

Six-month-old infants' perception of structural regularities in speech

Irene de la Cruz-Pavía^{a,b,c,*}, Judit Gervain^{c,d,e}

^a Faculty of Social and Human Sciences, University of Deusto, Bilbao, Spain

^b Basque Foundation for Science Ikerbasque, Bilbao, Spain

^c Integrative Neuroscience and Cognition Center, CNRS & Université Paris Cité, Paris, France

^d Department of Social and Developmental Psychology, University of Padua, Padua, Italy

^e Padova Neuroscience Center, University of Padua, Padua, Italy

ARTICLE INFO

Keywords:

Grammar

Prelexical infants

Repetition-based regularities

Diversity-based regularities

NIRS (near-infrared spectroscopy)

ABSTRACT

In order to acquire grammar, infants need to extract regularities from the linguistic input. From birth, infants can detect regularities in speech based on identity relations, and show strong neural activation to syllable sequences containing adjacent repetitions of identical syllables (e.g. ABB: *mubaba*). Meanwhile, newborns' neural responses to sequences of different syllables (e.g. ABC: *mubage*, i.e. diversity-based relations) do not differ from baseline. However, this latter ability needs to emerge during development, as most linguistic units, such as words, are composed of highly variable sequences. As infants begin to learn their first word forms at 6 months, we hypothesize that the ability to represent sequences of different syllables might become important for them at this age. Using near-infrared spectroscopy (NIRS), we measured 6-month-old infants' brain responses to repetition- and diversity-based sequences in the bilateral temporal, parietal and frontal areas. We found that 6-month-olds discriminated the repetition- and diversity-based structures in frontal and parietal regions, and exhibited equally strong activation to both grammars as compared to baseline. These results show that by 6 months of age, infants encode sequences with diversity-based structures. They thus provide the earliest evidence that prelexical infants represent difference in speech stimuli, which behavioral studies first attest at 11 months of age.

1. Introduction

Language acquisition begins prenatally, once infants' hearing becomes operational around the 24th–28th week of gestation (Eggermont & Moore, 2012). Throughout the first year of their lives, infants reach important milestones in their acquisition of the three core components of language. By 12 months of age, unguided and seemingly effortlessly, infants have discovered the sounds of their language (Werker & Tees, 1984), know a handful of content words (Bergelson & Swingle, 2012; Tincoff & Jusczyk, 1999) as well as the most frequent function words (Shi, 2014), and have built a rudimentary representation of a fundamental property of their native syntax, its basic word order (de la Cruz-Pavía, Marino, & Gervain, 2021; Gervain, Nespor, Mazuka, Horie, & Mehler, 2008).

Determining when and how infants acquire the abstract rules of grammar has been a central question for theories of language acquisition. A substantial body of work has tackled this question by investigating young infants' abilities to detect abstract regularities in the linguistic input. Since Marcus, Vijayan, Rao, and Vishton (1999) seminal

study showing that 7-month-old infants rapidly extract a regularity based on the identity relation (A = A) presented in trisyllabic sequences (e.g. AAB: *gagati*, ABA: *gatiga*), a wealth of work has demonstrated prelexical infants' robust ability to detect repetition-based regularities from linguistic (speech and sign; Berent, de la Cruz-Pavía, Brentari, & Gervain, 2021) and non-linguistic inputs (tones, animal sounds, animal pictures, faces, etc.; see de la Cruz-Pavía & Gervain, 2021, for a review, and Rabagliati, Ferguson, & Lew-Williams, 2019, for a recent meta-analysis).

A series of studies tested newborn infants using the brain imaging technique of near infrared spectroscopy (NIRS) to assess when this ability first emerges (Gervain, Berent, & Werker, 2012; Gervain, Macagno, Cogoi, Peña, & Mehler, 2008). Gervain, Macagno, et al. (2008) measured Italian newborns' neural activation while listening to blocks of trisyllabic sequences that either contained an adjacent repetition (e.g. ABB: *mubaba*), or comprised different syllables (e.g. ABC: *mubage*). Both ABB and ABC sequences contain an abstract regularity comprising symbolic variables, and their structures can be described by postulating a rule. ABC, i.e. diversity-based, sequences, have a structure in which its

* Corresponding author at: Faculty of Social and Human Sciences, University of Deusto, Avenida de las Universidades 24, 48007 Bilbao, Spain.

E-mail address: irene.delacruz.pavia@deusto.es (I. de la Cruz-Pavía).

components differ from one another. In turn, ABB, repetition-based sequences have a structure in which two of its syllables repeat. Newborns showed significantly stronger brain activation in response to the sequences containing adjacent repetitions as compared with the diversity-based sequences in bilateral temporal and left inferior frontal regions. Moreover, newborns exhibited robust neural activation to repetition-based sequences as compared to baseline, whereas their activation for the diversity-based sequences only very weakly differed from baseline. This pattern suggests that infants spontaneously encode adjacent repetitions from birth, but lack the ability to represent patterns based on diversity (i.e. in which all elements are different, as in the ABC sequences).

In subsequent studies, Gervain and colleagues showed that newborns also encode adjacent repetitions placed in sequence-initial position (AAB: *mumuba*), exhibiting stronger neural response to AAB tokens as compared to ABC controls (Gervain et al., 2012). By contrast, their automatic encoding of identity-based relations does not extend to non-adjacent repetitions (ABA: *bamuba*), as newborns' neural responses to ABA sequences does not differ significantly from their responses to the ABC controls (Gervain, Macagno, et al., 2008).

