Developmental Cognitive Neuroscience 26 (2017) 45–51

Contents lists available at ScienceDirect

Developmental Cognitive Neuroscience

journal homepage: www.elsevier.com/locate/dcn

# A novel form of perceptual attunement: Context-dependent perception of a native contrast in 14-month-old infants

Mathilde Fort[a,*,1], Perrine Brusini[b,1], M. Julia Carbajal[c], Yue Sun[c,d], Sharon Peperkamp[c,e]

[a] Universidad Pompeu Fabra, Center for Brain and Cognition, Barcelona, Spain
[b] Center for Neuroscience in Education, University of Cambridge, Downing St, Cambridge CB2 3EB, United Kingdom
[c] Laboratoire de Sciences Cognitives et Psycholinguistique (Département d'Etudes Cognitives, Ecole Normale Supérieure – PSL Research University/Ecole des Hautes Etudes en Sciences Sociales/Centre National de la Recherche Scientifique), Paris, France
[d] Max-Planck-Institut für empirische Ästhetik, Frankfurt am Main, Germany
[e] Maternité Port-Royal, AP-HP, Faculté de Médecine Descartes, Paris, France

ARTICLE INFO

*Keywords:*
Perceptual attunement
Voicing assimilation
Phonetic learning
Electroencephalography

ABSTRACT

By the end of their first year of life, infants have become experts in discriminating the sounds of their native language, while they have lost the ability to discriminate non-native contrasts. This type of phonetic learning is referred to as perceptual attunement. In the present study, we investigated the emergence of a context-dependent form of perceptual attunement in infancy. Indeed, some native contrasts are not discriminated in certain phonological contexts by adults, due to the presence of a language-specific process that neutralizes the contrasts in those contexts. We used a mismatch design and recorded high-density Electroencephalography (EEG) in French-learning 14-month-olds. Our results show that similarly to French adults, infants fail to discriminate a native voicing contrast (e.g., [f] vs. [v]) when it occurs in a specific phonological context (e.g. [ofbe] vs. [ovbe], no mismatch response), while they successfully detected it in other phonological contexts (e.g., [ofne] vs. [ovne], mismatch response). The present results demonstrate for the first time that by the age of 14 months, infants' phonetic learning does not only rely on the processing of individual sounds, but also takes into account in a language-specific manner the phonological contexts in which these sounds occur.

## 1. Introduction

Speech is an inherently variable signal: tokens of identical phonemes and words are acoustically distinct. When acquiring their native language, infants have thus to detect the equivalence of these different tokens in spite of their variability. For instance, the acoustic properties of consonants differ as a function of the following vowel. After only a few months of life, infants normalize this acoustic variation (Bertoncini et al., 1988; Hochmann and Papeo, 2014; Mersad and Dehaene-Lambertz, 2016). Infants must also detect which aspects of phonetic variation in the speech signal are meaningful, i.e. reflect distinctions among the phonemes and hence words of their native language. Numerous behavioral studies report that by the end of the first year of life, infants have learned to interpret which phonetic distinctions are relevant (or not) to recognize the phonemes of their native language (e.g., Kuhl et al., 1992; Polka et al., 1994; Werker and Tees, 1984). More recently, electrophysiological measures have confirmed these findings (Bosseler et al., 2013; Conboy and Kuhl, 2011; Peña et al., 2012; Rivera-Gaxiola et al., 2007a, 2005a,b). For instance, Peña et al. (2012) collected electro-encephalographic (EEG) measures in 9- and 12-month-old Spanish-learning infants using a mismatch paradigm. They computed infants' auditory mismatch response (MMR), an ERP component that reflects the automatic detection of perceptual change in both adults (Näätänen et al., 1997, 2012) and infants (Dehaene-Lambertz and Baillet, 1998; Dehaene-Lambertz and Dehaene, 1994). In each trial, infants listened to a series of three identical syllables ([ɖa]) that were followed by a fourth syllable that was either identical or differed from them in the first consonant. Crucially, the deviant consonant always differed on the same phonetic feature (i.e., place of articulation) but could be either native (i.e., [ba]) or non-native (i.e., [ɖa], with [ɖ] a retroflex consonant that does not occur in Spanish). The authors observed that at 9 months, infants showed a MMR in response to both the native and the non-native deviant consonants; hence, they detected the phonetic change in place of articulation from [ɖa] to both

* Corresponding author at: Center for Brain and Cognition (CBC), Departament de Tecnologies de la Informació i les Comunicacions (DTIC), Speech Acquisition and Perception group (SAP), Universitat Pompeu Fabra, C/Tànger 122-140, 08018 Barcelona, Spain.
*E-mail address:* mathilde.frt@gmail.com (M. Fort).
[1] Joint authors

https://doi.org/10.1016/j.dcn.2017.04.006
Received 27 June 2016; Received in revised form 11 February 2017; Accepted 13 April 2017
Available online 26 April 2017
1878-9293/ © 2017 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/).

native [ba] and non-native [ɖa]. By contrast, 12-month-olds showed a MMR in response to the change to native [ba], but failed to detect the phonetic change to non-native [ɖa].

Thus, by the age of 12 months, infants have lost the ability to discriminate non-native phonetic contrasts. This type of phonetic learning is often referred to as *perceptual attunement* (or *perceptual narrowing*). It should be noted that perceptual attunement likewise occurs in the case of *native* phonetic contrasts that are not used to distinguish word meaning. For instance, in English, the consonant [t] is usually aspirated when it occurs before a vowel (as in *top* [tʰɔp]), but it is unaspirated if preceded by [s], (as in *stop* [stɔp]). Acoustically, this unaspirated [t] is very similar to [d], which never occurs preceded by [s], and native English adults have difficulty discriminating the contrast between unaspirated [t] and [d] despite the fact that they both occur in their language. Importantly, English-learning infants discriminate the contrast at 6–8 months of age, but fail to do so by the time they are 10–months of age, thus showing perceptual attunement (Pegg and Werker, 1997). In a similar vein, English-learning infants lose their sensitivity to the contrast between oral and nasal vowels between the ages of 4 and 11 months (Seidl et al., 2009). While both types of vowels occur in their language, the contrast between them is never used to distinguish meaning: nasal vowels are phonetic variants of oral vowels, occurring before nasal consonants (cf. *beet* and *bean*, whose vowels only differ in nasality).

