

## COGNITIVE NEUROSCIENCE

# Infant cognition includes the potentially human-unique ability to encode embedding

M. Winkler<sup>1,2</sup>, J. L. Mueller<sup>1,3</sup>, A. D. Friederici<sup>1</sup>, C. Männel<sup>1,4,5\*</sup>

Human cognition relies on the ability to encode complex regularities in the input. Regularities above a certain complexity level can involve the feature of embedding, defined by nested relations between sequential elements. While comparative studies suggest the cognitive processing of embedding to be human specific, evidence of its ontogenesis is lacking. To assess infants' ability to process embedding, we implemented nested relations in tone sequences, minimizing perceptual and memory requirements. We measured 5-month-olds' brain responses in two auditory oddball paradigms, presenting standard sequences with one or two levels of embedding, interspersed with infrequent deviant sequences violating the established embedding rules. Brain potentials indicate that infants detect embedding violations and thus appear to track nested relations. This shows that the ability to encode embedding may be part of the basic human cognitive makeup, which might serve as scaffolding for the acquisition of complex regularities in language or music.

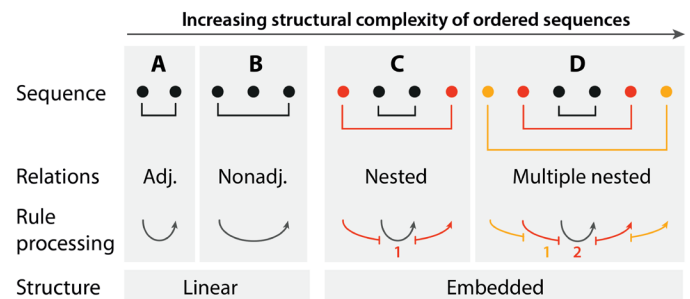
## INTRODUCTION

For humans, the ability to process complex regularities is a prerequisite for higher cognitive functions, such as language (1, 2), music (3), or mental arithmetic (4). While many of our perceptual and cognitive abilities can also be found in nonhuman species, the level of structural complexity that humans are able to master appears to be unique (1, 5). One crucial difference between humans and other species appears to lie in our ability to solve embedding, involving nested relations between sequential elements [see (6, 7); Fig. 1]. Nested relations are an essential ingredient of the syntax of human language, as can be seen in this embedded sentence [*The boy [the girl chased] kicked the ball.*] (brackets indicate the inner embedded sentence and the outer main sentence) (Fig. 2A). Such sentences can only be understood once the language system is able to recognize the underlying nested relations. However, the ability to encode nested relations in nonlinguistic auditory input may function as a precursor of the capacity to solve them in language. In the current study, we aim to investigate whether preverbal infants' cognitive capacities already include the ability to compute nested relations in the auditory domain. To date, there is no evidence thereof. Shedding light on the developmental origins of complex regularity processing would add to our understanding of the ontogenesis and the phylogenesis of human language.

According to formal language theory (8), the phenomenon of embedding introduces a distinctive boundary between so-called regular and context-free grammars. Regular grammars are equivalent to finite-state automata, which can generate linear structures in which each element depends directly on the previous one (Fig. 1, A and B). Infants have been shown to process those structures in adjacent and nonadjacent element relations from an early age (9–13). Context-