Furthermore, newborns can also discriminate sequences containing adjacent repetitions in initial vs. final position (i.e. AAB vs. ABB: *mumuba* vs. *mubaba*; Gervain et al., 2012). If newborns' perceptual system merely detects the occurrence of adjacent repetitions, for instance as if it was a Gestalt-like perceptual primitive, without encoding the sequence's structure, infants should not be able to discriminate between the two sequence types. These results, therefore, show that newborns can encode both the repetition and its sequential position and can thus form a representation that is at least one step more abstract than these two simple stimulus features. Importantly from a methodological perspective, both sequence-initial and -final repetitions trigger robust responses, and when compared directly in a simple block design, no difference is observed between them (Gervain et al., 2012). However, when using an alternating/non-alternating design, successful discrimination was found. In this design, infants hear two types of blocks: blocks containing sequences instantiating a single structure, either AAB or ABB sequences (e.g. ABB: *mubaba*_{ABB1}, *fipepe*_{ABB2}, *totomi*_{ABB3}...), and blocks containing sequences with both structures presented in strict alternation (e.g. ABB-AAB: *mubaba*_{ABB1}, *totomi*_{AAB1}, *fipepe*_{ABB2}...). If infants discriminate the two types of structures, differing neural activation is predicted in response to blocks containing one (i.e. non-alternating blocks) as opposed to both structures (i.e. alternating blocks). This is indeed the result found by Gervain et al. (2012): infants' neural responses were significantly stronger in response to non-alternating blocks as compared with alternating blocks. In order to discriminate these two types of structures, newborns had to detect the adjacent repetition, encode its position and combine these into a single representation, which is thus more abstract than either of the two individual features.

Taken together, these studies revealed a robust ability to encode adjacent repetition at birth, but lack of encoding of diversity-based relations. A single study to date has examined the developmental trajectory of these abilities. Wagner, Fox, Tager-Flusberg, and Nelson (2011) probed infants' detection of repetition- and diversity-based regularities at 7 and 9 months of age, using NIRS and Gervain, Macagno, et al.'s (2008) original ABB and ABC stimuli. Their results suggest that a potential developmental shift in the salience of ABB and ABC grammars might take place throughout the first year of life. Similarly to newborns (Gervain, Macagno, et al., 2008), 7-month-old infants exhibited stronger neural activity than baseline only in response to repetition-based structures, but not for diversity-based structures. By contrast, 9-month-old infants exhibited the reverse pattern, that is, heightened activation only to ABC sequences. While these results hint at a possible change in the encoding of repetition- and diversity-based structures between 7 and 9 months of age, they need to be taken with caution and should be replicated, as sample sizes were quite small (7-month-olds: $n = 13$; 9-month-olds: $n = 15$), and the reported results were found only for deoxyhemoglobin. Oxyhemoglobin showed no significant

differences. Note that effects in infant studies are typically more robust with oxyhemoglobin as compared with deoxyhemoglobin, presumably due to its higher signal to noise ratio (Gervain et al., 2011; Lloyd-Fox, Blasi, & Elwell, 2010). Moreover, Wagner et al. (2011) report no significant differences when the ABB and ABC sequences are directly compared.

The developmental trajectory of infants' ability to encode repetition- and diversity-based structures in linguistic input remains, to a large extent, unexplored. The present study, therefore, seeks to trace potential developmental changes in the detection of these structural regularities. Specifically, we investigate whether by 6 months of age infants represent regularities based on diversity relations. Infants begin to learn their first word forms at 6 months of age, and link them to meaning. Indeed, 6-month-old infants associate the labels "mommy" and "daddy" correctly to their parents (Tincoff & Jusczyk, 1999), and a number of common nouns including body-part or food terms to pictures of their referents (Bergelson & Swingley, 2012). We hypothesize thus that the ability to represent difference might coincide in development with infants' learning of their first words.

To test this hypothesis, we examine 6-month-old infants' neural processing of ABB (i.e. repetition-based) and ABC (i.e. diversity-based) structures using NIRS. If 6-month-old infants have begun to represent sequences of different syllables, in addition to sequences of identical syllables, then (1) infants' neural response to both repeated and diversity-based sequences should be significantly stronger than baseline, and (2) infants should discriminate the two types of structures. Following Gervain et al. (2012), we choose an alternating/non-alternating design, because if 6-month-olds are able to represent both structures and have equally strong neural responses to them, a simple block design would not reveal any differences between them. Indeed, Wagner et al. (2011) used a simple block design and found no differences in brain activation between the ABB and ABC grammars, either at 7 or 9 months. By contrast, an alternating/non-alternating design allows us to: (1) test infants' discrimination of the two structures by comparing their neural activation in the alternating vs. the non-alternating blocks, as well as (2) to directly assess infants' representations of the two structures by investigating responses in non-alternating blocks and comparing them to baseline, similarly to what the simple block design allows.

2. Materials and methods

2.1. Participants

24 French-learning infants (10 females; mean age: 5 months 28 days; range: 5 months 14 days – 6 months 15 days) contributed to the final analyses. All infants were born full-term and exposed to French at least 80% of the time. Twenty-six additional infants were tested but excluded from the final analysis due to fussiness and crying (10), an insufficient number of valid trials / poor data quality (14), parental interference (1), and equipment failure (1). Rejection due to poor data quality was performed in batch, following the same criteria for all infants (see Data Processing and Analysis), and prior to statistical analysis. Parents gave informed consent before participation. The study was approved by the CERES ethics board (Université Paris Cité, France).

2.2. Stimuli and design

Following Gervain and colleagues (Gervain et al., 2012; Gervain, Macagno, et al., 2008), we created an inventory of 20 Consonant-Vowel (CV) syllables and sorted them into 10 syllable pairs (Fig. 1). In each pair, the two syllables had different vowels but contained either the same consonant (e.g. *pe* – *pi*) or at least consonants from the same class (e.g. nasal: *mu* – *na*). We designated one of the members of each pair as syllable A, the other member as syllable B. The 20 syllables were thus split evenly into A and B syllables. In order to create the repetition