In this article we focus on yet another type of perceptual attunement, which has hitherto not been investigated. Specifically, depending upon the phonological context in which they appear, even native contrasts that *are* used to distinguish word meaning can be difficult to discriminate (Mitterer and Blomert, 2003; Sun et al., 2015). For instance, Sun et al. (2015) collected EEG measures in a mismatch paradigm and showed that French adults do not detect a change from [ofbe] to [ovbe] (i.e., they fail to show a MMR), despite the fact that the distinction between [f] and [v], a voicing[2] distinction, is phonemic in their language and hence used to distinguish meaning (e.g., *greffe* [gʁɛf] 'graft' – *grève* [gʁɛv] 'strike'); by contrast, they correctly detect a change from [ofne] to [ovne], as evidenced by the presence of a MMR. This result can be explained by the fact that in French, native voicing contrasts – such as the one between [f] and [v] – can be neutralized in certain phonological contexts. Specifically, in speech production, voiceless consonants can become voiced when followed by certain voiced consonants, a process called *voicing assimilation*.[3] For instance, the voiceless [f] in *oeuf* [œf] 'egg' is typically produced as its voiced counterpart [v] in *oeuf blanc* [œvblɑ̃] 'white egg', where it is followed by [b]. This process does not apply, however, in *oeuf noir* [œfnwaʁ] 'black egg', where it is followed by [n]. Specifically, French voicing assimilation applies before voiced obstruents (such as [b]) but not in other contexts (such as before [n]). The findings by Sun et al. (2015), then, show that this process affects speech perception: when presented with consonants in an assimilation context (i.e., followed by a voiced obstruent), French adults fail to detect a change in voicing. In other words, the capacity to discriminate a native contrast depends upon the phonological context in which the contrast occurs.

For phonetic learning, this raises the question of whether infants' perceptual attunement is likewise context-dependent, taking into account language-specific phonological processes and the contexts in which they apply. Nine-month-old infants are sensitive to phonological context: they have acquired constraints upon the sequencing of phonemes within syllables and words in their language (i.e., phonotactics), and prefer to listen to phoneme sequences that are phonotactically legal or frequent as opposed to illegal or infrequent (Friederici and Wessels, 1993; Jusczyk et al., 1993; Jusczyk and Luce, 1994). Moreover, 14-month-old infants show some evidence of context-sensitive perceptual attunement: Japanese infants of this age perceive an illusory vowel within consonant clusters that are illegal in their language (Mazuka et al., 2011), similarly to Japanese adults (Dehaene-Lambertz et al., 2000; Dupoux et al., 1999). However, whether and how language-specific phonological processes (like French voicing assimilation) influence infants' capacity to discriminate native phoneme contrasts in the context of these processes has not been investigated.

In the present study, we thus tested whether French 14-month-olds, who have already lost the ability to distinguish non-native contrasts,[4] are also insensitive to the native voicing contrast when it occurs in an assimilation context. We recorded high-density EEG, using a mismatch trial design as in Peña et al. (2012). We analyzed the presence or absence of a MMR in response to a voicing change in two types of context, one that allows and one that does not allow for voicing assimilation. The MMR is particularly suited to study native phoneme discrimination, given its sensitivity to the linguistic relevance of the phonetic change in the participant's native language (Dehaene-Lambertz and Baillet, 1998; Näätänen et al., 1997). If 14-month-olds have already developed sensitivity to the native process of voicing assimilation, they should – like French adults – exhibit a MMR in response to a voicing change only in contexts that cannot trigger voicing assimilation in French. Alternatively, if they have not yet developed this sensitivity, they should exhibit a MMR to a voicing change regardless of context.

## 2. Material and methods

### 2.1. Participants

Forty-three healthy infants (23 females; mean age: 427 days, range: 396–441 days) raised in a monolingual French-speaking environment participated in the study. All infants were born full-term (37–42 weeks gestation) with normal birthweight (> 2500 g). All parents gave informed consent before the study. According to parental report, infants had normal vision and audition and had no exposure to other languages. Forty additional infants were excluded from the analyses because of an insufficient number of trials (N = 27), parental interference (N = 2), or because they presented less than 13 artifact-free EEG trials per condition (N = 11). We received 29 others infants who did not provide any data because they refused to wear the net.

### 2.2. Design and stimuli

The stimuli were a subset of the ones used by Sun et al. (2015); they all had a vowel-consonant-consonant-vowel structure ($V_1C_1C_2V_2$). Given infants' limited attention span, we used two out of the eight original item pairs, i.e. [ofbe]-[ofne] and [ikdo]-[ikmo]. Each of the precursor stimuli was spoken by four female speakers, while the test stimuli was spoken by a male speaker (we used one token per speaker); the use of multiple speakers promotes the detection of phonological rather than purely acoustical changes (Dehaene-Lambertz, 2000; Eulitz and Lahiri, 2004). All $V_1C_1C_2V_2$ tokens was created by means of cross-splicing of entirely voiceless $C_1C_2$ consonant clusters with entirely voiced ones. Thus, in all stimuli (both precursor and test stimuli), $C_1$

[2] Voicing is the phonetic feature that refers to the vibration of the vocal cords during the production of a given speech sound. Sounds that are produced with vibration of the vocal cords are called voiced, and those without vibration are called voiceless.

[3] Similarly, certain voiced consonants can become voiceless when followed by a voiceless consonant. In this article, however, we are only concerned with the change from voiceless to voiced.

[4] To our knowledge, there is no ERP study investigating perceptual attunement for phoneme categories in French-learning infants, but given that it has been documented at 12 months of age for infants learning a number of languages, including English, Spanish, and Japanese, we have every reason to assume that French 14-month-olds have likewise acquired their native phoneme categories.