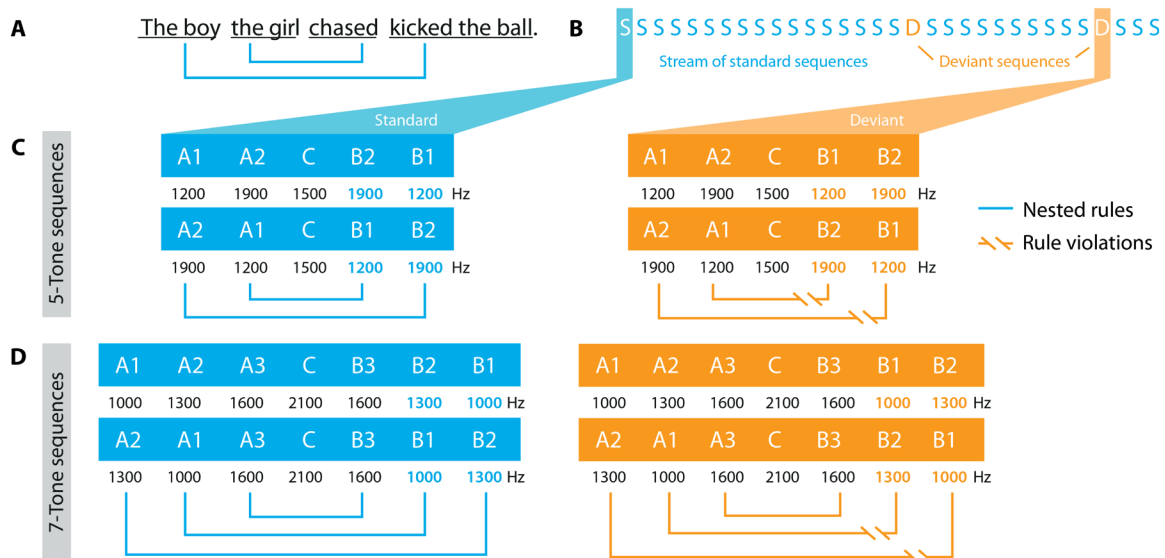
<sup>1</sup>Department of Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstr. 1a, 04103 Leipzig, Germany. <sup>2</sup>International Max Planck Research School on Neuroscience of Communication (IMPRS NeuroCom), Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstr. 1a, 04103 Leipzig, Germany. <sup>3</sup>Institute of Cognitive Science, University of Osnabrück, Wachsbleiche 27, 49090 Osnabrück, Germany. <sup>4</sup>Department of Neurology, Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstr. 1a, 04103 Leipzig, Germany. <sup>5</sup>Clinic for Cognitive Neurology, Medical Faculty, University Leipzig, Liebigstr. 16, 04103 Leipzig, Germany.

\*Corresponding author. Email: maennel@cbs.mpg.de



**Fig. 1. Illustration of different sequential regularities with increasing structural complexity.** (A) Example of a two-element sequence implementing a linear structure with adjacent relations. (B) Example of a three-element sequence implementing a linear structure with nonadjacent relations. (C) Example of a four-element sequence implementing an embedded structure with nested relations. The processing of the outer dependency (red) is temporarily interrupted (1) for processing an inner dependency (black) of a similar kind. (D) Example of a six-element sequence implementing an embedded structure with multiple nested relations. The processing of each outer dependency (yellow and red) is temporarily interrupted (1, 2) for processing an inner dependency of a similar kind (red and black).

free grammars, however, are equivalent to pushdown automata, which involve an additional memory component and allow neighboring elements to not directly depend on each other (6, 14). Thus, in contrast to linear adjacent and nonadjacent dependencies (Fig. 1, A and B), nested dependencies require processing of the inner dependency rule (embedded) before completion of the outer dependency rule (Fig. 1, C and D). The elements coding these relations in spoken language are speech units, realized as spectrally complex, dynamic sounds. However, syntactic relations in natural language require additional processes, because these relations are established not only between sequentially ordered sounds but also between syntactically categorized lexical items, for example, between a noun and a verb (see Fig. 2A). To acquire grammatical relations between lexical elements during language acquisition, the corresponding speech sound regularities in the input stream must first be decoded. This renders the examination of nested relations in auditory sequences a promising starting point for the empirical investigation of the developmental origin of complex regularity processing.



**Fig. 2. Experimental paradigm and sequence structures.** (A) Example of a center-embedded structure from natural language containing nested dependencies (blue lines), forming one level of embedding. (B) Illustration of the oddball paradigm containing tone sequences as frequent standards (S; blue) and infrequent deviants (D; orange). (C) Examples of five-tone sequences (experiment 1): Blue sequences indicate both standard forms, involving nested dependencies, which implement one level of center-embedding. A1, A2, B1, and B2 represent the rule-defining tones, and C is the center-marker at 1500 Hz. Orange sequences indicate both deviant forms, which violate the nested rules by exchanging the order of the last two tones. Note that the last two tones that define the rule violation in the deviant have previously appeared in the standards' second form and are thereby not informative about the rule per se. (D) Examples of seven-tone sequences (experiment 2): Blue sequences indicate both standard forms, involving nested dependencies, which implement two levels of center-embedding. A1, A2, A3, B1, B2, and B3 represent the rule-defining tones, and C is the center-marker at 2100 Hz. Orange sequences indicate both deviant forms, which violate the nested rules by exchanging the order of the last two tones.