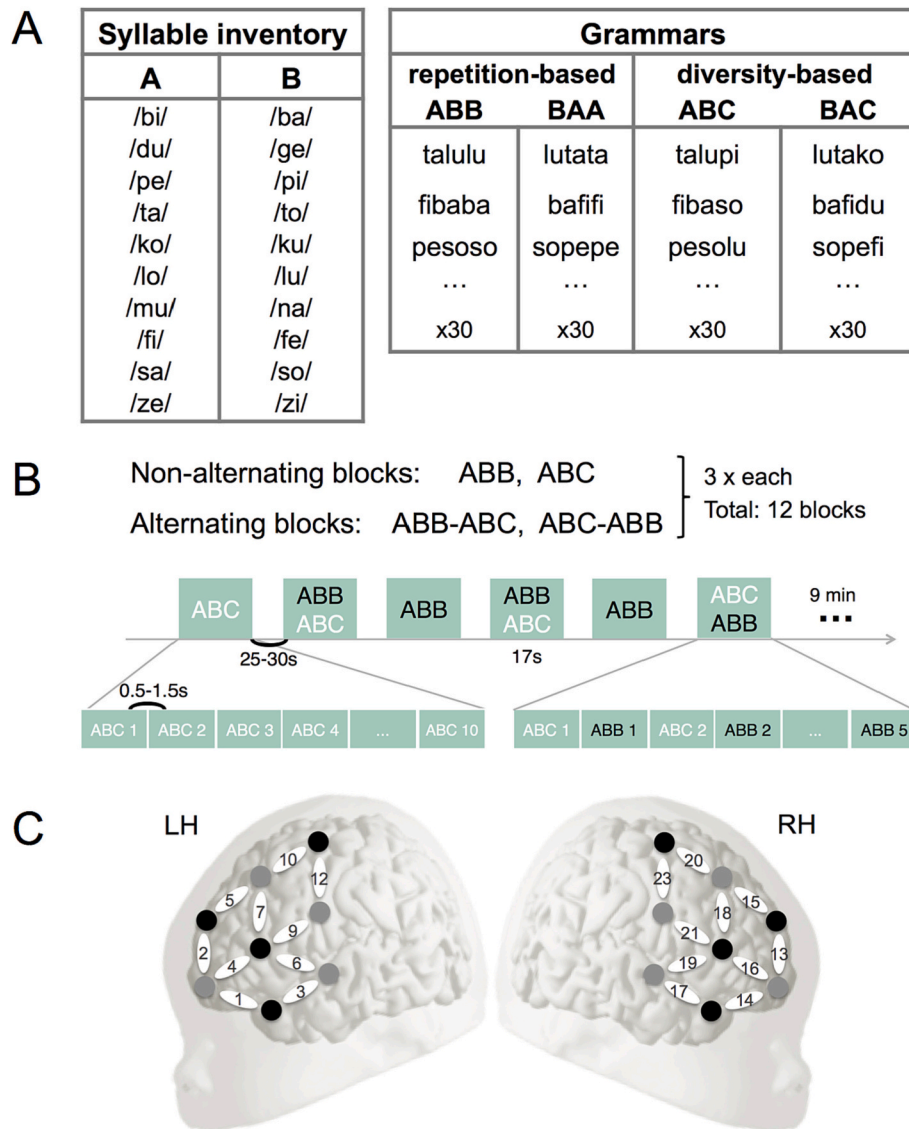


Fig. 1. (A) Stimuli: syllable inventory (left) and grammars (right). (B) Experimental design and procedure. (C) Layout of NIRS channels. For the sake of cross-study comparability, channel numbers correspond to those reported in Gervain et al. (2008, 2012). Note that channel numbers are not continuous, as in the present study we recorded 20 channels, a subset of the 24 recorded by Gervain et al. (2012) and Gervain, Macagno, et al. (2008).

grammar we combined them into trisyllabic sequences with an ABB structure, that is, containing an adjacent, sequence-final repetition of the B syllable (Fig. 1). Each ABB sequence consisted of an A and B syllable that did not belong to the same pair so as to maximize discriminability, and did not contain the same vowel (e.g. *pe – so*). We paired each of the 10 A syllables with 3 different B syllables (e.g. *pe – na/lu/so*), yielding a total of 30 ABB sequences. We then created a BAA counterpart for each ABB sequence by switching syllable order (e.g. ABB: *pesoso* became BAA: *sopepe*), resulting in a total of 60 repetition-based sequences. As a result, A and B syllables occurred equally frequently in all possible positions within a sequence.

The diversity-based grammar also consisted of 60 sequences, directly derived from the repetition grammar by shuffling around the third syllables of the sequences: 30 of the sequences were derived from ABB tokens and had an ABC structure; the remaining 30 sequences were derived from BAA sequences and had a BAC structure. Specifically, ABC sequences resulted from the fusion of two ABB sequences, e.g. the ABC sequence *penaku* was created by replacing the last syllable of the ABB sequence *penana* with the last syllable of the ABB sequence *zekuku*. Similarly, BAC sequences resulted from the fusion of two BAA sequences

(e.g. the BAC sequence *napefi* resulted from the replacement of the last syllable of the BAA sequence *napepe* with the last syllable of the sequence *bafifi*).

The full set of stimuli thus consisted of 120 sequences: 30 ABB, 30 BAA, 30 ABC, 30 BAC (see Appendix A for the full list of sequences). Following Gervain et al. (2012) and Gervain, Macagno, et al. (2008), we synthesized the sequences using the fr4 French female voice of the MBROLA database (Dutoit, 1997), at a constant pitch of 200 Hz and syllable duration of 270 ms (consonant: 120 ms, vowel: 150 ms; see the Supporting Information for sample sound files). The repetition- and diversity-based grammars were identical in the overall frequency of the syllables, the frequency of each syllable in all possible positions, and in all phonological and prosodic features. Moreover, the distribution of transitional probabilities (TPs), i.e. the likelihood of one syllable following another, was also equated in the two grammars.

The 120 sequences were sorted into 12 blocks, each with exactly 10 sequences. We chose an alternating/non-alternating design (Fig. 1). Half of the blocks contained sequences of a single grammar (either repetition- or diversity-based), whereas the remaining half contained sequences of the two grammars in strict alternation. This design is frequently used to

test fine-grained discrimination in behavioral measures (Best & Jones, 1998; Maye, Werker, & Gerken, 2002) as well as with NIRS (Gervain et al., 2012; Sato, Sogabe, & Mazuka, 2009).

Infants heard 6 alternating and 6 non-alternating blocks presented in strict alternation. Three of the alternating blocks began with a repetition-based sequence, the other 3 with a diversity-based sequence. Similarly, 3 of the non-alternating blocks contained sequences from the repetition grammar, the remaining 3 blocks from the diversity-based grammar. Order of presentation of tokens within blocks and order of presentation of the 12 blocks were randomized and counterbalanced across infants. Half of the infants heard an alternating block first, the other half a non-alternating block first. Within blocks, sequences were separated by pauses of varying length (0.5 or 1.5 s), yielding blocks of about 17 s (Fig. 1). Blocks were also separated by pauses of varying length (25 or 30 s), to avoid phase-locked brain responses. The whole experiment lasted around 9 min.

