COGNITIVE NEUROSCIENCE

Prosodic cues enhance infants' sensitivity to nonadjacent regularities

Anna Martinez-Alvarez[1,2,3], Judit Gervain[1,2], Elena Koulaguina[4], Ferran Pons[3,5], Ruth de Diego-Balaguer[3,5,6,7]*

In language, grammatical dependencies often hold between items that are not immediately adjacent to each other. Acquiring these nonadjacent dependencies is crucial for learning grammar. However, there are potentially infinitely many dependencies in the language input. How does the infant brain solve this computational learning problem? Here, we demonstrate that while rudimentary sensitivity to nonadjacent regularities may be present relatively early, robust and reliable learning can only be achieved when convergent statistical and perceptual, specifically prosodic cues, are both present, helping the infant brain detect the building blocks that form a nonadjacent dependency. This study contributes to our understanding of the neural foundations of rule learning that pave the way for language acquisition.

INTRODUCTION

One of the most puzzling problems that infants face when acquiring language is determining which linguistic units are linked to each other. When the units are adjacent to one another, computing their dependence is straightforward and allows listeners to learn regularities between them (*1*, *2*). However, how does the infant brain face the challenge of computing regularities over linguistic units that are distant from one another in a sentence? The subject, for instance, is often separated from the verb by a varying number of words yet can agree with it in number and person irrespective of their distance (e.g., She sleeps. She sometimes sleeps. She, who never drinks coffee, sleeps more). The possible candidates for a nonadjacent dependency are theoretically infinite (*3*, *4*). At the heart of this problem lies the challenge of identifying among the infinite number of possible candidates which units the dependency needs to be computed over.

Nonadjacent regularities between nonidentical items, generally referred to as AxB rules, are extremely difficult to learn. AxB dependencies refer to the surface relationship between two distinct items (A and B) separated by unrelated intermediate items (*x*) varying in number (*5*). Infants fail to detect a nonadjacent dependency in artificial grammars when there is a reduced variability in the pool of intermediate *x* elements (*6*). Even in natural language, they do not succeed when the intermediate elements are composed of material not previously exposed (*7*), a two-syllable adverb (*8*), or are longer than three syllables (e.g., Granma is almost always singing) (*9*). Successful learning is only observed under certain restricted circumstances. For example, infants discriminate between grammatical (e.g., is eating) versus ungrammatical sequences (*can eating) when there is either a high variability (i.e., $x > 18$) of intervening items or no intermediate variability at all (i.e., $x = 1$) (*6*), when

the A and B items are phonologically similar (*10*), or when infants had prior exposure to A and B (*11*).

Thus, numerous studies have attempted to understand how infants learn nonadjacent relations. However, the inconsistent patterns of success and failure in AxB learning have led to an enduring debate about the mechanisms underlying the extraction of nonadjacent rules from speech. Considerable evidence supports the role of statistical learning (*12*, *13*). However, the fact that learning can only occur in restricted situations has led to the proposal that infants need additional mechanisms to be able to learn nonadjacent regularities in language. The existing studies in which learning was successful appear to highlight the dependent items. Successful learning thus occurs in situations in which the dependent items are more similar to one another than to the intermediate ones due to their intrinsic perceptual similarity or because they are grouped together by the high variability of the intermediate items. Therefore, Gestalt principles of perception (*14*, *15*), as well as attention processes (*16*), may facilitate nonadjacent dependency learning.

Prosody, often referred to as "the music of speech," helps infants acquire language. Newborns already group speech sounds on the basis of the acoustic cues that carry prosodic prominence in their native language (*17*). Prosodic bootstrapping has also been shown to support word learning (*18*), and prosody cues the acquisition of the word order of the native language (*19*). In these cases, prosodic prominence, either through its position or through its acoustic realization, supports learning by highlighting information relevant for grammar or the lexicon. Successful detection of nonadjacent regularities is observed when pauses co-occur with the dependencies (*20*, *21*). Given the central role of early sensitivity to nonadjacent rules for subsequent language development (*22*), it is essential to understand whether prosody helps solve the learning problems of nonadjacent dependencies.

Here, we put forth the hypothesis that if Gestalt principles of perception or attentional processes play a role in enhancing learning, then prosody may be a highly relevant cue that supports learning nonadjacent dependencies very early on in language development. The highlighting function of the sound patterns of language could assist infants solve the computational learning problem of nonadjacent dependencies. Our study investigates this hypothesis by

[1]Department of Social and Developmental Psychology, University of Padua, Padua, Italy. [2]Integrative Neuroscience and Cognition Center, Université Paris Cité and CNRS, Paris, France. [3]Department of Cognition, Development, and Educational Psychology, University of Barcelona, Barcelona, Spain. [4]Cognitive Neuroimaging Unit, CNRS ERL 9003, INSERM U992, CEA, Université Paris-Saclay, NeuroSpin Center, Gif-sur-Yvette, Île-de-France, France. [5]Institute of Neurosciences, University of Barcelona, Barcelona, Spain. [6]Cognition and Brain Plasticity Unit, IDIBELL, L'Hospitalet de Llobregat, Barcelona, Spain. [7]ICREA, Barcelona, Spain.
*Corresponding author. Email: ruth.dediego@ub.edu

comparing infants' sensitivity to nonadjacent dependencies with or without concurrent pitch cues. We tested four groups of 9-month-old infants exposed to trisyllabic rule sequences conforming to an AxB structure, whereby the A and B tokens predicted one another with certainty (e.g., "pedibu" and "pegabu"). Learning in these artificial language paradigms has been linked to natural language learning (10, 23), to syntactic and morphosyntactic deficits (24, 25), and to later language development in infants (22), suggesting that these paradigms successfully capture learning abilities that underlie natural language acquisition. We chose the age of 9 months since it is the optimal age where, behaviorally, infants do not succeed in learning, but the available studies do not include convergent cues that can help infants to extract the pattern of regularities. Infants' discrimination abilities were measured behaviorally using looking time measures and neurally using functional near-infrared spectroscopy (fNIRS).

The comparison of behavioral and fNIRS results is a key methodological innovation of our study. Since neural measures do not require overt responses, they have the potential to reveal sensitivities in infants that behavioral tests fail to find at an early stage of development (26). At the same time, neural responses without corresponding behavioral counterparts are challenging to interpret. However, behavioral and neural responses are usually not reported within the same study, and differences in the procedures and materials used in separate studies often make comparisons difficult. Behavioral and neural measures provide complementary but equally relevant information about infants' perceptual and cognitive abilities. In the current study, the behavioral and NIRS versions of the experiments were matched as closely as possible, placing similar demands on the infants and providing the same amount of exposure to the same stimuli. Specifically, we compared the extraction of an AxB dependency in a flat, i.e., no pitch, and a pitch experiments behaviorally and using NIRS. We expected the behavioral data to uncover whether convergent prosodic cues to the nonadjacent dependency improve learning as early as 9 months of age. In addition, we used NIRS to reveal whether the brain is sensitive to the presence of dependencies even in the absence of overt discrimination in behavior and to identify the brain network responsible for improved sensitivity to nonadjacent dependencies in the pitch experiment.