Infants have been shown to exhibit astonishing speech decoding abilities from early on, detecting linguistic units in the speech stream and extracting their interrelations. For example, infants successfully segment syllables (15), words (15, 16), and clauses (17–19) from continuous speech at around half a year of age. In addition to these segmentation abilities, infants have been ascribed impressive abilities in processing the dependencies between different linguistic units. The first groundbreaking behavioral evidence of infants' learning of novel linguistic regularities, from mere listening, were reports of statistical learning showing that 8-month-olds were able to process transitional probabilities between adjacent linguistic or nonlinguistic elements (20, 21). More recently, neurophysiological measures could evidence this ability shortly after birth (10, 12, 13). Evidence for the extraction of nonadjacent dependencies between linguistic units has been observed in behavioral studies starting from 12 months of age and becoming more stable at around 18 months of age (22–25). As for adjacent dependencies, neurophysiological evidence for nonadjacent dependency learning suggests an earlier onset of this ability, namely 3 to 4 months (9, 11). The ontogeny of processing embedding, involving nested dependencies, has yet to be examined. The present study uses center-embedding as a prototypical example of a nested structure in language (see Fig. 2A). Evidence of the processing of center-embedding in natural language suggests that these regularities are understood by children after the age of 5 years (26, 27) but can still be challenging for adults, especially with increasing levels of embedding (28, 29). Similarly, artificial grammar learning (AGL) studies in adults have reported learning failure (30–32) unless sufficient perceptual cues were provided (33, 34), which undermines the challenge of learning center-embedded structures. Nonetheless, given infants' impressive input decoding abilities, we here investigate whether the core ability to process embedding can

be traced back to infant age, which might be possible by using behavior-independent neurophysiological methods and a paradigm that reduces computational demands. Thus, we hypothesize that infants will demonstrate the ability to identify embedding violations. Such a finding would be consistent with the extraction of nested dependencies from the auditory signal as a core cognitive ability at the root of human language.

In AGL studies, the reported failures to process certain structures may not stem from a lack of computational abilities per se but from other contributing factors, such as limitations of attention and memory capacities (35) or perceptual demands of the rule-coding elements (34). We therefore used the passive-listening oddball paradigm, an experimental setup that minimizes long-term memory requirements, originally used to demonstrate the preattentive processing of simple and complex regularities in the auditory domain (36–38). In this paradigm, a stream of frequent standard stimuli is occasionally interrupted by the occurrence of an infrequent deviant stimulus, varying from the standards on one or more feature dimensions. The resulting event-related potential (ERP) in adults, the mismatch negativity, is taken as a marker of preattentive feature discrimination. This paradigm has been widely applied in the study of infant auditory processing and language abilities (39, 40), and infant mismatch responses can be measured from birth (40, 41).