Infants watched a silent attention-getter cartoon to maintain their attention and reduce movement artifacts (Obrig et al., 2017).

2.3. Procedure

Infants were tested in a quiet, dimly lit testing booth, seated on a caregiver's lap. During the testing session, infants were presented with a silent video of a cartoon that provided no meaningful content. The cartoon was displayed on a screen placed in front of the infants at approximately 80 cm. The sound stimuli were administered through two loudspeakers positioned at both sides of the screen. A computer played the cartoon and the sound stimuli through *E-Prime*, and another computer recorded the NIRS signal.

Infants' neural responses to the alternating and non-alternating blocks were measured with a NIRx NIRScout 8–16 machine (source-detector separation: 3 cm; wavelengths of 760 nm and 850 nm; sampling rate: 15.625 Hz). The optical sensors were inserted into a stretchy cap and placed on the infants' head using surface landmarks (nasion and the preauricular points). We recorded activity in the bilateral temporal, frontal and parietal cortices (10 channels per hemisphere; Fig. 1C), that is, the language areas shown to respond to repetition grammars in newborns (Gervain et al., 2012; Gervain, Macagno, et al., 2008) and in 6-month-olds (Berent et al., 2021). Localization analysis was performed for this age range in Berent et al. (2021) and revealed that channels 1, 2, 4, 5 and 13–16 query the frontal lobe, channels 3, 8, 11, 17, 22, and 24 the temporal lobe, channels 10, 12, 20 and 23 are parietal, whereas channels 6, 7, 9, 18, 19 and 21 span two lobes.

2.4. Data processing and analysis

The NIRS machine measured the absorption of red and near-infrared light, from which the changes in concentration of oxygenated hemoglobin (oxyHb) and deoxygenated hemoglobin (deoxyHb) were calculated using the modified Beer-Lambert Law. We band-pass filtered the data between 0.01 and 0.7 Hz to eliminate noise (e.g. heartbeat) and overall trends. We then removed movement artifacts, defined as concentration changes larger than 0.1 mmol*mm over 200 milliseconds, by rejecting block-channel pairs in which artifacts occurred. For valid, non-rejected blocks, a baseline was linearly fitted between the means of the 5 s preceding the onset of the block and the 5 s starting 15 s after offset of the block. This 15-s-long resting period is essential to allow the metabolic response to return to baseline before the next simulation. This preprocessing routine has been shown to yield accurate recovery of the infant hemodynamic response (Gemignani & Gervain, 2021), and has been used in a number of studies (Berent et al., 2021; Gemignani et al., 2023; Martínez-Alvarez, Benavides-Varela, Lapillonne, & Gervain, 2021; Martínez-Alvarez, Gervain, Koulaguina, Pons, & de Diego-Balaguer, 2023). Only infants that had at least 33% valid data (minimally 2 trials per condition) were included in the analysis. The 24 infants included in the final sample provided on average 64% valid data.

We carried out statistical analyses of the two metabolic indicators of neural activity, oxyHb and deoxyHb. We conducted cluster-based permutation analyses (Maris & Oostenveld, 2007) comparing each condition to baseline, and the two conditions between them. This analysis identifies spatially adjacent channels in which significant activation is observed in temporally adjacent samples. By identifying regions and time windows of interest in a non-arbitrary, data-driven fashion, permutation tests avoid the multiple comparison problem. Since the permutation test serves to identify time windows and channels with a significant activation, no a priori assumptions need to be made about the channels and time windows of interest, thus all channels and the full time series of blocks (i.e. full stimulation period plus 15-s resting period) were entered into the analyses. To perform the permutation test, we used paired-sample *t*-tests with $t = 2$ as threshold for significance, and ran 1000 permutations under the null hypothesis. This analysis is standard procedure in the NIRS literature (Cabrera & Gervain, 2020; Berent et al., 2021; Ferry et al., 2016; Mahmoudzadeh et al., 2013; amongst others).

3. Results

3.1. Discrimination of repetition- and diversity-based structures

A cluster-based permutation test comparing infants' responses to the alternating and non-alternating blocks revealed greater activation in oxyHb in response to non-alternating blocks in right frontal areas, specifically, a cluster containing channels 13 and 16 in the RH ($t(23) = 640$, $p < .001$; Fig. 2). A similar permutation test over deoxyHb also revealed greater activation in response to non-alternating blocks in the left frontal areas, specifically channel 4 ($t(23) = 190$, $p < .001$), and the right fronto-parietal areas, specifically channel 15 ($t(23) = 285$, $p < .001$). These results indicate that infants discriminated the blocks containing only one vs. both types of structures.

Further cluster-based permutation analyses showed that infants exhibited significantly greater activation than baseline in response to both alternating and non-alternating blocks in oxyHb as well as deoxyHb. The statistical details of these analyses are reported in Appendix B.

3.2. Encoding of repetition- and diversity-based structures: analysis of non-alternating blocks

In order to determine how infants encoded the two grammars, we compared activation for each grammar in the non-alternating blocks to baseline. We also compared infants' responses to the two grammars directly.

OxyHb. Permutation analyses revealed clusters of activation significantly greater than baseline in both hemispheres for both grammars (i.e. repetition-based grammar: clusters containing channels 3, 6, and 9 in the LH, $t(23) = 1045$, $p < .001$, and channels 14, 16 and 17 in the RH, $t(23) = 601$, $p < .001$; diversity-based grammar: channels 1, 3 and 4 in the LH, $t(23) = 366$, $p < .001$, and channels 13, 14, 16, 18 and 21 in the RH, $t(23) = 1917$, $p < .001$). An additional permutation analysis revealed no significant differences in the direct comparison of the repetition- and diversity-based grammars.

DeoxyHb. Permutation analyses revealed clusters of activation significantly lower than baseline in both hemispheres, although only in response to the repetition-based grammar (i.e. clusters containing channels 1, 3, 4, 6, and 9 in the LH, $t(23) = 2062$, $p < .001$, and channels 13, 14, 16 and 18 in the RH, $t(23) = 1223$, $p < .001$). No significant differences emerged in the direct comparison of the repetition- and diversity-based grammars.