Experiments 1 and 2 (Fig. 1A) measured infants' behavioral and neural responses, respectively, to AxB stimuli with flat pitch. Experiments 3 and 4 (Fig. 1A) measured, respectively, infants' behavioral and neural responses when pitch cues highlighted the nonadjacent regularity. Specifically, the first and last items, A and B, respectively, in the trisyllabic sequences were marked by elevated pitch compared to the intermediate items. In all experiments, two conditions were presented. The rule condition implemented the AxB regularity (e.g., pedibu and pegabu), while the no-rule condition included trisyllabic sequences made up of the same syllable set as in the rule condition but in a random order, i.e., not respecting the AxB nonadjacent dependency (e.g., tabupe and bupego). The pitch manipulation in experiments 3 and 4 was thus uninformative in the no-rule condition as syllables varied randomly in the A and B positions and formed no regularity. Therefore, both the rule and no-rule conditions contained the same prosodic information, i.e., higher pitch in the A and B positions but only in the rule condition, prosody aligned with the nonadjacent regularity, rendering the prosodic and statistical cues convergent. The same syllables were used in all conditions and experiments, and they were presented with the same frequency.

In the two behavioral experiments (experiments 1 and 3), infants were familiarized with trisyllabic rule, i.e., AxB sequences for 2.4 min and subsequently tested with trisyllabic sequences not previously presented but following or violating the rule (Fig. 1A). AxB rule detection abilities were assessed by comparing looking times to rule versus no-rule test sequences (Fig. 1B). In the two NIRS experiments (experiments 2 and 4), sequences were presented in 30-s blocks separated by silences of a jittered duration of 25 to 35 s (Fig. 1C). We assessed AxB dependency sensitivity by comparing the hemodynamic responses to the rule and no-rule conditions (Fig. 1D). NIRS recorded infants' neural activity (Fig. 1E) in bilateral temporal, frontal, and parietal areas (Fig. 1F).

RESULTS

### Infants' responses to nonadjacent dependencies in flat speech

When infants were presented with flat speech in the behavioral experiment (experiment 1), a paired sample t test on infants' averaged looking times for rule and no-rule test trials did not show any difference between them [$t_{(19)} = 0.17$, $P = 0.87$, Cohen's $d = 0.039$; rule: $M = 15.34$ s, SE = 1.32; no rule: $M = 15.14$ s, SE = 1.2]. Replicating previous behavioral studies, these results suggest that 9-month-olds seem unable to learn the nonadjacent dependency rule in the absence of acoustic cues.

In the NIRS experiment with flat speech (experiment 2), we performed a permutation test over oxyhemoglobin (oxyHb) change, based on two-tailed paired sample t tests comparing the no-rule and the rule conditions. In contrast to what we observed in the behavioral experiment, this analysis revealed a significantly higher activation for the no-rule than for the rule condition in a cluster in left frontotemporal regions (channels 1 and 4; $P < 0.001$) (Fig. 2). The permutation test over deoxyhemoglobin (deoxyHb) comparing the rule and no-rule conditions did not yield any significant activation, as is often the case in infant NIRS studies (26–28). Permutation tests comparing the no-rule condition to baseline and the rule condition to baseline are reported in the Supplementary Materials.

Infants' responses to nonadjacent dependencies in speech with pitch cues

Infants displayed greater sensitivity in the experiments where the speech stream contained pitch. In the behavioral experiment with pitch (experiment 3), a paired sample t test on infants' averaged looking times for the rule and no-rule test trials showed a significant preference for no-rule ($M = 17.28$ s, SE = 1.28) over the rule condition ($M = 15.65$ s, SE = 1.38), $t_{(19)} = -2.31$, $P = 0.033$, Cohen's $d = 0.52$. These results indicate that with concurrent pitch cues, preverbal infants are able to learn the nonadjacent dependency rule.

In the NIRS experiment with pitch (experiment 4), we performed a permutation test over oxyHb change based on two-tailed, paired sample t tests comparing the no-rule and rule conditions (Fig. 3). This analysis revealed a significantly greater activation for the rule than for the no-rule condition in a spatial cluster in left temporal regions (channels 3 and 6; $P < 0.001$) and in a spatial cluster in right temporal regions (channels 19 and 21; $P < 0.001$) (Fig. 3). The permutation test over deoxyHb comparing the no-rule and rule conditions did not yield significant differences, as is often the case in infant NIRS studies (26–28). Permutation tests
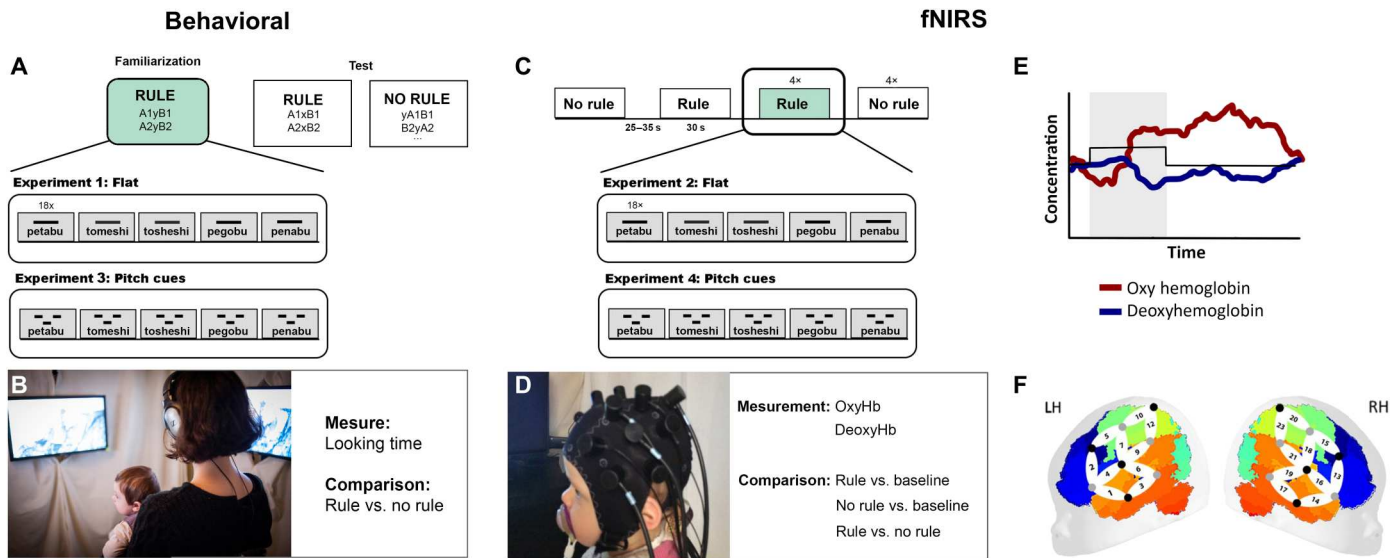
Fig. 1. Experimental design. (**A**) Experimental design of behavioral experiments (experiments 1 and 3). (**B**) Central fixation procedure with the measurement and comparison between test conditions. (**C**) Experimental design of fNIRS experiments (experiments 2 and 4). The stimuli either lacked (experiment 2) or contained (experiment 3) elevated pitch on the first and last syllables. In all four experiments, in the rule condition, all trisyllabic sequences conformed to an AxB structure, whereby A and B tokens predicted one another with certainty (e.g., pedibu and pegabu). In the no-rule condition, the same syllables were presented pseudo-randomly (e.g., "dibupe" and "bugape"). The stimuli either lacked (experiment 1 and 2) or contained (experiments 3 and 4) elevated pitch on the first and last (A and B) syllables. The familiarization phase in experiments 1 and 3 lasted 2.4 min and corresponded to the stimuli presented in the rule condition in experiments 2 and 4. (**D**) Picture of a participant wearing the NIRS cap with the measurement and comparison between conditions. (**E**) A typical hemodynamic response function in infants, with time on the x axis, Hemoglobin (Hb) concentration on the y axis, and time of stimulation marked in gray. (**F**) Probe configuration overlaid on an age-appropriate segmented anatomical brain image using localization analysis. Gray circles indicate sources, black circles detectors. White ovals indicate channel numbers (10 per hemisphere). Frontal lobe in blue, temporal lobe in orange, and parietal lobe in green. Localization analysis following (*27*) and (*17*) and using age-appropriate structural magnetic resonance imaging and stereotaxic atlases (*37*, *38*), indicates that channels 2 and 13 are located over the frontal lobe; channels 1, 3, 6, 9, 14, 17, 19, and 21 are placed over the temporal lobe; channels 10, 12, 20, and 23 query the parietal lobe; and channels 4, 5, 7, 15, 16, and 18 span two lobes.