**Table 1**
Design.

| Pair | Context | Precursor (x4) | Test stimuli |
|---|---|---|---|
| *ofbe-ofne* | | | No Change: [ɔfbe] |
| | Assimilation | [ɔfbe] | Voicing Change: [ɔvbe] |
| | | | Place Change *localizer*: [ɔsbe] |
| | | | No Change: [ɔfne] |
| | No-assimilation | [ɔfne] | Voicing Change: [ɔvne] |
| | | | Place Change *localizer*: [ɔsne] |
| *ikdo-ikmo* | | | No Change: [ikdo] |
| | Assimilation | [ikdo] | Voicing Change: [igdo] |
| | | | Place Change *localizer*: [ipdo] |
| | | | No Change: [ikmo] |
| | No-assimilation | [ikmo] | Voicing Change: [igmo] |
| | | | Place Change *localizer*: [ipmo] |

was always completely voiceless or voiced across all the conditions. For instance, [ofbe] was created by combining [of]- from [ofpe] and − [be] from [ovbe], such that [f] was completely voiceless (see Sun et al. (2015); for more details). The duration of the items was normalized to 500 ms; the duration of the $VC_1$ part varied (mean: 229 ms; range: 190–279 ms). Similarly, the intensity of the items was normalized to mean RMS amplitude (60 dB).

The stimuli were presented in series of five, with an ISI of 400 ms and a final 2100 ms of silence. In each series, four identical precursor stimuli (e.g., [ofbe] or [ofne] four consecutive times) were followed by a fifth stimulus. In the No Change condition, the last item was identical to the precursors; in the Voicing Change condition, it differed in the voicing of $C_1$ (e.g., [ovbe] or [ovne], respectively). Given the voicing assimilation process in French, a voicing change is expected to be perceived in [ofne]-[ovne] (No Assimilation context), but not in [ofbe]-[ovbe] (Assimilation context, cf. Table 1). We also used a control Place Change condition in which the fifth stimulus differed from the previous ones in its place of articulation of $C_1$ (e.g., [osbe] or [osne], respectively). Given that these place of articulation changes are phonemic in French and do not reflect any assimilation process, they are expected to be perceived in both the Assimilation context (e.g., [ofbe]-[osbe]) and the No Assimilation context (e.g., [ofne]-[osne]); specifically, see Dehaene-Lambertz and Baillet (1998) for evidence that 3- to 4-month-old French infants are sensitive to a change in place of articulation.

### 2.3. Procedure

Infants were seated on their parent's lap in a sound-proof room. To limit body movement, they were presented with silent animated cartoons on a screen, and/or passively observed the experimenter making soap bubbles, and/or played with a small silent toy during testing. They heard 240 trials of 5 items, 40 in each of the six conditions (two types of Contexts, i.e., Assimilation and No Assimilation; by three types of test stimuli, i.e., No Change, Voicing Change and Place Change). The order of the female voices for the precursor stimuli was counterbalanced across trials. Trials were presented in a pseudo-random order, such that no more than two consecutive trials used the same precursor. The experiment was run using the MATLAB Psychophysics Toolbox extensions (Brainard, 1997). The maximum duration was approximately 20 minutes, but the experiment was stopped earlier if the infant became restless.

### 2.4. EEG recording and preprocessing

EEG was recorded using the Net Station 4.3 (EGI software, net amp 200 system) and the EGI (Electrical Geodesics Incorporated, USA) 128-channel HCGSN sensor net adapted to infants' cranial perimeter. During the recording, the EEG signal was referenced to the vertex (Cz) and digitized at a sampling rate of 250 Hz.

After acquisition, data were preprocessed using MATLAB and the EEGLab toolbox (Delorme and Makeig, 2004). Recordings were digitally band-pass filtered (0.3–20 Hz) and segmented into epochs starting 500 ms prior to the onset of the first consonant of each test stimulus and ending 600 ms after it. For each epoch, channels contaminated by eye or motion artifacts (defined as local deviations higher than 80 μV) were automatically excluded, and trials with more than 40% contaminated channels were rejected. Channels containing fewer than 50% good trials were also automatically rejected for the entire recording. The artifact-free epochs (total mean number: 169, total range: 100–239, mean number per condition: 28, range per condition: 13–40) were then averaged for each participant in each of the six conditions (Context: Assimilation, No Assimilation x Phonological Change: No Change, Voicing Change, Place Change), detrended, baseline-corrected (from 500 ms before the onset of the critical consonant $C_1$), and finally transformed into reference-independent values using the average reference method (Luck, 2005).

### 2.5. EEG data analysis

Due to the multiple dimensions that can potentially be considered, ERP data analyses are especially sensitive to Type I errors (false alarms), notably when selecting the time window and cluster of electrodes to include in the analysis. To discard this multiple comparison problem, we used our control condition (Place Change) as a localizer to determine the time window and the topography of the MMR in our study. Specifically, we compared the recorded activity in the Place Change condition relative to the No Change condition, collapsing the data of the Assimilation and the No Assimilation contexts in order to maximize statistical power and avoid a selection biased towards one particular context. Indeed, a change in place of articulation should be detected in both the Assimilation and the No Assimilation contexts. We visually selected the time windows and electrode clusters featuring the difference between these two conditions, and averaged the voltage over the selected electrodes and time windows for each participant (Fig. 1). We then used the same time windows and electrode clusters to analyze our contrasts of interest in a 2 × 2 ANOVA with Context (Assimilation vs. No Assimilation) and Phonological Change (No Change vs. Voicing Change) as within-participant factors.

## 3. Results

### 3.1. Determination of the channel-clusters and time windows of interest with the place change condition

The visual inspection of the two-dimensional reconstructions of the difference between the Place Change and the No Change conditions (averaged across both Assimilation and No Assimilation contexts) revealed two distinct negativity peaks around the right centro-frontal electrodes, including C4 and T8 synchronous with a positivity on the posterior electrodes. The first negative response lasted from 0 to 200 ms after the onset of the critical consonant ($C_1$); the second negativity response lasted from 250 until 400 ms after the onset of the critical consonant (cf. Fig. 1).