4. Discussion

Newborn infants have a robust ability for encoding sequences containing identity-based relations (i.e. $A = A$), and specifically adjacent

Non-Alternating vs. Alternating blocks

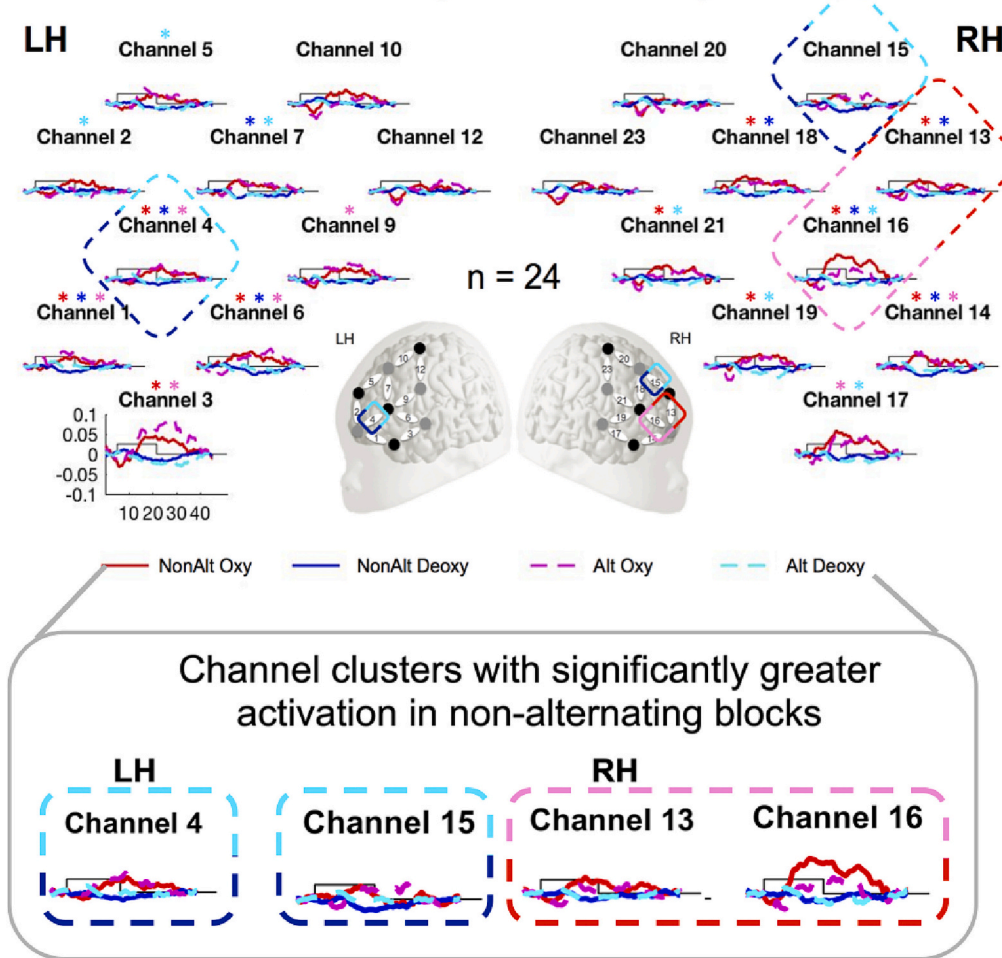
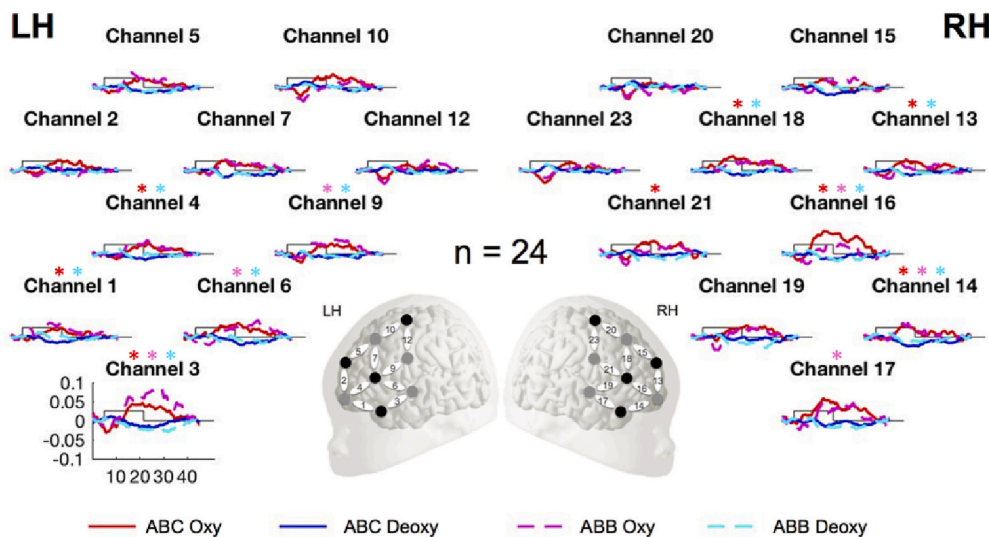


Fig. 2. Grand average responses. The figure depicts grand average hemodynamic responses to non-alternating vs. alternating blocks (upper plot), and to the repetition- (ABB) vs. diversity-based (ABC) grammars in non-alternating blocks (lower plot). Channels are plotted following the placement depicted in Fig. 1C. The x-axis shows time in seconds, with the rectangle indicating the window of stimulation. The y-axis represents concentration in mmol³mm. Continuous red and blue lines depict oxyHb and deoxyHb concentrations, respectively, in response to the non-alternating blocks in the upper plot, and to the ABC blocks in the lower plot. The dashed magenta and cyan curves represent oxyHb and deoxyHb concentrations, respectively, in response to the alternating blocks in the upper plot, and to ABB blocks in the lower plot. Asterisks placed above individual channels indicate significant differences between the condition indicated by the asterisk's color vs. baseline. Upper plot: the red-pink box encircles the cluster of channels in which oxyHb concentration significantly differed between alternating and non-alternating blocks, while the dark blue-light blue boxes depict channels with significantly different deoxyHb concentrations in alternating and non-alternating blocks. Channel clusters with significantly greater activation in non-alternating blocks are shown enlarged below. Lower plot: direct comparison of activation in response to ABB and ABC blocks yielded no significant differences in oxyHb or deoxyHb concentration. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Non-Alternating blocks: ABB vs. ABC blocks