comparing the no-rule condition to baseline and the rule condition to baseline are reported in the Supplementary Materials.

Comparison between behavioral experiments

To directly assess the effect of pitch cues behaviorally, we compared infants' orientation times in response to the structural regularities in the absence versus presence of pitch cues in a mixed analysis of variance (ANOVA) with the between-subjects factor pitch (flat experiment 1/pitch experiment 3) and the within-subjects factor rule (rule/no rule). We observed no main effect of rule, $F_{(1,38)} = 1.06$, $P = 0.31$, neither a main effect of experiment, $F_{(1,38)} = 0.52$, $P = 0.48$, nor the rule-by-experiment interaction, $F_{(1,38)} = 1.76$, $P = 0.19$ (Fig. 4).

Comparison between NIRS experiments

To directly assess the effect of pitch cues using NIRS, we compared oxyHb concentrations in response to the structural regularities in the absence versus presence of pitch cues in a mixed ANOVA with the between-subjects factor pitch (flat experiment 2/pitch experiment 4) and the within-subjects factors rule (rule/no rule) and hemisphere [left hemisphere (LH)/right hemisphere (RH)]. For this analysis, we extracted data from the significant clusters obtained in the permutation test for each experiment, therefore defining regions of interest (ROIs) in a data-driven, functional way. When a significant cluster was not present (flat experiment, RH), the analogous channels on the contralateral side were used (Fig. 5A). Specifically, for the flat experiment, the ROI in the LH included channels 1 and

4, and the ROI in the RH included channels 14 and 16 (time window, 6 to 30 s). For the pitch experiment, the ROI in the LH included channels 3 and 6, and the ROI in the RH included channels 19 and 21 (time window, 12 to 38 s). The ANOVA revealed a main effect of rule [$F_{(1,37)} = 4.734$, $P = 0.031$, $\eta^2 = 0.026$] and a significant pitch-by-rule interaction [$F_{(1,37)} = 18.82$, $P < 0.001$, $\eta^2 = 0.105$] (Fig. 5B). Scheffé's post hoc tests indicated that when pitch cues were present (experiment 4), the oxyHb concentration was significantly higher in the rule condition than in the no-rule condition ($P < 0.001$) and that, in the rule condition, the oxyHb concentration was significantly higher in experiment 4 (pitch) than in experiment 2 (flat) ($P < 0.001$) (Fig. 5C). No other main effect or interaction was significant.

Together, our results show that when asked to show an overt behavior, 9-month-olds only discriminate nonadjacent regularities from random controls if pitch cues highlight the nonadjacent items. By contrast, their brain responses show earlier sensitivities, discriminating between the nonadjacent regularity and the random controls at this age even without a convergent pitch cue. However, the magnitude and the localization of the responses differ considerably as a function of the absence or presence of the concurrent pitch cues highlighting the regularity (see Fig. 5).

DISCUSSION

Powerful learning mechanisms are available to young infants, which allow them to learn language efficiently in the first years of life. How
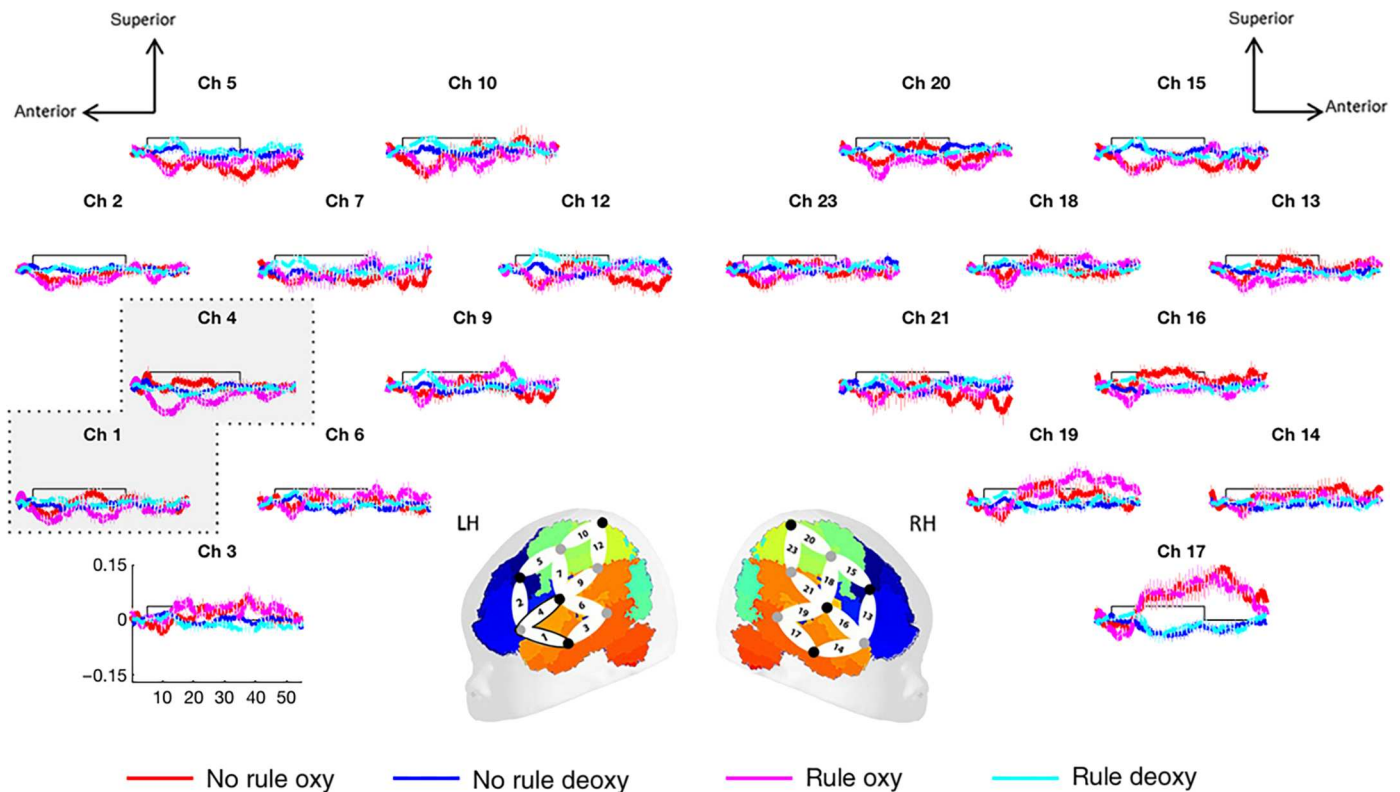
Fig. 2. Hemodynamic responses. experiment 2 (flat) Grand average hemodynamic responses evoked by the rule and no-rule conditions in each channel in experiment 2 (flat). Channels are plotted following the probe placement indicated in the inset (frontal lobe in blue, temporal lobe in orange, and parietal lobe in green). The x axis represents time in seconds. The y axis shows concentration change in millimoles × millimeters. The rectangle along the x axis indicates time of stimulation (in seconds). The curves indicate grand average responses for rule blocks and no-rule blocks. Error bars indicate SEMs. Gray-shaded rectangles depict significant spatial clusters obtained in the cluster-based nonparametric permutation tests comparing the rule and no-rule conditions. N = 20.