We infer that the first negativity response is automatically generated by the detection of the acoustic differences of $C_1$ in the first syllable of the precursor and test items (e.g., [of], [ov] and [os] for the [ofbe]-[ofne] pair). Its bilateral topography and latency (0–200 ms) are compatible with the early auditory response usually observed in infants of the same age (Rivera-Gaxiola et al., 2007b; Rivera-Gaxiola et al., 2005a,b). The second response consisted of a right anterior negative pole and a reverse of polarity over the left temporal region. This topography and latency is compatible with a mismatch response, which is interpreted as an error signal between the predictions based on the repetition of the precursors and the actual perception of the target stimulus (Basirat et al., 2014; Bendixen et al., 2010; Lieder et al., 2013; Wacongne et al., 2012; Winkler, 2007). Given the nature of the

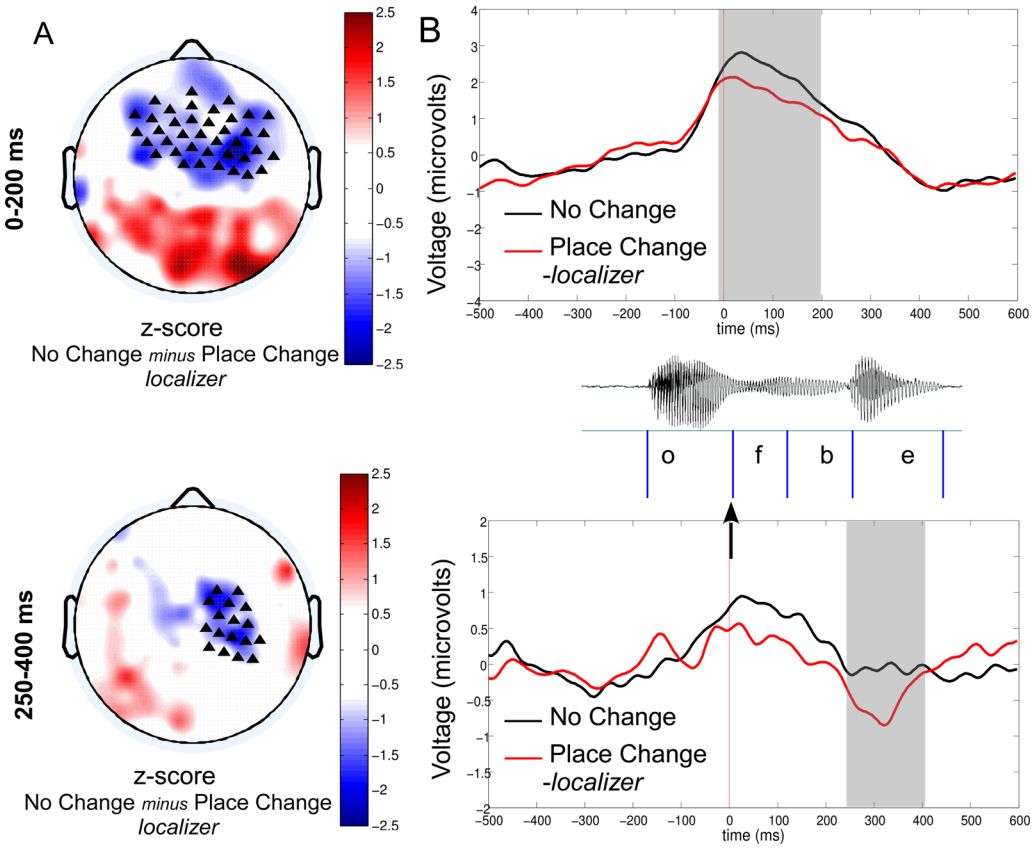


**Fig. 1.** MMR obtained for the localizer comparison between No Change *minus* Place Change, averaged across the Assimilation and the No-assimilation Contexts, for the second negative response (250–400 ms). The left side displays a 2-D map of the corresponding z-scores (A) while the right side displays the time course of the voltage averaged over the respective electrode clusters for the No Change vs. Place Change conditions (B). In these maps, triangles mark the location of the electrodes selected for statistical analysis. The onset (cf. black arrow) of the MMR was measured from the onset of the first consonant $C_1$ as shown in the example for the test stimulus *ofbe*.

predicted effects (failure to detect a voicing change in a language-specific assimilation context), we thus focused our analysis of the conditions of interest on this second response.

We selected the corresponding electrodes cluster featuring the largest difference ($p < 0.005$ for each channel) between the No Change and the Place Change conditions. These clusters contained 17 channels (including F4, C4, P4 and T8; $t(42) = 3.25$, $p < 0.005$, see the black triangles on Fig. 1). We used the same time window and electrode clusters extracted from this localizer to test for the presence of a mismatch response in each condition of interest.

### 3.2. Analysis of the conditions of interest

The 2 × 2 ANOVA with context (Assimilation vs. No Assimilation) and phonological change (No Change vs. Voicing Change) conducted on the MMR response [250–400 ms] (cf. Fig. 2) yielded no main effect of Context ($F < 1$) but a marginal effect of Phonological Change ($F(1.42) = 2.85$ $p = 0.09$), as well as an interaction between these two factors ($F(1, 42) = 4.75$, $p < 0.05$). This interaction was due to an effect of Phonological Change in the No Assimilation context ($t(42) = 6.82$, $p < 0.05$) but not in the Assimilation context ($t < 1$). Thus, infants detected the voicing change in the No Assimilation context but not in the Assimilation context.

## 4. Discussion

We showed that 14 month-old French-learning infants' capacity to discriminate a native voicing contrast is context-dependent. That is, similarly to French adults (Sun et al., 2015), they exhibit a MMR to a voicing change when the following sound is a nasal (No Assimilation context) but not when the following sound is an obstruent (Assimilation context). This, then, is evidence for a novel type of perceptual attunement. More specifically, the present results show that at 14 months of age, French-learning infants have not only constructed their native phoneme inventory, they have also acquired some knowledge about language-specific interactions among phonemes. Indeed, French voicing assimilation – a phonological process that changes the voicing of a consonant – applies before obstruents but not before nasals; infants have thus acquired this context-dependent neutralization of the voicing contrast.