repetitions (AAB, ABB), but do not encode structures comprised of different syllables (ABC), i.e. they do not yet represent diversity-based regularities (Gervain et al., 2012; Gervain, Macagno, et al., 2008). We hypothesized that this ability might emerge by 6 months of age, time at

which infants begin to learn their first word forms (Bergelson & Swingley, 2012; Tincoff & Jusczyk, 1999), readily learning sequences of different syllables as possible word form candidates, as well as the foundations of grammar (Gervain, Nespor, et al., 2008). To test this

hypothesis, we presented 6-month-old infants with trisyllabic sequences containing an adjacent repetition (ABB: *mubaba*, repetition-based regularity), and sequences containing three different syllables (ABC: *mubage*, diversity-based regularity), and measured neural activation in response to the two grammars using NIRS in an alternating/non-alternating design.

We found that infants showed stronger activation in response to blocks containing tokens from a single grammar (only ABB or ABC sequences) as compared with blocks containing alternating sequences from both grammars (ABB and ABC). Infants can thus discriminate the repetition- and diversity-based grammars, as predicted. This differential response was localized to the frontal regions and showed a pattern, similar to that reported by Gervain et al. (2012): greater activation in response to non-alternating than to alternating blocks in the frontal regions. These results are in line with the general architecture of the language network, as frontal areas have been shown to be responsible for sequence processing and the computation of higher order structure already in preverbal infants, similarly to adults (Dehaene-Lambertz et al., 2006; Dehaene-Lambertz, Dehaene, & Hertz-Pannier, 2002; Friederici, 2002). Non-alternating blocks may trigger stronger activation than alternating ones, because they allow the extraction of a clear regularity, whereas alternating blocks provide conflicting information, and/or the extraction of two regularities may be too complex for infants. In general, 6-month-olds' ability to discriminate between ABB and ABC patterns is expected, as newborns can already perform this discrimination.

Importantly from the point of view of our hypothesis and unlike in newborns, however, infants also exhibited significant neural activation to both ABB and ABC regularities when presented in the non-alternating blocks, showing that infants encoded both grammars. This activation occurred primarily in temporal and frontal regions bilaterally. Furthermore, when compared directly, infants' neural responses to the two grammars did not differ.

While both findings separately could be explained without assuming that 6-month-olds can represent diversity-based structures, taken together, they rule out these other potential explanations. Thus, it could be argued that the similar activation found in response to the repetition- vs. diversity-based structures presented in non-alternating blocks results from the auditory processing of the sequences, i.e. just a general response to speech sounds, without encoding their structure. However, if infants' response to the two sequence types was auditory only, they would exhibit similar activation in alternating vs. non-alternating blocks — i.e. they would not discriminate between them — unlike what was found in the present study. It could also be argued that infants discriminated between the two structures in the alternating vs. non-alternating blocks by encoding only repetitions. However, a processing advantage — i.e. greater activation — for repetition would then be predicted, as found in newborns (Gervain et al., 2008). By contrast, infants exhibit similar neural responses to repetition- and diversity-based structures. Taken together, these results thus highlight an important developmental change. While strong activation was already observed at birth to repetition grammars, diversity-based grammars only trigger robust activation, one that is similar to repetition-based grammars, at 6 months, but not yet at birth.

Prior to the present research, a single study had examined the developmental trajectory of infants' ability to encode ABB and ABC structures after birth. Wagner et al. (2011) presented 7- and 9-month-old infants with a simple block design in which blocks contained tokens from either the ABB or the ABC grammar. Their results are only partially replicated in the present study. Thus, Wagner et al. (2011) find no differential activation to the two grammars at 7 or at 9 months of age, similarly to the present study. By contrast, Wagner et al. (2011) report stronger neural activation than baseline only for the repetition grammar at 7 months, and only for the diversity-based grammar at 9 months. It

needs to be noted, however, that these effects are only observed in deoxyHb and the study is likely underpowered. In the present study, we observe equally strong activation to the two grammars. The origin of this discrepancy is likely Wagner et al.'s (2011) very reduced sample sizes (7-month-olds: $n = 13$, 9-month-olds: $n = 15$) as compared with the present study ($n = 24$).

In sum, the present study shows that by six months of age infants encode a diversity-based pattern in speech stimuli. Behavioral studies first attest this ability at 11 months of age. Hochmann and Toro (2021) presented 11-month-olds with sequences of 3, 4, or 5 syllables, which consisted of the same syllable repeated, except for the last one in the sequence (i.e. AA-A, AAA-A, AAAA-A, e.g. AAAA-A: *fifififilo*). They intermixed these with sequences comprising four identical syllables (i.e. AAAA, e.g. *babababa*) at a 3:1 ratio. That is, while 75% of the trials comprised a different final syllable (*standard* trials), the remaining 25% consisted of sequences containing only repeated syllables (*deviant* trials). Infants showed larger pupil dilation in response to the infrequent sequences of repeated syllables as compared with the frequently-occurring sequences containing a different final syllable, suggesting that infants expected the sequence to end with a syllable that differed from the previous ones, and were surprised when it was absent. This finding shows that the 11-month-old infants encoded the abstract structure of the sequences and represented the relation *different*, as found in the present brain imaging study with 6-month-old infants. Note that other behavioral studies probing infants' ability to detect diversity-based rules have failed to detect this ability even at 12 months of age, both with linguistic (Kovács, 2014) and non-linguistic stimuli (e.g. geometrical shapes; Hochmann, Carey, & Mehler, 2018).

In conclusion, our results provide the earliest evidence that young infants represent difference in speech stimuli, and the emergence of this ability coincides with the time at which infants begin to learn their first word forms. While we hypothesize that the ability to represent difference might coincide in development with infants' learning of their first words, we do not claim a causal relation between the two, as such a broad implication falls outside the scope of the present study and remains to be tested in future studies.