they extract the complex rules of grammar, many of which involve dependencies between items that are not adjacent to each other, remains unknown. The present study reports two behavioral and two NIRS experiments that uncovered two of the potential neural mechanisms young infants are equipped with to solve this task. We found that while rudimentary sensitivity to nonadjacent regularities without any supporting perceptual cues may be present already at 9 months, robust and reliable learning can only be achieved at this age when convergent statistical and prosodic cues are both present, helping the infant brain detect the building blocks that form a nonadjacent dependency. These findings shed light on the understanding of the role of prosody in language acquisition. Specifically, we have directly examined how prosody affects the way nonadjacent dependencies are tracked. The current findings provide evidence for the crucial impact of subtle pitch changes in the processing of statistical information in early infancy.

Our results converge with previous studies regarding the role of the auditory signal in learning nonadjacent regularities. To explain successful detection of violations of nonadjacent dependencies at 3 to 4 months observed electrophysiologically, we hypothesized that infants may rely on phoneme-level similarities between the dependent elements, e.g., shared vowels between the auxiliary and the suffix (10). Furthermore, in young infants, a correlation between the ability to detect pitch violations and nonadjacent dependency violations has also been observed (5). Our study, by systematically

manipulating prosody, reveals a direct link between prosody and nonadjacent rule learning.

How does prosody help learning? Earlier studies observed nonadjacent rule learning only after the first year of life (29). This developmental change has been linked to the maturation of infants' attentional system (14). Our findings mesh well with these accounts. Pitch may provide a salient cue to highlight the relevant elements and their dependency, enhancing infants' performance. Specifically, while executive components of attention develop through childhood, exogenous attention mechanisms, already in place at birth, may help younger infants orient to the prosodic cues in the speech stream (30). These findings thus suggest that the powerful attention system infants are equipped with early in life may assist language learning in preverbal infants.

More specifically, we propose that exogenous attention helps trigger learning when the speech signal has certain characteristics that highlight the dependent elements, rendering them different from the intermediate ones. This mechanism could be engaged either by the perceptual features of the dependent elements (i.e., externally driven) or engaged by previous knowledge of the dependent elements (i.e., internally driven). For instance, 7-month-olds, but not yet newborns (31), may be able to learn nonadjacent repetition-based rules, i.e., ABA nonadjacent regularities (19), because the reduplication of the dependent elements renders them phonologically different from the intermediate one, making it a salient cue
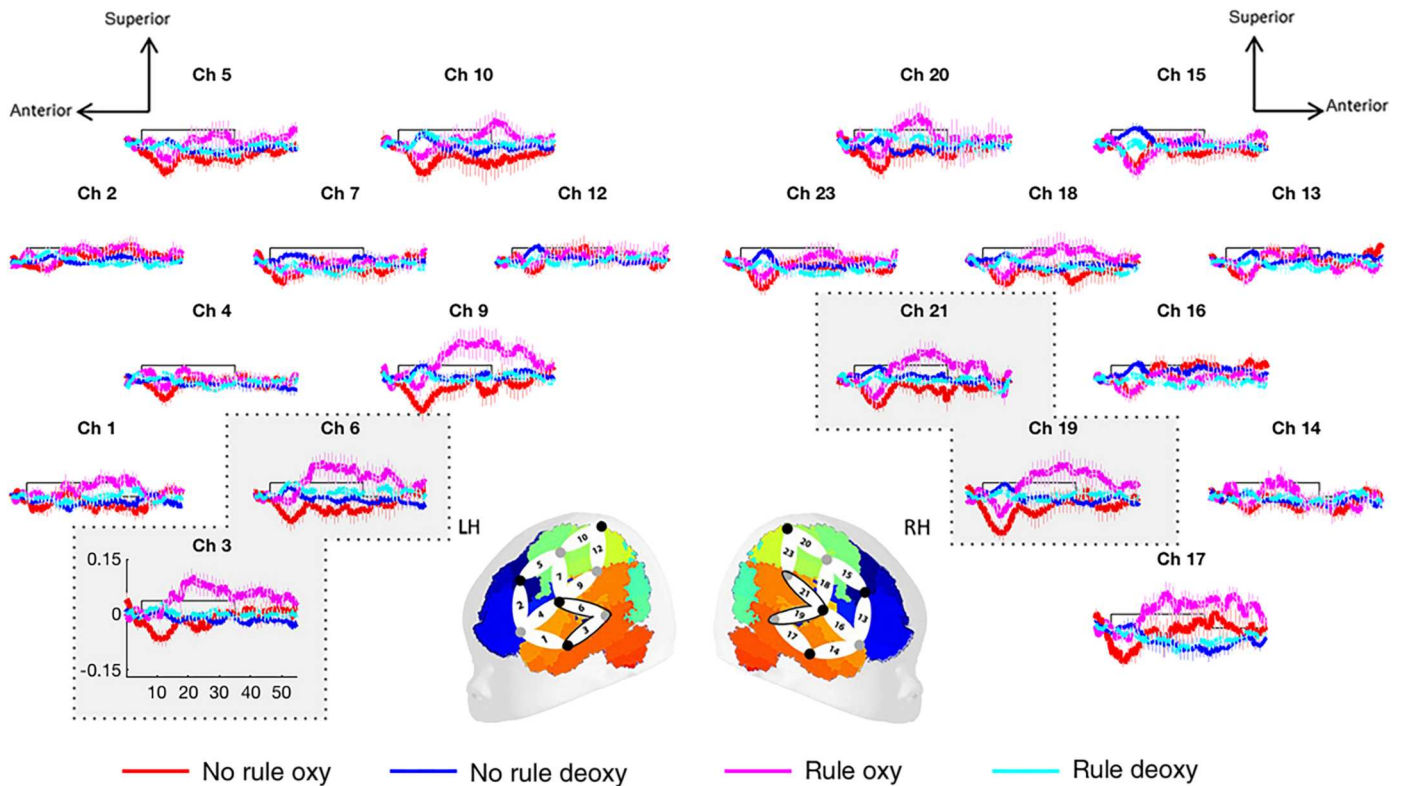
Fig. 3. Hemodynamic responses. experiment 4 (pitch) Grand average hemodynamic responses evoked by the rule and no-rule conditions in each channel in experiment 4 (pitch). Channels are plotted following the probe placement indicated in the inset (frontal lobe in blue, temporal lobe in orange, and parietal lobe in green). The *x* axis represents time in seconds. The *y* axis shows concentration change in millimoles × millimeters. The rectangle along the *x* axis indicates time of stimulation (in seconds). The curves indicate grand average responses for rule blocks and no-rule blocks. Error bars indicate SEMs. Gray-shaded rectangles depict significant spatial clusters obtained in the cluster-based nonparametric permutation tests comparing the rule and no-rule conditions. $N = 20$.
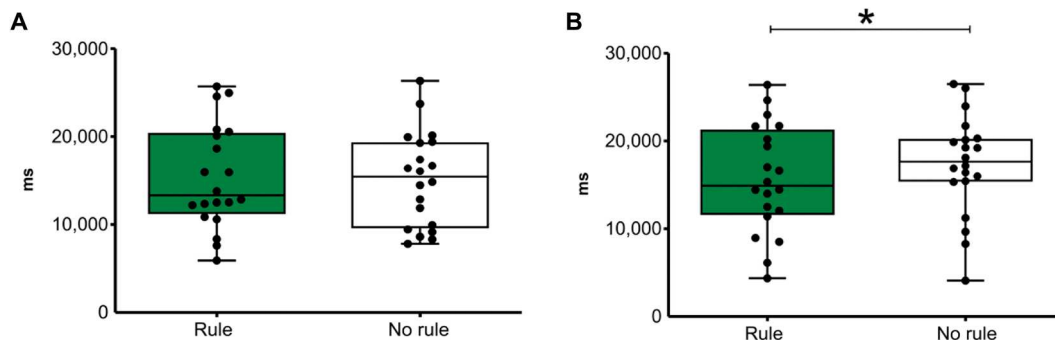