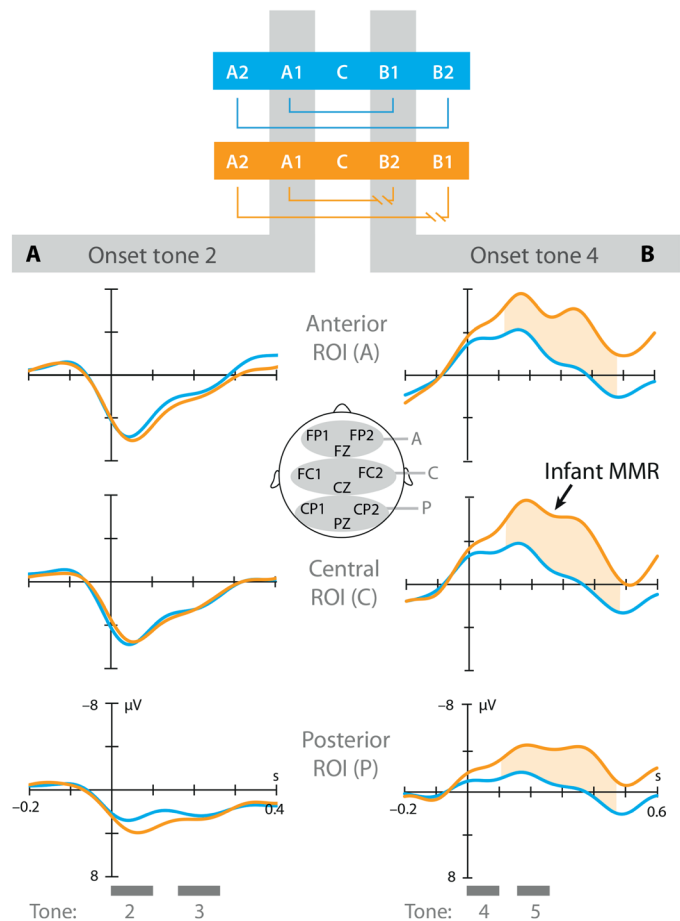
To increase the likelihood of infant learning of center-embedding, we minimized perceptual and memory requirements by using an oddball paradigm with perceptually simple tone stimuli as rule-coding element sets and identity relations linking these elements. The resulting structure can be classified as a mirror grammar, which establishes the nested dependencies in an item-wise manner, ruling out simpler solving strategies, such as count-and-compare (6, 14). In addition, to ensure that infants could not just use a strategy of

detecting unfamiliar repetition patterns [cf. (30)], the middle position of each structure was held constant by using the same frequency tone. The embeddings were implemented as sine tone sequences. The sequences as a whole, in contrast to single tones, defined the standard and deviant stimuli within an oddball design (Fig. 2B). Standard sequences established the embeddings by defining nested relations between the rule elements. Deviant sequences were characterized by violations of the nested dependencies induced by a reversion of the two final elements. To probe the complexity limits of nested dependency processing, we tested both five-tone sequences with one center-embedding (experiment 1, Fig. 2C) and seven-tone sequences with two center-embeddings (experiment 2, Fig. 2D). Given that deviants differed from standards only in the validity of the nested dependencies, the presence of mismatch responses to rule violations would indicate infants' successful processing of these complex regularities by tracking the mirror structure.

## RESULTS

To evaluate infants' processing of nested dependencies, we used a repeated-measures analysis of variance (RM-ANOVA) to compare ERP responses locked to the rule-violating tones of the deviant sequences (fourth tone in experiment 1 and sixth tone in experiment 2), relative to those from the same position arising from the standard sequences. There were two factors: Condition (standards and deviants) and Region (anterior, central, and posterior). The 30-ms consecutive time-window analyses revealed the effects of Condition in the 120- to 480-ms time window for experiment 1 and in the 210- to 570-ms time window for experiment 2. The RM-ANOVA across the time window of 120 to 480 ms confirmed a main effect of Condition ( $F_{1,37} = 9.281, P = 0.004, \eta^2_{\text{partial}} = 0.201$ ), while no significant interaction of Condition  $\times$  Region was observed. Likewise, in experiment 2, for the time window of 210 to 570 ms, an effect of Condition ( $F_{1,37} = 8.235, P = 0.007, \eta^2_{\text{partial}} = 0.182$ ) was confirmed, and no significant interaction effect of Condition  $\times$  Region was observed. As an additional control for alpha-error accumulation resulting from multiple comparisons, we ran a time-domain, cluster-based permutation test (42) on the factor Condition using the FieldTrip toolbox for EEG/MEG analyses (Donders Institute for Brain, Cognition and Behaviour; [www.ru.nl/neuroimaging/fieldtrip](http://www.ru.nl/neuroimaging/fieldtrip)). We performed cluster statistics on a trial length of 0 to 600 ms (with 301 time points) across the single-participant average of all electrode sites. We ran 10,000 permutations and dependent sample *t* tests for each time point with a cluster-forming threshold of  $P < 0.05$  to determine the cluster mass. These cluster-based permutation tests confirmed the reported ANOVA effects of Condition for each experiment (experiment 1, 162 to 460 ms, cluster value  $P = 0.012$ ; experiment 2, 222 to 558 ms, cluster value  $P = 0.006$ ), showing that alpha-error accumulation from multiple comparisons cannot explain the reported effects.