A straightforward interpretation of our results is that infants compensate for voicing changes in assimilation contexts. That is, given that [ofbe] is usually pronounced as [ovbe], they automatically retrieve the form [ofbe] upon hearing [ovbe]. Direct evidence for such compensation at an older age has been shown with a word recognition task (Skoruppa et al., 2013). In this study, 24-month-old French toddlers were tested in a looking-while-listening paradigm. Target pictures of familiar objects whose name ends in a consonant that can undergo assimilation (e.g., an egg – French: *oeuf* [œf]) were presented side-by-side with distracter pictures of unfamiliar objects. In the probe sentences, the target name was pronounced with a final voicing change either in an assimilation context, i.e. before an obstruent (e.g., *Regarde l'oeu**v** **d**evant toi.* 'Look at the egg in front of you.') or in a no-assimilation context, i.e. before a nasal (e.g., *Regarde l'oeu**v** **m**aintenant.* 'Look at the egg now.'). Toddlers looked longer at the target pictures in the assimilation than in the no-assimilation context. Thus, they more often considered the assimilated form of the target words as a possible instance of these words when the context allowed for voicing assimilation than when it did not. In other words, they showed context-specific compensation for their native process of voicing assimilation (see Darcy et al. (2009), for a similar study with French adults). As shown by the EEG study with adults reported in Sun et al. (2015), compensation for assimilation influences not only word recognition but also phonetic discrimination. Using a subset of Sun et al. (2015)'s stimuli and an adapted paradigm, we provide evidence that this type of context-dependent perceptual attunement is already in place at 14 months of age.

These results leave open two questions concerning *how* and *when* infants acquire native phonological processes and hence become

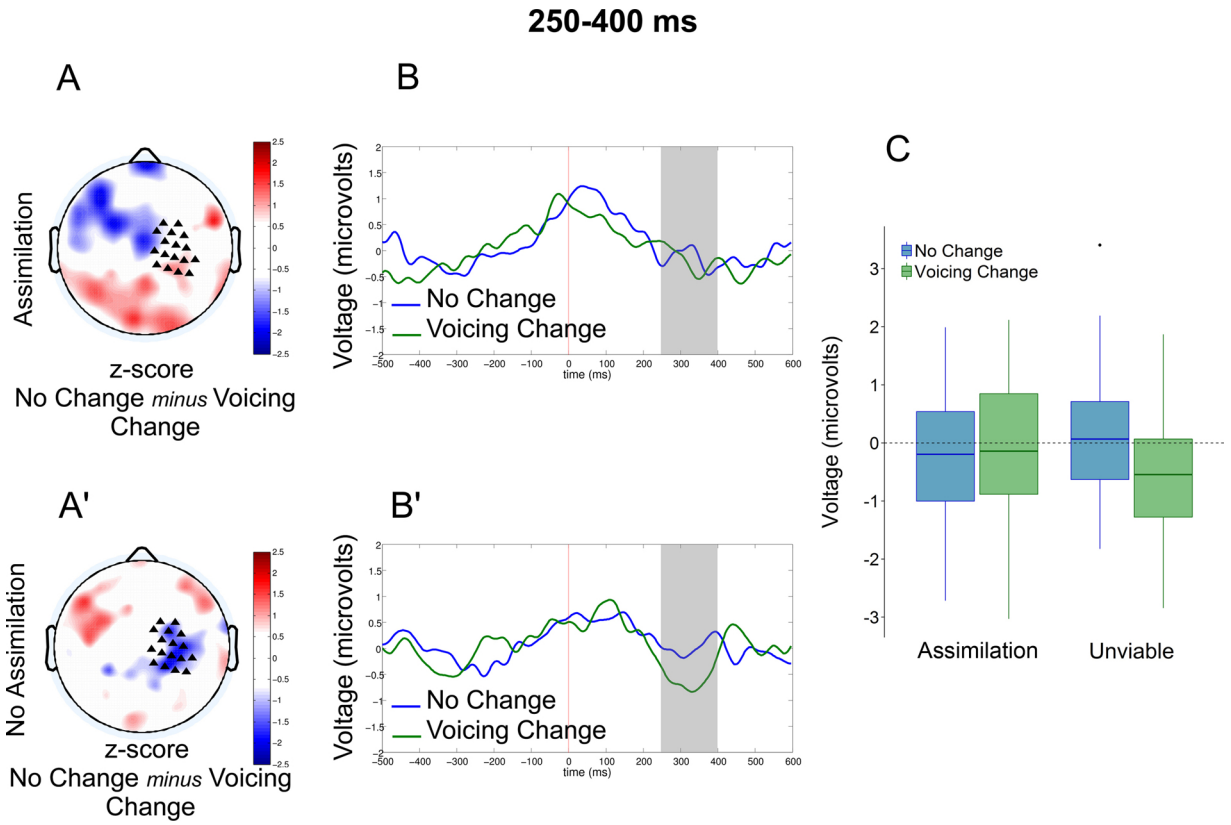


**Fig. 2.** MMR obtained for the comparison No Change *minus* Voicing Change, in the Assimilation Context (legal context for voicing assimilation, A and B) and the No-assimilation Context (illegal context for voicing assimilation, A' and B'), for the time window 250–400 ms. In each panel, the left side displays a 2-D map of the corresponding z-scores (A and A'). In these maps, triangles mark the location of the electrodes selected for statistical analysis. The central columns display the time course of the voltage averaged over the selected electrodes for each of the conditions of interest (B and B'). The rightmost part of the figure displays the graphical representation of this activity averaged for the whole duration of the response (C). Error bars represent standard error of the mean.

insensitive to native sound contrasts in particular contexts. Let us start by considering possible learning mechanisms. For the classic case of perceptual attunement to native sound categories, distributional learning theories (Kuhl et al., 1992; Maye et al., 2008, 2002) postulate that infants attend to the distribution of speech sounds in the acoustic space. As tokens of a given sound category cluster together, this would allow them to discover the speech sound categories used in their language. By taking into account the contexts in which the sounds of their language occur, infants could similarly infer which contrasts are meaningful, and hence distinguish words (Peperkamp et al., 2006). For instance, before knowing any words English-learning infants could infer that the contrast between oral and nasal vowels is not meaningful in their language by observing that nasal vowels occur before a nasal consonant while oral vowels occur in all other contexts. Noticing this complementary distribution, then, would result in loss of sensitivity to the oral-nasal contrast. Distributional learning mechanisms have been argued to be supplemented by top-down information. For instance, infants could benefit from consistent pairing of speech with visible objects (Yeung et al., 2013; Yeung and Werker, 2009), or they might use their knowledge about (real or even approximate) word forms (Feldman et al., 2013a,b; Martin et al., 2013; Swingley, 2009). Could distributional learning, with or without top-down influences, also account for the acquisition of native phonological processes and hence for the context-dependent perceptual attunement evidenced by our results? Note that unlike the English oral and nasal vowels mentioned above, French voiced and voiceless obstruents do not have complementary distributions. Indeed, they can both occur before, say, nasal consonants, where the voicing assimilation never occurs. Due to voicing assimilation, however, voiceless obstruents tend not to occur before voiced obstruents.[5] Observing this distributional fact might be useful in the path towards the acquisition of voicing assimilation, but it is not enough. Indeed, in order for infants to infer that, say, [fb] is rare because it is most often changed into [vb] rather than into something else, they need more information. There are at least two (mutually non-exclusive) sources of top-down information they could use to infer the assimilation process.