Fig. 4. Looking times in experiment 1 (flat) and experiment 3 (pitch). (**A**) Grand average looking times (in milliseconds) for rule and no-rule conditions in the absence of pitch cues ($N = 20$, two-tailed paired sample *t* test; $P < 0.8$) and (**B**) in the presence of pitch cues ($N = 20$, two-tailed paired sample *t* test; $*P < 0.03$). Dots represent each infant's averaged looking time. ANOVA with between-subjects pitch and within-subjects rule: rule-by-pitch interaction ($P < 0.2$).

to attract exogenous attention. In the case of nonidentical nonadjacent dependencies as the AxB rule tested here, such an attention-based mechanism may be engaged when the intermediate xitem is highly variable (*6*), leading to successful rule discrimination. Attentional mechanisms can also account for the opposite learning scenario, whereby the intermediate element is totally invariable, while the nonadjacent ones vary (*32*). Since languages differ in the way nonadjacent regularities are marked in the speech signal, our

attention system might allow us to orient dynamically to the speech input for successful learning.

The only previous study seeking to identify the brain network underlying nonadjacent rule learning in childhood tested 3- to 4-year-olds' ability to discriminate grammatical versus ungrammatical AxB sequences and found a bilateral network including frontal, temporal, and parietal regions involved in the rule violation detection (*33*). In the current study, 9-month-olds' activation of temporal
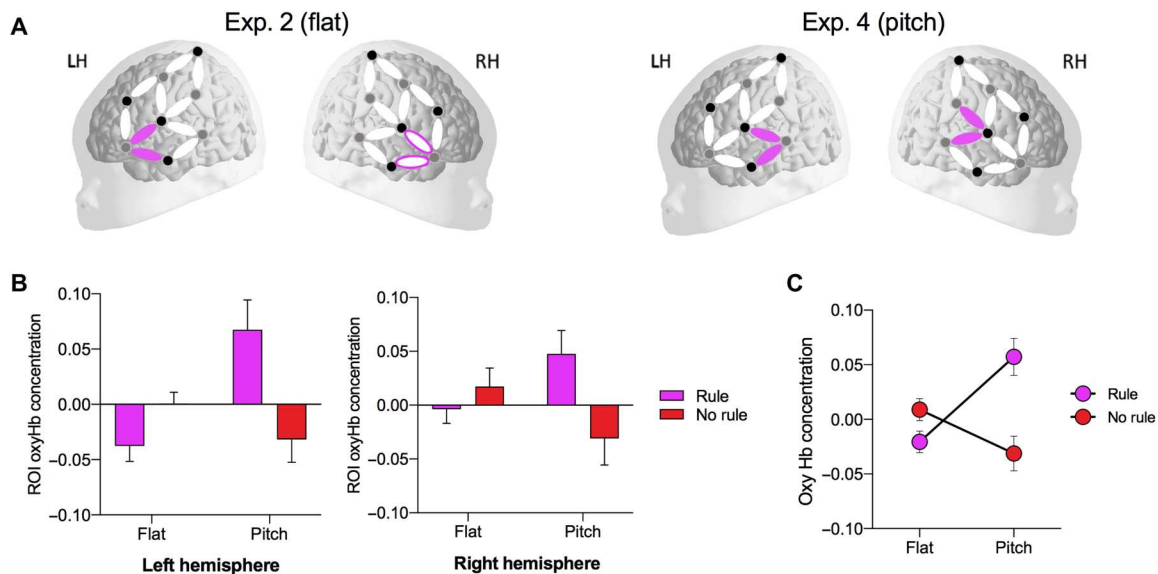
Fig. 5. Pitch-by-rule interaction. (**A**) Significant clusters obtained in the permutation test for each experiment. When no significant cluster appeared (flat experiment, RH), the analogous channels on the contralateral side were used (encircled in pink). These constituted our ROI. (**B**) ROI oxyHb concentrations in response to the rule (pink) and no-rule (red) conditions in the pitch ($N = 20$) versus flat ($N = 20$) experiments in the LH and RH. ANOVA with pitch, rule, and hemisphere. Rule effect: $P < 0.03$. (**C**) Significant pitch-by-rule interaction ($P < 0.001$).

regions found in experiment 2 for discriminating rule versus no-rule sequences partially overlaps with this bilateral network (*33*), suggesting developmental continuity in the neural circuitry involved in this task across infancy and early childhood, when most of the native language grammar is acquired.

Whether the same underlying mechanisms support both adjacent and nonadjacent dependency learning is unknown. Our findings suggest that the two may overlap. Both adjacent and nonadjacent rule extraction seem to recruit a bilateral temporal network in infancy (*31*). Our results suggest, however, that this temporal network is only engaged when concurrent prosodic cues highlight the dependent elements.

These are the first results comparing behavioral and fNIRS in an infant grammar study. This methodological innovation has yielded interesting findings. While neural measures revealed a pitch-by-rule interaction, the same interaction did not appear when infants were required to show an overt response. These results suggest that neuroimaging techniques such as fNIRS have the potential to reveal sensitivities in infants that behavioral tests fail to find at an early stage of development (*26*). By using both behavioral and neural data, we have thus been able to, first, uncover infants' overt responses showing that convergent prosodic cues to the nonadjacent dependency improve learning and, second, uncover the brain network responsible for improved sensitivity to nonadjacent dependencies in the pitch condition.

