for speech sound changes in fast multifeature MMN paradigm. *Clin. Neurophysiol.* 124, 1578–1585. doi: 10.1016/j.clinph.2013.02.014

Pihko, E., Sambeth, A., Leppänen, P., Okada, Y., and Lauronen, L. (2004). Auditory evoked magnetic fields to speech stimuli in newborns – effect of sleep stages. *Neurol. Clin. Neurophysiol.* 6, 1–5.

Pisoni, D. B. (1977). Identification and discrimination of the relative onset time of two component tones: implications for voicing perception in stops. *J. Acoust. Soc. Am.* 61, 1352–1361. doi: 10.1121/1.381409

Polka, L., and Bohn, O.-S. (1996). A cross-language comparison of vowel perception in English-learning and German-learning infants. *J. Acoust. Soc. Am.* 100, 577–592. doi: 10.1121/1.415884

Polka, L., and Bohn, O.-S. (2003). Asymmetries in vowel perception. *Speech Commun.* 41, 221–231. doi: 10.1016/S0167-6393(02)00105-X

Polka, L., and Bohn, O.-S. (2011). Natural Referent Vowel (NRV) framework: an emergent view of early phonetic development. *J. Phonet.* 39, 467–478. doi: 10.1016/j.wocn.2010.08.007

Polka, L., Colantonio, C., and Sundara, M. (2001). A cross-language comparison of /d/-/ð/ perception: evidence for a new developmental pattern. *J. Acoust. Soc. Am.* 109, 2190–2201. doi: 10.1121/1.1362689

Polka, L., and Werker, J. F. (1994). Developmental changes in perception of non-native vowel contrasts. *J. Exp. Psychol. Hum. Percept. Perform.* 20, 421–435. doi: 10.1037/0096-1523.20.2.421

Pons, F., Albareda-Castellot, B., and Sebastián-Gallés, N. (2012). The interplay between input and initial biases: asymmetries in vowel perception during the first year of life. *Child Dev.* 1–12. doi: 10.1111/j.1467-8624.2012.01740.x

Rodenbeck, A., Binder, R., Geisler, P., Danker-Hopfe, H., Lund, R., Raschke, F., et al. (2007). A review of sleep EEG patterns. Part I: a compilation of amended rules for their visual recognition according to Rechtschaffen and Kales. *Somnologie* 10, 159–175. doi: 10.1111/j.1439-054X.2006.00101.x

Sambeth, A., Pakarinen, S., Ruohio, K., Fellman, V., Van Zuijen, T. L., and Huotilainen, M. (2009). Change detection in newborns using a multiple deviant paradigm: a study using magnetoencephalography. *Clin. Neurophysiol.* 120, 530–538. doi: 10.1016/j.clinph.2008.12.033

Schouten, M. E. H. (1975). *Native-Language Interference in the Perception of Second-Language Vowels: An Investigation of Certain Aspects of the Acquisition of a Second Language.* Doctoral dissertation, Utrecht University.

Schröger, E. (1997). On the detection of auditory deviations: a pre-attentive activation model. *Psychophysiology* 34, 245–257. doi: 10.1111/j.1469-8986.1997.tb02395.x

Shafer, V. L., Yu, Y. H., and Datta, H. (2011). The development of English vowel perception in monolingual and bilingual infants: neurophysiological correlates. *J. Phonet.* 39, 527–545. doi: 10.1016/j.wocn.2010.11.010

Sokolov, E., Spinks, J., Näätänen, R., and Lyytinen, H. (2002). *The Orienting Response in Information Processing.* Mahwah, New Jersey/London: Lawrence Erlbaum Associates.

Sundara, M., Polka, L., and Genesee, F. (2006). Language-experience facilitates discrimination of /d-ð/ in monolingual and bilingual acquisition of English. *Cognition* 100, 369–388. doi: 10.1016/j.cognition.2005.04.007

Swoboda, P. J., Kass, J., Morse, P. A., and Leavitt, L. A. (1978). Memory factors in vowel discrimination of normal and at-risk infants. *Child Dev.* 49, 332–339. doi: 10.2307/1128695

Swoboda, P. J., Morse, P. A., and Leavitt, L. A. (1976). Continuous vowel discrimination in normal and at risk infants. *Child Dev.* 47, 459–465. doi: 10.2307/1128802

Tsao, F.-M., Liu, H.-M., and Kuhl, P. K. (2006). Perception of native and non-native affricate-fricative contrasts: cross-language tests on adults and infants. *J. Acoust. Soc. Am.* 120, 2285–2294. doi: 10.1121/1.2338290

Tsushima, T., Takizawa, O., Sasaki, M., Shiraki, S., Nishi, K., Kohno, M., et al. (1994). "Discrimination of English /r-l/ and /w-y/ by Japanese infants at 6-12 months: language-specific developmental changes in speech perception abilities," *International Conference on Spoken Language Processing (ICSLP),* (Yokohama), 1695–1698.

Van Leeuwen, T., Been, P., Van Herten, M., Zwarts, F., Maassen, B., and Van der Leij, A. (2008). Two-month-old infants at risk for dyslexia do not discriminate /bAk/ from /dAk/: a brain-mapping study. *J. Neurolinguist.* 21, 333–348. doi: 10.1016/j.jneuroling.2007.07.004

Wanrooij, K., and Boersma, P. (2013). Distributional training of speech sounds can be done with continuous distributions. *J. Acoust. Soc. Am.* 133, EL398–EL404. doi: 10.1121/1.4798618

Wanrooij, K., Escudero, P., and Raijmakers, M. E. J. (2013). What do listeners learn from exposure to a vowel distribution? An analysis of listening strategies in distributional learning. *J. Phon.* 41, 307–319. doi: 10.1016/j.wocn.2013.03.005

Weber, A., and Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *J. Mem. Lang.* 50, 1–25. doi: 10.1016/S0749-596X(03)00105-0

Weber, C., Hahne, A., Friedrich, M., and Friederici, A. D. (2004). Discrimination of word stress in early infant perception: electrophysiological evidence. *Cogn. Brain Res.* 18, 149–161. doi: 10.1016/j.cogbrainres.2003.10.001

Werker, J. F., and Tees, R. C. (1984/2002). Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behav.*

Dev. 25, 121–133. (First published in 1984: *Infant Behav. Dev.* 7, 49–63). doi: 10.1016/S0163-6383(84)80022-3

Yeung, H. H., Chen, K. H., and Werker, J. F. (2013). When does native language input affect phonetic perception? The precocious case of lexical tone. *J. Mem. Lang.* 68, 123–139. doi: 10.1016/j.jml.2012.09.004

Yoshida, K. A., Pons, F., Maye, J., and Werker, J. F. (2010). Distributional phonetic learning at 10 months of age. *Infancy* 15, 420–433. doi: 10.1111/j.1532-7078.2009.00024.x

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 07 November 2013; accepted: 20 January 2014; published online: 25 February 2014.

*Citation: Wanrooij K, Boersma P and van Zuijen TL (2014) Fast phonetic learning occurs already in 2-to-3-month old infants: an ERP study. Front. Psychol. **5**:77. doi: 10.3389/fpsyg.2014.00077*

This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology.

Copyright ©2014 Wanrooij, Boersma and van Zuijen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.