On the one hand, infants might rely on word meaning. For instance, knowing that the phrases *oeuf blanc*, [œvblɑ̃] 'white egg' and *oeuf noir* [œfnwar] 'black egg' refer to the same type of object (i.e., eggs) would allow French-learning infants to deduce that [œv] and [œf] are different instances of the same word, and hence that, more generally, a difference in voicing before voiced obstruents does not signal a difference in word meaning. According to this rationale, this specific form of perceptual attunement would depend upon infants' ability to associate auditory labels to a corresponding meaning. While words that can potentially undergo assimilation may be frequent in the infant's environment (approximately 28% of the words contained in the French version of the Communicative Development Inventory, CDI (Fenson et al., 2007) can undergo voicing assimilation),[6] most of these words are not yet known at the age of 14 months: in vocabulary data for a sample of 103 monolingual 14-month-olds (mean number of known words: 74, SD = 64) tested in the same lab as those in the present study,

[5] They sometimes occur in this position, because voicing assimilation is not obligatory (Hallé and Adda-Decker, 2011).

[6] For adjectives, we considered both masculine and feminine forms, and for verbs, all forms of the indicative mode. For the following analyses, we used an augmented version of the corpus analyzed by Ngon et al. (2013), containing over 425,000 word tokens, and considered the full extent of the French voicing assimilation process, including both voicing and devoicing changes (see note 1).

only 13% of the assimilable words contained in the French CDI are reported to be known by more than a quarter of the infants, and only 2% by more than half of them. Thus, an account based on infants' knowledge of word meanings appears unlikely.

On the other hand, infants may rely on a natural clustering of tokens of assimilable words: if both assimilated and non-assimilated forms of the same word often appear within a short period of time (for example, within a short conversation), it might be possible for them to consider them as two instances for the same word. Importantly, this mechanism is independent on infants' knowledge of word meanings. Some support for this hypothesis is provided by data from an orthographically transcribed corpus of spontaneous conversations in the presence of young children, taken from the CHILDES database (MacWhinney, 2000). In this corpus, 7% of the word tokens end in a consonant that can undergo assimilation and of these, 10.2% occur in an assimilation context. Crucially, pairs of tokens of the same word, one occurring in an assimilation and one in a no-assimilation context (within a single recording session), tend to cluster together. That is, nearly 40% of these co-occurrences are found within an estimated two-minute stretch of speech.[7] Moreover, co-occurring tokens of *distinct* words that differ only in their final consonant (e.g. *mousse* [mus] 'foam' vs. *mouche* [muʃ] 'fly') are three times less likely to appear within the same time frame. The clustering of assimilated and unassimilated tokens, then, might be a powerful cue for infants to infer that word-final voicing differences do not signal differences in word meaning, but, rather, reflect systematic variation induced by voicing assimilation.

Finally, as to the question of *when* infants acquire native phonological processes and hence show context-sensitive perceptual attunement, the scenario sketched above relies on word segmentation and sensitivity to the clustering of assimilated and non-assimilated word forms within short stretches of speech. Word segmentation abilities emerge between 6 and 12 months of age (for a review, see Nazzi et al. (2013); we therefore cannot exclude that the acquisition of native phonetic categories and of phonological processes proceed simultaneously, such that context-dependent perceptual attunement evidenced here arises at the same age as 'standard' perceptual attunement, i.e. before the end of the first year of life. Alternatively, the acquisition of phonological processes might lag behind the acquisition of native phonetic categories. We deem this to be more likely, as it takes time not only to fully develop word segmentation abilities but also to gather enough data on clustered tokens of assimilated and non-assimilated word forms. We therefore predict that 12-month-old French-learning infants behave differently from the 14-month-old ones that we tested, and discriminate native voicing contrasts regardless of the context in which they appear.

## 5. Conclusion

To conclude, this research provides evidence for a novel type of context-dependent perceptual attunement. Specifically, we demonstrated that by the age of 14 months, infants' native phoneme categorization does not only rely on the processing of individual sounds, but also takes into account the phonological contexts in which these sounds occur, in a language-specific manner. Further research is needed to determine the precise age range at which this context-dependent perceptual attunement emerges, and to understand the underlying learning mechanism.

## Conflict of Interest:

None.

[7] Based on Hart and Risley (1995); we considered that two minutes of child-directed speech contains on average 42 word tokens.

## Acknowledgements

This study was funded by grants from the Agence Nationale de la Recherche (ANR-13-APPR-0012; ANR-10-LABX-0087 IEC; ANR-10-IDEX-0001-02 PSL*) and the Fondation de France. We thank Ghislaine Deheane-Lambertz for comments on analyses and previous versions of this manuscript. We are grateful to the parents and infants who participated in this study, as well as to Anne-Caroline Fievet and Luce Legros for their help in recruiting and testing participants and Michel Dutat for technical support.

## References

Basirat, A., Dehaene, S., Dehaene-Lambertz, G., 2014. A hierarchy of cortical responses to sequence violations in three-month-old infants. Cognition 132 (2), 137–150. https://doi.org/10.1016/j.cognition.2014.03.013.

Bendixen, A., Jones, S.J., Klump, G., Winkler, I., 2010. Probability dependence and functional separation of the object-related and mismatch negativity event-related potential components. NeuroImage 50 (1), 285–290. https://doi.org/10.1016/j.neuroimage.2009.12.037.

Bertoncini, J., Bijeljac-Babic, R., Jusczyk, P.W., Kennedy, L.J., Mehler, J., 1988. An investigation of young infants' perceptual representations of speech sounds. J. Exp. Psychol. Gen. 117 (1), 21–33. https://doi.org/10.1037/0096-3445.117.1.21.