Prosody converging with grammar (*34*) serves to chunk the speech signal into morphosyntactically meaningful units and marks certain syntactic functions, thereby facilitating speech processing during language use. Our results imply that for the young learner, prosody may also serve the function of highlighting relevant information, thus contributing to the solution of the learning problem. This idea converges well with prosodic bootstrapping approaches to language acquisition (*35*), suggesting that the speech signal is rich in cues that help infants break into the grammar of their native language.

Here, we show that when preverbal infants face the challenge of extracting statistical regularities over nonadjacent elements from ongoing speech, their brain engages different neural mechanisms depending on the prosodic cues available in the speech stream. We found that infants' differential brain responses to nonadjacent regularities were relatively weak in the absence of pitch cues highlighting the dependent elements, and such discrimination seems absent when assessing infants' overt behavior. However, when concurrent pitch cues are provided, infants robustly discriminate nonadjacent dependencies from random controls in the bilateral temporal regions. This pattern of results suggests that the infant brain exploits signal-driven cues to compute the regularities over nonadjacent items in the language input. Uncovering this learning ability in early infancy has important implications developmentally, as it has been associated to infants' later language development (*22*).

MATERIALS AND METHODS
Experimental design
Our study investigates the role of prosody in infant rule learning. To do so, we compare 9-month-old infants' sensitivity to nonadjacent dependencies with or without concurrent pitch cues. We tested four groups exposed to trisyllabic rule sequences conforming to an AxB structure, whereby the A and B tokens predicted one another with certainty (e.g., pedibu and pegabu). Infants' discrimination abilities were measured behaviorally using looking time measures and neurally using NIRS. The ethics committee of the Universitat de Barcelona (Institutional Review Board 00003099) approved the study.

Participants

We calculated the sample size for the behavioral studies considering a mixed ANOVA with the between-subjects factor pitch (flat experiment 2/pitch experiment 4) and the within-subjects factor rule (rule/no rule). On the basis of previous work with the same measure and similar paradigm in infants, we expected a small-to-moderate effect size of 0.44 (29) and assumed a power of at least 0.8 with a low correlation of 0.2 between repeated measures. These calculations require a total sample of 22 across two groups. We included 40 infants in the final sample, achieving a 0.99 power.

We determined the sample size for the NIRS studies considering the mixed ANOVA with the between-subjects factor pitch (flat experiment 2/pitch experiment 4) and the within-subjects factors rule (rule/no rule) and hemisphere (LH/RH). Assuming a target power of at least 0.8, a low correlation of 0.2 between repeated measures and a small-to-moderate effect size of 0.30 derived from the previous NIRS study testing rule-learning in infants (31), the necessary sample size was calculated to be a total of 22 infants across the two groups. We thus included a total of 40 infants in the final sample, achieving a 0.98 power. The software G*Power was used for all calculations.

We ran the behavioral (experiment 1 and 3) experiments in parallel, and we assigned infants to conditions randomly. We ran the NIRS experiment flat and the NIRS experiment with pitch at different periods. We thus assigned infants to each experiment upon recruitment. All parents signed informed consent for the participation of the infant before the enrollment in the experimental procedure.

Experiment 1 (flat)

Twenty healthy full-term French-hearing monolingual 9-month-old infants (M age = 9.8; SD = 0.32; 12 boys and 8 girls) were included in the analysis. Three additional infants were tested but excluded because of fussiness.

Experiment 2 (flat)

Twenty healthy full-term French-hearing monolingual 9-month-old infants (M age = 8.21; SD = 0.15; 6 boys and 14 girls) were included in the analysis. Thirty-two additional infants were tested but excluded from the final analysis due to fussiness or crying ($n = 12$), parental interference ($n = 1$), insufficient data (i.e., less than two trials per condition), or bad data quality ($n = 19$).

Experiment 3 (pitch)

Twenty healthy full-term French-hearing monolingual 9-month-old infants (M age = 9.7; SD = 0.32; 9 boys and 11 girls) were included in the analysis. Two additional infants were tested but excluded because of fussiness ($n = 1$) and to "sticky fixation," i.e., reaching maximal duration in all test trials ($n = 1$).

Experiment 4 (pitch)

Twenty healthy full-term French-hearing monolingual 9-month-old infants (M age = 9.03; SD = 0.21; 9 boys and 11 girls) were included in the analysis. Forty-five additional infants were tested but excluded from the final analysis due to fussiness or crying ($n = 23$), parental interference ($n = 1$), insufficient data (i.e., less than two trials per condition), or bad data quality ($n = 21$).

Stimuli

All auditory stimuli in familiarization and test were synthesized with MBROLA (36) using the fr4 French-speaking female voice of the MBROLA diphone database (31). The duration of consonants was set to 120 ms and vowels to 150 ms (31). In experiment 1 and 2, pitch was uniformly set to 200 Hz. In experiment 3 and 4, pitch was not monotonous; on the intervening syllables, pitch was set to 200 Hz, while for the first and last syllables, pitch was raising (from 200 to 220 Hz) (5).

Behavioral experiments. experiment 1

Familiarization consisted of 72 trisyllabic items following an AxB rule. Each item contained one of two nonadjacent dependencies in the A and B positions (pe x bu; to x shi): e.g., petabu, tomeshi. To enhance rule learnability, the intermediate x element was randomly chosen of a pool of 18 items (ba, bo, di, du, fe, go, ke, lo, la, mu, me, na, pi, ta, she, shu, zo, and zi), given that high variability of intermediate elements aids learning (6). There were no pauses between the three syllables within an AxB sequence (31). The interstimulus interval (ISI) between sequences was randomly chosen to be either 0.5 or 1.5 s (31). In addition, a colorful animation was used as a visual attention-getter to keep infants focused on the task. At test, the auditory stimuli were eight not previously presented sequences. Four of them had an AxB structure (rule trials) with two sequences with each of the nonadjacent dependencies from familiarization (pe x bu; to x shi) and four x items (da, mi, fo, and ze) not previously presented with the particular structure to test infants′ ability to generalize the underlying structure rather than simply recall items learned through rote memorization during familiarization. The other four test items contained violations of the AXB rule (no-rule trials) using the same syllables as in rule trials but with their order violating the AxB rule in all the five possible combinations: ABx, xAB, xBA, BxA, and BAx. The ISI between test trials was 1000 ms. A fragment of the same animation video as in familiarization was played as a visual attention-getter (identical in all test trials across all infants). For experiment 3, the auditory stimuli and the experimental design were the same as in experiment 1, except that the intervening syllables pitch was set to 200 Hz, while for the first and last syllables, the pitch was raising (from 200 to 220 Hz) (5), as in experiment 4.

NIRS experiments. experiment 2

The auditory stimuli in the rule condition were the same 72 trisyllabic sequences as in the familiarization phase of experiment 1. The auditory stimuli in the no-rule condition contained the same syllables with the same frequency as in the rule condition but without following an AxB rule. The possible combinations were ABx, xAB, xBA, BxA, and BAx (for details on the combinatorics, see table S1). For experiment 4, the auditory stimuli and the experimental design were the same as in experiment 2, except that the intervening syllables pitch was set to 200 Hz, while for the first and last syllables, the pitch was raising (from 200 to 220 Hz) (5), as in experiment 3.