We performed additional control analyses relative to the onsets of the tones preceding the center-marker (second tone in experiment 1 and third tone in experiment 2). However, no effect of Condition was observed in either experiment when comparing ERPs locked to the tones preceding the center-markers (second tone in experiment 1 and third tone in experiment 2), for which the validity of the underlying dependencies did not differ between standards and deviants. Figure 3 shows the grand-average ERP responses for standards and deviants relative to the onsets of the second (control

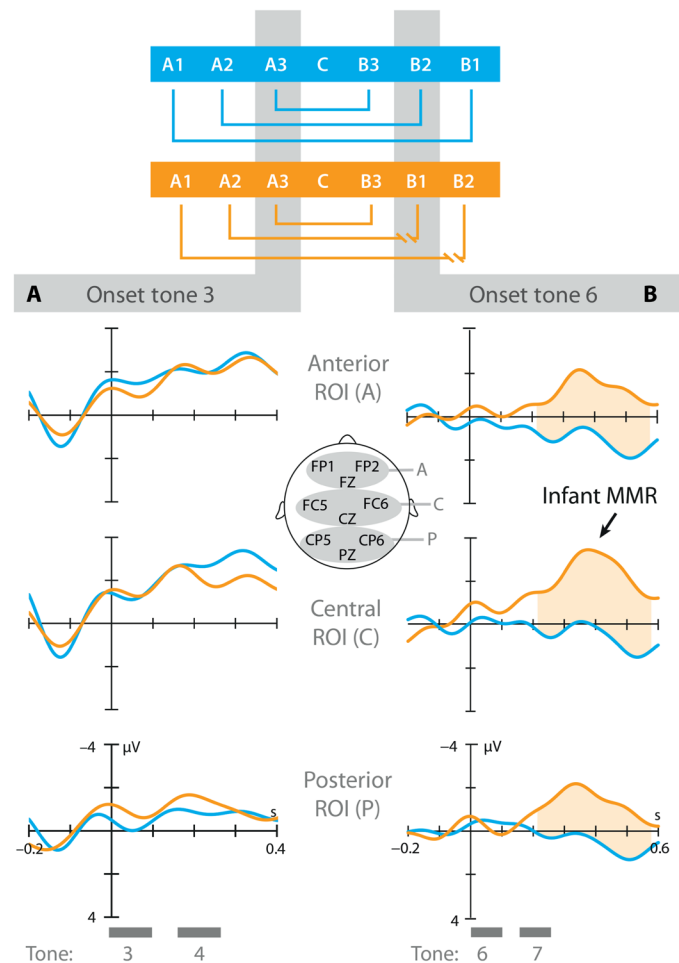


**Fig. 3. ERP results of experiment 1.** (A) Control analysis: Grand-average ERP responses for standards (blue) plotted against rule deviants (orange) for anterior (A), central (C), and posterior (P) regions of interest (ROI). Negativity is plotted upward. The y axis corresponds to onset of tone 2. Timing of rule-conforming tones 2 and 3 is illustrated by gray rectangles at the bottom of the illustration. (B) Analysis of rule violation: Grand-average ERP responses for standards (blue) plotted against rule deviants (orange). Negativity is plotted upward. Window of significant condition effects (120 to 480 ms) is highlighted in light orange. The y axis corresponds to onset of tone 4. Timing of rule-violating tones 4 and 5 is illustrated by gray rectangles at the bottom of the illustration. MMR, mismatch response.

analysis, Fig. 3A) and the fourth tones (Fig. 3B) for the five-tone sequences of experiment 1. The same information but relative to the onsets of the third and the sixth tones of the seven-tone sequences of experiment 2 is shown in Fig. 4. Thus, the observed condition effects of both experiments appear as negative mismatch responses, broadly distributed over infants' scalps, with more negative responses for rule-violating than standard sequences.