Bosseler, A.N., Taulu, S., Pihko, E., Mäkelä, J.P., Imada, T., Ahonen, A., Kuhl, P.K., 2013. Theta brain rhythms index perceptual narrowing in infant speech perception. Front. Psychol. 4 (October), 690. https://doi.org/10.3389/fpsyg.2013.00690.

Brainard, D.H., 1997. The psychophysics toolbox. Spat. Vis. 10, 433–436. https://doi.org/10.1163/156856897X00357.

Conboy, B.T., Kuhl, P.K., 2011. Impact of second-language experience in infancy: brain measures of first-and second-language speech perception. Dev. Sci. 14 (2), 242–248. https://doi.org/10.1111/j.1467-7687.2010.00973. x.

Darcy, I., Ramus, F., Christophe, A., Kinzler, K., Dupoux, E., 2009. Phonological knowledge in compensation for native and non-native assimilation. Variation and Gradience in Phonetics and Phonology. pp. 265–309.

Dehaene-Lambertz, G., Baillet, S., 1998. A phonological representation in the infant brain. Neuroreport 9 (8), 1885–1888. https://doi.org/10.1097/00001756-199806010-00040.

Dehaene-Lambertz, G., Dehaene, S., 1994. Speed and cerebral correlates of syllable discrimination in infants. Nature 370 (6487), 292–295. https://doi.org/10.1038/370292a0.

Dehaene-Lambertz, G., Dupoux, E., Gout, A., 2000. Electrophysiological correlates of phonological processing: a cross-linguistic study. J. Cogn. Neurosci. 12 (4), 635–647. https://doi.org/10.1162/089892900562390.

Dehaene-Lambertz, G., 2000. Cerebral specialization for speech and non-speech stimuli in infants. J. Cogn. Neurosci. 12 (3), 449–460. https://doi.org/10.1162/089892900562264.

Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J. Neurosci. Methods 134 (1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009.

Dupoux, E., Hirose, Y., Kakehi, K., Pallier, C., Mehler, J., 1999. Epenthetic vowels in Japanese: a perceptual illusion? J. Exp. Psychol. Hum. Percept. Perform. 25 (6), 1568–1578. https://doi.org/10.1037/0096-1523.25.6.1568.

Eulitz, C., Lahiri, A., 2004. Neurobiological evidence for abstract phonological representations in the mental lexicon during speech recognition. J. Cogn. Neurosci. 16, 577–583. https://doi.org/10.1162/089892904323057308.

Feldman, N.H., Griffiths, T.L., Goldwater, S., Morgan, J.L., 2013a. A role for the developing lexicon in phonetic category acquisition. Psychol. Rev. 120 (4), 751–778. https://doi.org/10.1037/a0034245.

Feldman, N.H., Myers, E.B., White, K.S., Griffiths, T.L., Morgan, J.L., 2013b. Word-level information influences phonetic learning in adults and infants. Cognition 127 (3), 427–438. https://doi.org/10.1016/j.cognition.2013.02.007.

Fenson, L., Marchman, V.A., Dale, P.S., Reznick, J.S., Bates, E., 2007. MacArthur-Bates communicative development inventories. User's Guide and Technical Manual, 2nd ed. Brookes, Baltimore, MD.

Friederici, A.D., Wessels, J.M., 1993. Phonotactic knowledge of word boundaries and its use in infant speech perception. Percept. Psychophys. 54 (3), 287–295. https://doi.org/10.3758/BF03205263.

Hart, B., Risley, T.R., 1995. 42 American families. In Meaningful Differences in the Everyday Experience of Young American Children. pp. 53–74.

Hochmann, J.-R., Papeo, L., 2014. The invariance problem in infancy: a pupillometry study. Psychol. Sci. 25 (11), 2038–2046. https://doi.org/10.1177/0956797614547918.

Jusczyk, P.W., Luce, P.A., 1994. Infants' sensitivity to phonotactic patterns in the native language. J. Mem. Lang. 33 (5), 630–645. https://doi.org/10.1006/jmla.1994.1030.

Jusczyk, P.W., Friederici, A.D., Wessels, J.M.I., Svenkerud, V.Y., Jusczyk, A.M., 1993. Infants' sensitivity to the sound patterns of native language words. J. Mem. Lang. 32, 402–420. https://doi.org/10.1006/jmla.1993.1022.

Kuhl, P.K., Williams, K.A., Lacerda, F., Stevens, K.N., Lindblom, B., 1992. Linguistic experience alters phonetic perception in infants by 6 months of age. Sour. Sci. New Ser. 255 (5044), 606–608. https://doi.org/10.1126/science.1736364.

Lieder, F., Stephan, K.E., Daunizeau, J., Garrido, M.I., Friston, K.J., 2013. A neurocomputational model of the mismatch negativity. PLoS Comput. Biol. 9 (11). https://