Procedure

Behavioral experiments (experiment 1 and 3)

Each infant was tested individually in a quiet, dimmed room. Infants were seated on a caregiver's lap in the center of the experimental booth. Each infant was tested individually with a central fixation procedure. The stimuli were presented using Habit 2.0 (34) software, which automatically recorded the orientation times to a screen placed at approximately 80 cm in front of the infant. The auditory stimuli were played from speakers placed on the left and right sides of the screen covered by black curtains. The caregiver was instructed to not interact with or talk to the child. During the experiment, caregivers listened to masking music over tight-fitting headphones to prevent them from hearing the stimuli. An observer

operated the experiment from outside the test booth, controlling stimulus presentation and observing infants' behavior through an online video stream. The observer pressed/released predefined buttons to signal infants' gaze fixation on the screen and their look-away. Infants' looking times were coded online by an experimenter blind to trial type. The experiment consisted of familiarization and test. Each trial was initiated when the infant looked at the screen. Test trials were fully infant controlled: They were terminated if the infant looked away from the screen for more than 2 s. If the look-away was less than 2 s, the trial was not terminated; the look-away time was included in the duration of infant's orientation time for that trial. If infants did not look away for more than 2 s, then the trial lasted for its maximal duration. Familiarization had a fixed duration for all infants irrespective of infants' looking behavior. Familiarization consisted of an animation video presented simultaneously with the auditory stimuli. Its duration was fixed to 143 s. The test phase contained eight test trials: four rule trials with one of the two familiarized nonadjacent dependencies and four no-rule trials containing violations of the dependencies. In each test trial, the trisyllabic sequence could be repeated for up to 15 times if the infant did not look away from the screen. The order of rule and no-rule trials was counterbalanced across participants. The maximal duration of each test trial was 26.5 s.

NIRS experiments (experiments 2 and 4)
Each infant was tested individually in a quiet, dimmed room. Infants were seated on a caregiver's lap in the center of the experimental booth. The stimuli were presented using E-Prime software on a screen placed at approximately 80 cm in front of the infant. The auditory stimuli were played from speakers placed on the left and right sides of the screen. The caregiver was instructed to not interact with or talk to the child. During the experiment, caregivers listened to masking music over tight-fitting headphones to prevent them from hearing the stimuli.

A NIRx NIRScout NIRS machine with eight sources and eight detectors (source-detector separation, 3 cm; two wavelengths of 760 and 850 nm; sampling rate, 15.625 Hz) combined into 20 channels in a stretchy cap (Fig. 1D) measured infants' brain responses to the stimuli in the frontal, temporal, and parietal areas. Localization analysis, following (27) and (17) and using age-appropriate structural magnetic resonance imaging and stereotaxic atlases (37, 38), indicates that channels 2 and 13 are located over the frontal lobe; channels 1, 3, 6, 9, 14, 17, 19, and 21 are placed over the temporal lobe; channels 10, 12, 20, and 23 query the parietal lobe; and channels 4, 5, 7, 15, 16, and 18 span two lobes (Fig. 1E).

Stimuli were presented in blocks (Fig. 1C). One block consisted of 18 trisyllabic sequences, separated by silences jittered between 0.5 and 1.5 s (31). The between-block interval was jittered between 25 and 35 s (31). The experiment consisted of eight blocks (four rule blocks and four no-rule blocks). The experiment lasted 8.34 min. Blocks were presented in an interleaved fashion with no more than two consecutive blocks of the same condition (31). Blocks were pseudorandomized and counterbalanced across participants. To prevent infants from moving and to maintain their attention, participants were presented with a silent animation video during the experiment.

NIRS data preprocessing
NIRS analysis was performed on all infants in batch using MATLAB scripts requiring no user input once basic analysis parameters were defined. The authors were thus blinded to the experimental conditions during data analysis. Changes of oxyHb and deoxyHb were calculated by means of the modified Beer-Lambert law. A band-pass filter between 0.01 and 0.7 Hz eliminated noise and systemic blood flow variations (e.g., infants' heartbeat) and overall trends. Block-channel pairs were rejected when movement artifacts (defined by intensity changes above 0.1 mmol × mm over two samples) occurred. For the remaining nonrejected blocks, a baseline was linearly fitted between the means of the 5 s preceding the onset of the block and the 5 s after the 15 s following the offset of the block. Only participants with at least 50% artifact-free blocks were entered into the final analysis.

Statistical analyses
Behavioral experiments
First, paired sample t tests (two-tailed) were conducted to compare infants' looking times in the rule and no-rule conditions for experiments 1 and 3, respectively. Second, to directly compare infants' orientation times to structural regularities in the absence versus presence of pitch cues, we ran a mixed ANOVA with between-subjects factor experiment (flat/pitch) and within-subjects factor rule (rule/no rule).

NIRS experiments
To establish ROIs and time windows of interest in a data-driven way, we first performed three cluster-based nonparametric permutation tests (39), which established spatially adjacent channels in which significant differences, as determined by two-tailed, paired sample t tests between the conditions being compared, were observed in temporally adjacent samples, separately for oxyHb and deoxyHb. We ran 1000 permutations under the null hypothesis. Cluster-based permutation tests (39), a data-driven method, is now the standard in the NIRS literature to identify significant ROIs as it suitably safeguards against the problem of multiple comparisons and thus offers a stringent statistical test. Second, to directly compare infants' brain responses to structural regularities in the absence versus presence of pitch cues, we ran a mixed ANOVA with between-subjects factor experiment (flat/pitch) and within-subjects factors rule (rule/no rule) and hemisphere (LH/RH) over oxyHb concentrations in the ROI and time window of interest identified by the permutation analyses. In the experiment with flat prosody, a significant cluster emerged only in the LH, we therefore took the corresponding cluster and time window in the RH as the contralateral ROI. We used normalized concentration changes in the ANOVA, as is appropriate for between-subject comparison with NIRS data (17, 40, 41). Given the short attention span of infants and the limited testing time that results from it, it would not have been possible to test all four conditions in a single experiment to achieve a full within-subject design.

Supplementary Materials
This PDF file includes:
Supplementary Text
Table S1
Legend for data S1

Other Supplementary Material for this manuscript includes the following:
Data S1

View/request a protocol for this paper from *Bio-protocol*.