CRediT authorship contribution statement

Irene de la Cruz-Pavía: Conceptualization, Data curation, Formal analysis, Funding acquisition, Writing – original draft, Writing – review & editing. **Judit Gervain:** Conceptualization, Formal analysis, Funding acquisition, Resources, Supervision, Writing – review & editing.

Declaration of Competing Interest

None.

Data availability

Data will be made available on request.

Acknowledgements

This research was supported by the Agence Nationale de la Recherche (ANR) [SpeechCode—ANR-15-CE37-0009-01], the ANR's French Investissements d'Avenir – Labex EFL Program [ANR-10-LABX-0083], the European Research Council [Consolidator 773202 ERC-2017-COG 'BabyRhythm'], the ECOS-Sud grant nr. C20S02, the European Union - Next Generation EU - NRRP M6C2 - Investment 2.1 Enhancement and strengthening of biomedical research in the NHS, and the FARE Project [R204MPRHKE] to Judit Gervain, and the Spanish Ministry of Science and Innovation [PID2019-105100RJ-I00 and RYC2021-033951-I], and the Basque Foundation for Science Ikerbasque to Irene de la Cruz-Pavía.

Appendix A. Inventory of the 120 trisyllabic sequences

ABB	BAA	ABC	BAC
talulu	lutata	talupi	lutako
penana	napepe	penaku	napefi
bisoso	sobibi	bisoge	sobize
kofefe	fekoko	kofezi	fekomu
duzizi	zidudu	duzito	zidusa
mugege	gemumu	mugeba	gemuta
fibaba	bafifi	fibaso	bafidu
zepipi	pizeze	zepina	pizelo
satoto	tosasa	satolu	tosape
lokuku	kulolo	lokufe	kulobi
mubaba	bamumu	mubage	bamuko
binana	nabibi	binaso	nabipe
zekuku	kuzeze	zekuzi	kuzebi
lozizi	zilolo	lozife	zilomu
dufefe	fedudu	dufeto	fedusa
pelulu	lupepe	pelupi	lupefi
fitoto	tofiifi	fitoba	tofidu
tasoso	sotata	tasoku	sotaze
sapipi	pisasa	sapilu	pisalo
kogege	gekok	kogena	gekota
sakuku	kusasa	sakufe	kusape
bitoto	tobibi	bitona	tobisa
muzizi	zimumu	muziba	zimuta
kolulu	lukoko	kolupi	lukobi
figege	gefifi	figeso	gefidu
lopiipi	pilolo	lopige	piloze
pesoso	sopepe	pesolu	sopefi
zebaba	bazeze	zebato	bazeko
tafefe	fetata	tafeku	fetamu
dunana	nadudu	dunazi	nadulo

Appendix B. Results of the permutation analyses comparing infants' neural responses to alternating and non-alternating blocks with the silent baseline

OxyHb. A permutation analysis revealed clusters of activation significantly greater than baseline, present in both hemispheres and in response to both alternating and non-alternating blocks (i.e. alternating blocks: clusters containing channels 1, 3, 4, 6, and 9 in the LH, $t(23) = 3497$, $p < .001$, and channels 14 and 17 in the RH, $t(23) = 391$, $p < .001$; non-alternating blocks: channels 1, 3, 4 and 6 in the LH, $t(23) = 1469$, $p < .001$, and channels 13, 14, 16, 18, 19 and 21 in the RH, $t(23) = 2531$, $p < .001$).

DeOxyHb. Analysis of deoxyHb revealed a comparable picture. Thus, permutation tests revealed clusters of activation significantly lower than baseline, present in both hemispheres and in response to both alternating and non-alternating blocks (i.e. alternating blocks: clusters containing channels 2, 5 and 7 in the LH, $t(23) = 486$, $p < .001$, and channels 16, 17, 19 and 21 in the RH, $t(23) = 916$, $p < .001$; non-alternating blocks: channels 1, 4, 6 and 7 in the LH, $t(23) = 1053$, $p < .001$, and channels 13, 14, 16 and 18 in the RH, $t(23) = 952$, $p < .001$).

Appendix C. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.cognition.2023.105526>.

References

- Berent, I., de la Cruz-Pavía, I., Brentari, D., & Gervain, J. (2021). Infants differentially extract rules from language. *Scientific Reports*, *11*(20001), 10. <https://doi.org/10.1038/s41598-021-99539-8>
- Bergelson, E., & Swingle, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences*, *109*(9). <https://doi.org/10.1073/pnas.1113380109>. Article 9.
- Best, C. T., & Jones, C. (1998). Stimulus-alternation preference procedure to test infant speech discrimination. *Infant Behavior & Development*, *21*, 295. [https://doi.org/10.1016/S0163-6383\(98\)91508-9](https://doi.org/10.1016/S0163-6383(98)91508-9)
- Cabrera, L., & Gervain, J. (2020). Speech perception at birth: The brain encodes fast and slow temporal information. *Science Advances*, *6*(30). eaba7830.
- de la Cruz-Pavía, I., & Gervain, J. (2021). Infants' perception of repetition-based regularities in speech: A look from the perspective of the same/different distinction. *Current Opinion in Behavioral Sciences*, *37*, 125–132. <https://doi.org/10.1016/j.cobeha.2020.11.014>
- de la Cruz-Pavía, I., Marino, C., & Gervain, J. (2021). Learning word order: Early beginnings. *Trends in Cognitive Sciences*, *25*(9), 802–812. <https://doi.org/10.1016/j.tics.2021.04.011>
- Dehaene-Lambertz, G., Dehaene, S., & Hertz-Pannier, L. (2002). Functional neuroimaging of speech perception in infants. *Science*, *298*(5600). <https://doi.org/10.1126/science.1077066>. Article 5600.
- Dehaene-Lambertz, G., Hertz-Pannier, L., Dubois, J., Meriaux, S., Roche, A., Sigman, M., & Dehaene, S. (2006). Functional organization of perisylvian activation during presentation of sentences in preverbal infants. *Proceedings of the National Academy of Sciences of the United States of America*, *38*. <https://doi.org/10.1073/pnas.0606302103>. Article 38.
- Dutoit, T. (1997). *An introduction to text-to-speech synthesis* (Vol. 3). Kluwer Academic Publishers.
- Eggermont, J. J., & Moore, J. K. (2012). Morphological and functional development of the auditory nervous system. In *Human auditory development* (pp. 61–105). Springer. http://link.springer.com/chapter/10.1007/978-1-4614-1421-6_3.
- Ferry, A. L., Fló, A., Brusini, P., Cattarossi, L., Macagno, F., Nespor, M., & Mehler, J. (2016). *Developmental Science*, *19*, 488–503.
- Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, *6*(2). [https://doi.org/10.1016/s1364-6613\(00\)01839-8](https://doi.org/10.1016/s1364-6613(00)01839-8). Article 2.
- Gemignani, J., de la Cruz-Pavía, I., Martínez, A., Nallet, C., Pasquini, A., Lucarini, G., ... Gervain, J. (2023). The reproducibility of infant fNIRS studies: A meta-analytic approach. *Neurophotonics*, *10*(2), Article 023518.
- Gemignani, J., & Gervain, J. (2021). Comparing different pre-processing routines for infant fNIRS data. *Developmental Cognitive Neuroscience*, *48*, Article 100943.