## DISCUSSION

We investigated whether preverbal infants have the core cognitive ability to encode embeddings of varying complexity in auditory sequences. The ERP results of two experiments revealed infant mismatch responses, indicating processing differences between standard sequences involving nested relations and deviant sequences violating those relations. Crucially, these processing differences only occurred at the position of the rule violation, and not at a preceding



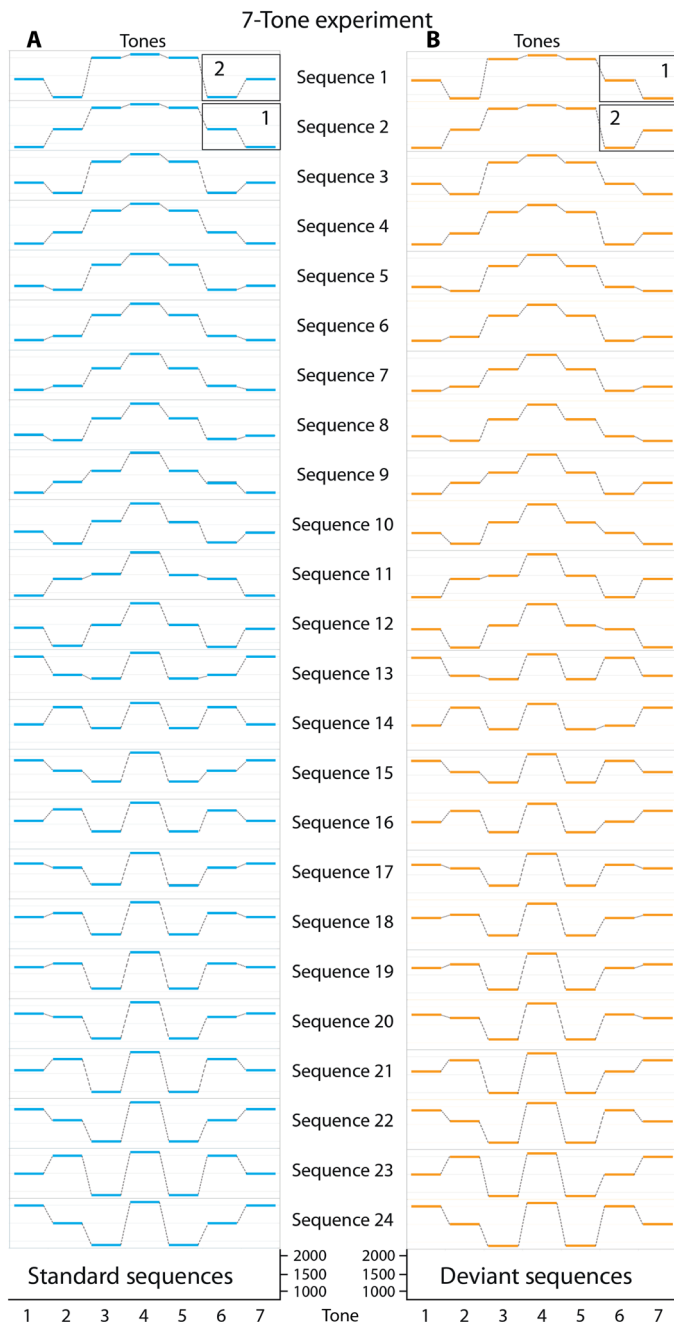
**Fig. 4. ERP results of experiment 2.** (A) Control analysis: Grand-average ERP responses for standards (blue) plotted against rule deviants (orange) for anterior (A), central (C), and posterior (P) regions of interest (ROI). Negativity is plotted upward. The y axis corresponds to onset of tone 3. Timing of rule-conforming tones 3 and 4 is illustrated by gray rectangles at the bottom of the illustration. (B) Analysis of rule violation: Grand-average ERP responses for standards (blue) plotted against rule deviants (orange). Negativity is plotted upward. Window of significant condition effects (210 to 570 ms) is highlighted in light orange. The y axis corresponds to onset of tone 6. Timing of rule-violating tones 6 and 7 is illustrated by gray rectangles at the bottom of the illustration.