REFERENCES AND NOTES

1. J. R. Saffran, R. N. Aslin, E. L. Newport, Statistical learning by 8-month-old infants. *Science* **274**, 1926–1928 (1996).
2. G. F. Marcus, S. Vijayan, S. Bandi Rao, P. M. Vishton, Rule learning by seven-month-old infants. *Science* **283**, 77–80 (1999).
3. R. B. Lees, N. Chomsky, Syntactic structures. *Language* **33**, 375 (1957).
4. M. D. Hauser, N. Chomsky, W. T. Fitch, The faculty of language: What is it, who has it, and how did it evolve? *Science* **298**, 1569–1579 (2002).
5. J. L. Mueller, A. D. Friederici, C. Männel, Auditory perception at the root of language learning. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 15953–15958 (2012).
6. R. L. Gómez, Variability and detection of invariant structure. *Psychol. Sci.* **13**, 431–436 (2002).
7. M. van Heugten, E. K. Johnson, Linking infants' distributional learning abilities to natural language acquisition. *J. Mem. Lang.* **63**, 197–209 (2010).
8. B. Hohle, M. Schmitz, L. M. Santelmann, J. Weissenborn, The recognition of discontinuous verbal dependencies by german 19-month-olds: Evidence for lexical and structural influences on children's early processing capacities. *Lang. Learn. Dev.* **2**, 277–300 (2009).
9. L. M. Santelmann, P. W. Jusczyk, Sensitivity to discontinuous dependencies in language learners: Evidence for limitations in processing space. *Cognition* **69**, 105–134 (1998).
10. A. D. Friederici, J. L. Mueller, R. Oberecker, Precursors to natural grammar learning: Preliminary evidence from 4-month-old infants. *PLOS ONE* **6**, e17920 (2011).
11. J. Lany, R. L. Gómez, Twelve-month-old infants benefit from prior experience in statistical learning. *Psychol. Sci.* **19**, 1247–1252 (2008).
12. E. L. Newport, R. N. Aslin, Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cogn. Psychol.* **48**, 127–162 (2004).
13. I. C. Grama, A. Kerkhoff, F. Wijnen, Gleaning structure from sound: The role of prosodic contrast in learning non-adjacent dependencies. *J. Psycholinguist. Res.* **45**, 1427–1449 (2016).
14. R. de Diego-Balaguer, A. Martinez-Alvarez, F. Pons, Temporal attention as a scaffold for language development. *Front. Psychol.* **7**, 44 (2016).
15. S. Pacton, P. Perruchet, An attention-based associative account of adjacent and nonadjacent dependency learning. *J. Exp. Psychol. Learn. Mem. Cogn.* **34**, 80–96 (2008).
16. P. Perruchet, M. D. Tyler, N. Galland, R. Peereman, Learning nonadjacent dependencies: No need for algebraic-like computations. *J. Exp. Psychol. Gen.* **133**, 573–583 (2004).
17. N. Abboub, T. Nazzi, J. Gervain, Prosodic grouping at birth. *Brain Lang.* **162**, 46–59 (2016).
18. S. L. Mattys, P. W. Jusczyk, P. A. Luce, J. L. Morgan, Phonotactic and prosodic effects on word segmentation in infants. *Cogn. Psychol.* **38**, 465–494 (1999).
19. J. Gervain, J. F. Werker, Prosody cues word order in 7-month-old bilingual infants. *Nat. Commun.* **4**, 1490 (2013).
20. E. Marchetto, L. L. Bonatti, Finding words and word structure in artificial speech: The development of infants' sensitivity to morphosyntactic regularities. *J. Child Lang.* **42**, 873–902 (2015).
21. C. Kabdebon, M. Pena, M. Buiatti, G. Dehaene-Lambertz, Electrophysiological evidence of statistical learning of long-distance dependencies in 8-month-old preterm and full-term infants. *Brain Lang.* **148**, 25–36 (2015).
22. R. L. A. Frost, A. Jessop, S. Durrant, M. S. Peter, A. Bidgood, J. M. Pine, C. F. Rowland, P. Monaghan, Nonadjacent dependency learning in infancy, and its link to language development. *Cogn. Psychol.* **120**, 101291 (2020).
23. B. Pelucchi, J. F. Hay, J. R. Saffran, Learning in reverse: Eight-month-old infants track backward transitional probabilities. *Cognition* **113**, 244–247 (2009).
24. H. J. Hsu, J. B. Tomblin, M. H. Christiansen, Impaired statistical learning of non-adjacent dependencies in adolescents with specific language impairment. *Front. Psychol.* **5**, 175 (2014).
25. I. Lammertink, P. Boersma, F. Wijnen, J. Rispens, Statistical learning in specific language impairment: A meta-analysis. *J. Speech Lang. Hear. Res.* **60**, 3474–3486 (2017).
26. J. Gervain, J. Mehler, J. F. Werker, C. A. Nelson, G. Csibra, S. Lloyd-Fox, M. Shukla, R. N. Aslin, Near-infrared spectroscopy: A report from the McDonnell infant methodology consortium. *Dev. Cogn. Neurosci.* **1**, 22–46 (2011).
27. S. Lloyd-Fox, J. E. Richards, A. Blasi, D. G. M. Murphy, C. E. Elwell, M. H. Johnson, Coregistering functional near-infrared spectroscopy with underlying cortical areas in infants. *Neurophotonics* **1**, 025006 (2014).
28. M. A. Yücel, A. V. Lühmann, F. Scholkmann, J. Gervain, I. Dan, H. Ayaz, D. Boas, R. J. Cooper, J. Culver, C. E. Elwell, A. Eggebrecht, M. A. Franceschini, C. Grova, F. Homae, F. Lesage, H. Obrig, I. Tachtsidis, S. Tak, Y. Tong, A. Torricelli, H. Wabnitz, M. Wolf, Best practices for fNIRS publications. *Neurophotonics* **8**, 012101 (2021).
29. R. Gómez, J. Maye, The developmental trajectory of nonadjacent dependency learning. *Inf. Dent.* **7**, 183–206 (2005).
30. A. Martinez-Alvarez, M. Sanz-Torrent, F. Pons, R. de Diego-Balaguer, Rethinking attention in time: Expectancy violations reconcile contradictory developmental evidence. *J. Exp. Child Psychol.* **206**, 105070 (2021).
31. J. Gervain, F. Macagno, S. Cogoi, M. Peña, J. Mehler, The neonate brain detects speech structure. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 14222–14227 (2008).
32. L. Onnis, P. Monaghan, M. H. Christiansen, Variability is the spice of learning, and a crucial ingredient for detecting and generalizing in nonadjacent dependencies. *Cognit. Sci.* , (2002).
33. A. van der Kant, C. Männel, M. Paul, A. D. Friederici, B. Höhle, I. Wartenburger, Linguistic and non-linguistic non-adjacent dependency learning in early development. *Dev. Cogn. Neurosci.* **45**, 100819 (2020).
34. L. B. Cohen, D. J. Atkinson, H. H. Chaput, Habit X: A new program for obtaining and organizing data in infant perception and cognition studies (version 1.0) [Computer Software] (Austin, University of Texas, 2004).
35. J. L. Morgan, K. Demuth, Eds. *Signal to Syntax: Bootstrapping From Speech to Grammar in Early Acquisition* (Psychology Press, 2014).
36. T. Dutoit, H. Leich, MBR-PSOLA: Text-to-speech synthesis based on an MBE re-synthesis of the segments database. *Speech Commun.* **13**, 435–440 (1993).
37. C. Kabdebon, F. Leroy, H. Simmonet, M. Perrot, J. Dubois, G. Dehaene-Lambertz, Anatomical correlates of the international 10–20 sensor placement system in infants. *Neuroimage* **99**, 342–356 (2014).
38. P. T. Fillmore, J. E. Richards, M. C. Phillips-Meek, A. Cryer, M. Stevens, F. Richards, Stereotaxic magnetic resonance imaging brain atlases for infants from 3 to 12 months. *Dev. Neurosci.* **37**, 515–532 (2015).
39. E. Maris, R. Oostenveld, Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* **164**, 177–190 (2007).
40. G. Matsuda, K. Hiraki, Sustained decrease in oxygenated hemoglobin during video games in the dorsal prefrontal cortex: A NIRS study of children. *Neuroimage* **29**, 706–711 (2006).
41. Y. Otsuka, E. Nakato, S. Kanazawa, M. K. Yamaguchi, S. Watanabe, R. Kakigi, Neural activation to upright and inverted faces in infants measured by near infrared spectroscopy. *Neuroimage* **34**, 399–406 (2007).