- Gervain, J., Berent, I., & Werker, J. F. (2012). Binding at birth: The newborn brain detects identity relations and sequential position in speech. *Journal of Cognitive Neuroscience*, 24(3). https://doi.org/10.1162/jocn_a.00157. Article 3.
- Gervain, J., Macagno, F., Cogoi, S., Peña, M., & Mehler, J. (2008). The neonate brain detects speech structure. *Proceedings of the National Academy of Sciences*, 105(37). <https://doi.org/10.1073/pnas.0806530105>. Article 37.
- Gervain, J., Mehler, J., Werker, J. F., Nelson, C. A., Csibra, G., Lloyd-Fox, S., ... Aslin, R. N. (2011). Near-infrared spectroscopy: A report from the McDonnell infant methodology consortium. *Developmental Cognitive Neuroscience*, 1, 22–46.
- Gervain, J., Nespore, M., Mazuka, R., Horie, R., & Mehler, J. (2008). Bootstrapping word order in prelexical infants: a Japanese–Italian cross-linguistic study. *Cognitive Psychology*, 57(1). <https://doi.org/10.1016/j.cogpsych.2007.12.001>. Article 1.
- Hochmann, J.-R., Carey, S., & Mehler, J. (2018). Infants learn a rule predicated on the relation same but fail to simultaneously learn a rule predicated on the relation different. *Cognition*, 177, 49–57. <https://doi.org/10.1016/j.cognition.2018.04.005>
- Hochmann, J.-R., & Toro, J. M. (2021). Negative mental representations in infancy. *Cognition*, 213, Article 104599. <https://doi.org/10.1016/j.cognition.2021.104599>
- Kovács, Á. M. (2014). Extracting regularities from noise: Do infants encode patterns based on same and different relations? *Language Learning*, 64(s2). <https://doi.org/10.1111/lang.12056>. Article s2.
- Lloyd-Fox, S., Blasi, A., & Elwell, C. E. (2010). Illuminating the developing brain: The past, present and future of functional near infrared spectroscopy. *Neuroscience and Biobehavioral Reviews*, 34, 269–284.
- Mahmoudzadeh, M., Dehaene-Lambertz, G., Fournier, M., Kongolo, G., Goudjil, S., et al. (2013). Syllabic discrimination in premature human infants prior to complete formation of cortical layers. *Proceedings of the National Academy of Sciences*, 110(12), 4846–4851.
- Marcus, G. F., Vijayan, S., Rao, S. B., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science*, 283(5398). <https://doi.org/10.1126/science.283.5398.77>. Article 5398.
- Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, 164(1). <https://doi.org/10.1016/j.jneumeth.2007.03.024>. Article 1.
- Martínez-Alvarez, A., Benavides-Varela, S., Lapillonne, A., & Gervain, J. (2021). Newborns discriminate utterance-level prosodic contours. *Developmental Science*, 26(2), Article e13304.
- Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, 82(3). [https://doi.org/10.1016/s0010-0277\(01\)00157-3](https://doi.org/10.1016/s0010-0277(01)00157-3). Article 3.
- Obriq, H., Mock, J., Stephan, F., Richter, M., Vignotto, M., & Rossi, S. (2017). Impact of associative word learning on phonotactic processing in 6-month-old infants: A combined EEG and fNIRS study. *Developmental Cognitive Neuroscience*, 25, 185–197. <https://doi.org/10.1016/j.dcn.2016.09.001>
- Rabagliati, H., Ferguson, B., & Lew-Williams, C. (2019). The profile of abstract rule learning in infancy: Meta-analytic and experimental evidence. *Developmental Science*, 22(1). <https://doi.org/10.1111/desc.12704>. Article 1.
- Sato, Y., Sogabe, Y., & Mazuka, R. (2009). Development of hemispheric specialization for lexical pitch-accent in Japanese infants. *Journal of Cognitive Neuroscience*. <https://doi.org/10.1162/jocn.2009.21377>
- Shi, R. (2014). Functional morphemes and early language acquisition. *Child Development Perspectives*, 8(1). <https://doi.org/10.1111/cdep.12052>. Article 1.
- Tincoff, R., & Jusczyk, P. W. (1999). Some beginnings of word comprehension in 6-month-olds. *Psychological Science*, 10(2). <https://doi.org/10.1111/1467-9280.00127>. Article 2.
- Wagner, J. B., Fox, S. E., Tager-Flusberg, H., & Nelson, C. A. (2011). Neural processing of repetition and non-repetition grammars in 7- and 9-month-old infants. *Frontiers in Psychology*, 2. <https://doi.org/10.3389/fpsyg.2011.00168>
- Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, 7(1). [https://doi.org/10.1016/S0163-6383\(84\)80022-3](https://doi.org/10.1016/S0163-6383(84)80022-3). Article 1.
- Martínez-Alvarez, A., Gervain, J., Koulaguina, E., Pons, F., & de Diego-Balaguer, R. (2023). Prosodic cues enhance infants' sensitivity to non-adjacent regularities. *Science Advances*, 9(15), eade4083. <https://doi.org/10.1126/sciadv.ade4083>