doi.org/10.1371/journal.pcbi.1003288.
Luck, S.J., 2005. An introduction to the event-related potential technique. Monogr. Soc. Res. Child Dev. 78 (3), 388. https://doi.org/10.1118/1.4736938.
MacWhinney, B., 2000. The CHILDES project: tools for analyzing talk (third edition): volume I: transcription format and programs, volume II: the database. Computational Linguistics 26 (4), 657. https://doi.org/10.1162/coli.2000.26.4.657.
Martin, A., Peperkamp, S., Dupoux, E., 2013. Learning phonemes with a proto-Lexicon. Cogn. Sci. 37 (1), 103–124. https://doi.org/10.1111/j.1551-6709.2012.01267.x.
Maye, J., Werker, J.F., Gerken, L., 2002. Infant sensitivity to distributional information can affect phonetic discrimination. Cognition 82, 101–111. https://doi.org/10.1016/S0010-0277(01)00157-3.
Maye, J., Weiss, D.J., Aslin, R.N., 2008. Statistical phonetic learning in infants: facilitation and feature generalization. Dev. Sci. 11, 122–134. https://doi.org/10.1111/j.1467-7687.2007.00653. x.
Mazuka, R., Cao, Y., Dupoux, E., Christophe, A., 2011. The development of a phonological illusion: a cross-linguistic study with Japanese and French infants. Dev. Sci. 14 (4), 693–699. https://doi.org/10.1111/j.1467-7687.2010.01015.x.
Mersad, K., Dehaene-Lambertz, G., 2016. Electrophysiological evidence of phonetic normalization across coarticulation in infants. Dev. Sci. 19 (5), 710–722. https://doi.org/10.1111/desc.12325.
Mitterer, H., Blomert, L., 2003. Coping with phonological assimilation in speech perception: evidence for early compensation. Percept. Psychophys. 65 (6), 956–969. https://doi.org/10.3758/BF03194826.
Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., Alho, K., et al., 1997. Language-specific phoneme representations revealed by electric and magnetic brain responses. Nature 385 (6615), 432–434. https://doi.org/10.1038/385432a0.
Näätänen, R., Kujala, T., Escera, C., Baldeweg, T., Kreegipuu, K., Carlson, S., Ponton, C., 2012. The mismatch negativity (MMN)–a unique window to disturbed central auditory processing in ageing and different clinical conditions. Clin. Neurophysiol. Off. J. Int. Fed. Clinic. Neurophysiol. 123 (3), 424–458. https://doi.org/10.1016/j.clinph.2011.09.020.
Ngon, C., Martin, A., Dupoux, E., Cabrol, D., Dutat, M., Peperkamp, S., 2013. (Non)words, (non)words, (non)words: evidence for a protolexicon during the first year of life. Dev. Sci. 16 (1), 24–34. https://doi.org/10.1111/j.1467-7687.2012.01189.x.
Peña, M., Werker, J.F., Dehaene-Lambertz, G., 2012. Earlier speech exposure does not accelerate speech acquisition. J. Neurosci. 32 (33), 11159–11163. https://doi.org/10.1523/JNEUROSCI.6516-11.2012.
Pegg, J.E., Werker, J.F., 1997. Adult and infant perception of two English phones. J. Acoust. Soc. Am. 102 (6), 3742–3753. https://doi.org/10.1121/1.420137.
Peperkamp, S., Le Calvez, R., Nadal, J.-P., Dupoux, E., 2006. The acquisition of allophonic rules: statistical learning with linguistic constraints. Cognition 101, B31–B41.
Polka, L., Werker, J.F., Fraser, J., Phillips, J., O'hara, M., Wolowidnyk, E., Tara, R., et al., 1994. Developmental changes in perception of nonnative vowel contrasts. J. Exp. Psychol.: Hum. Percept. Perform. 20 (2), 421–435. https://doi.org/10.1037/0096-1523.20.2.421.
Rivera-Gaxiola, M., Klarman, L., Garcia-Sierra, A., Kuhl, P.K., 2005a. Neural patterns to speech and vocabulary growth in American infants. Neuroreport 16 (5), 495–498. https://doi.org/10.1097/00001756-200504040-00015.
Rivera-Gaxiola, M., Silva-Pereyra, J., Kuhl, P.K., 2005b. Brain potentials to native and non-native speech contrasts in 7- and 11-month-old American infants. Dev. Sci. 8, 162–172. https://doi.org/10.1111/j.1467-7687.2005.00403.x.
Rivera-Gaxiola, M., Silva-Pereyra, J., Klarman, L., Garcia-Sierra, A., Lara-Ayala, L., Cadena-Salazar, C., Kuhl, P., 2007a. Principal component analyses and scalp distribution of the auditory P 150-250 and N250-550 to speech contrasts in Mexican and American infants. Dev. Neuropsychol. 31 (April), 363–378. https://doi.org/10.1080/87565640701229292.
Rivera-Gaxiola, M., Silva-Pereyra, J., Klarman, L., Garcia-Sierra, A., Lara-Ayala, L., Cadena-Salazar, C., Kuhl, P., 2007b. Principal component analyses and scalp distribution of the auditory P 150-250 and N250-550 to speech contrasts in Mexican and American infants. Dev. Neuropsychol. 31 (3), 363–378. https://doi.org/10.1080/87565640701229292.
Seidl, A., Cristià, A., Bernard, A., Onishi, K.H., 2009. Allophonic and phonemic contrasts in infants' learning of sound patterns. Lang. Learn. Dev. 5 (19), 191–202. https://doi.org/10.1080/15475440902754326.
Skoruppa, K., Mani, N., Plunkett, K., Cabrol, D., Peperkamp, S., 2013. Early word recognition in sentence context: French and English 24-month-olds' sensitivity to sentence-medial mispronunciations and assimilations. Infancy 18 (6), 1007–1029. https://doi.org/10.1111/infa.12020.
Sun, Y., Giavazzi, M., Adda-Decker, M., Barbosa, L.S., Kouider, S., Bachoud-Lévi, A.C., Peperkamp, S., et al., 2015. Complex linguistic rules modulate early auditory brain responses. Brain Lang. 149, 55–65. https://doi.org/10.1016/j.bandl.2015.06.009.
Swingley, D., 2009. Contributions of infant word learning to language development. Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. 364 (1536), 3617–3632. https://doi.org/10.1098/rstb.2009.0107.
Wacongne, C., Changeux, J.-P., Dehaene, S., 2012. A neuronal model of predictive coding accounting for the mismatch negativity. J. Neurosci. 32 (11), 3665–3678. https://doi.org/10.1523/JNEUROSCI.5003-11.2012.
Werker, J.F., Tees, R.C., 1984. Cross-language speech-perception – evidence for perceptual reorganization during the 1st year of life. Infant Behav. Dev. 7 (1), 49–63. https://doi.org/10.1016/S0163-6383(84)80022-3.
Winkler, I., 2007. Interpreting the mismatch negativity. J. Psychophysiol. 21 (3–4), 147–163. https://doi.org/10.1027/0269-8803.21.34.147.
Yeung, H.H., Chen, L.M., Werker, J.F., 2013. Referential labeling can facilitate phonetic learning in infancy. Child Dev. 0 (0), 1–14. https://doi.org/10.1111/cdev.12185.
Yeung, H.H., Werker, J.F., 2009. Learning words' sounds before learning how words sound: 9-month-olds use distinct objects as cues to categorize speech information. Cognition 113 (2), 234–243.