sequence position. Mismatch responses occurred for both complexity levels, involving one or two levels of embedding, albeit with different latencies. The fact that the mismatch responses started approximately 90 ms later, for seven-tone relative to five-tone sequences, might have resulted from both the increased complexity of the underlying structure and the respective processing difficulty. Previous studies using oddball paradigms have reported delayed onset latencies of mismatch responses for the discrimination of abstract regularities, compared to simple physical features [see (43)], as well as for increased discrimination difficulty from less distinct stimulus differences (44). The processing of two levels, as opposed to one level, of embedding is more demanding, given the higher number of sequence items that have to be kept in memory (i.e., seven compared to five tones) and the distance between the corresponding items. In dealing with three nested relations, two levels must be

temporarily interrupted and held in memory while processing the inner relation. Only then can the outer two be successively closed. Our findings indicate that 5-month-old infants were able to process the nested dependencies between tones, likely guided by the symmetry inherent in the mirror structure of the tone sequences used. Thus, our results show that the ability to process embedding—a core computational mechanism—may be present from very early on in human ontogenesis and not depend on language skills per se. Infants might have a specific proclivity toward decomposing high-level auditory structure, which would serve as an important building block for the later acquisition and processing of syntactic structure.

The present findings of infants' processing of nested relations may seem unexpected in the light of children's apparent late acquisition of center-embedding (26) and simpler, nonadjacent dependencies (45) in natural language. Even adults have been observed to encounter difficulties when processing more than two levels of center-embedding in the syntax of natural language (29). These contradictory findings of infants' and adults' capacities might be, in part, explained by differences in computational demands presupposed by particular experimental paradigms and stimulus features. Results by Frank and Gibson (35) on adult AGL suggest that the inability to maintain stimuli in memory long enough to learn the underlying rule might explain the null results of some studies. When the learning task was modified, such that the individual strings could be kept in memory longer, learning success increased. Similarly, in the current study, using a task with no long-term memory requirements (as testing was integrated in the learning phase via the oddball design) and limited short-term memory requirements (as even the seven-tone sequences were less than 2 s long) might have enabled the learning success at an unusually early age. Furthermore, the use of an oddball paradigm instead of a classical AGL paradigm might have contributed to the observed learning effect. In AGL paradigms, longer learning phases are followed by testing phases, which include several rule-violating elements in a row, whereas the oddball paradigm presents only one violating element at a time, with the subsequent element reestablishing the rule. It is conceivable that the former design prevents potential learning effects, within testing phases, because learning of violations as new rules overrides the effects of initial rule learning. Studies testing the consolidation of nonadjacent dependencies have shown relearning effects induced by testing items (46, 47). Furthermore, the present paradigm used a mirror structure with identity relations between nested rule elements, which are computationally easier to process than categorical relations in AGL or natural language paradigms. Accordingly, our current approach will need to be adapted stepwise to more natural and computationally demanding learning conditions to evaluate the conditions under which the current findings apply. Together, we propose that our use of a minimalistic oddball paradigm has unveiled the computational mechanism involved in the processing of embedding in early infancy.

In view of our findings of infants' early regularity-processing capacities, it is important to note that the design of the current paradigm can rule out several alternative processing strategies. First, the large number of different stimulus sequences (see Fig. 5 for an illustration of sequence variability in experiment 2) renders it unlikely that infants memorized individual sequences, rather than deriving the underlying rule. Moreover, we observed mismatch responses to the onset of the tones violating the inner center-embedding rule in experiment 1 (see Figs. 3 and 4). This rules out the possibility that



**Fig. 5. Sequence variability of experiment 2.** (A) Standard sequences (blue): Variability of 24 standard sequences illustrated by blue bars at the tone frequency (in Hz, y axis) for each tone of the sequence (x axis). Each inset row depicts one sequence. The outlined rectangles at the first and second sequences of the deviants and standards illustrate that the rule-violating part of the deviant is also present in the corresponding standards' second form. (B) Deviant sequences (orange): Variability of 24 deviant sequences illustrated by orange bars at the tone frequency (in Hz, y axis) for each tone of the sequence (x axis). Each inset row depicts one sequence.

infants only processed the first and last elements of the sequence [see (48) for this strategy], ignoring the middle elements. Likewise, the objection that infants only processed the inner, but not the outer, center-embedding rule is rebutted by the outcome of experiment 2, involving seven-tone sequences with two levels of embedding, for

which the observed mismatch responses occurred at the first outer embedding. Last, by using a mirror grammar, which establishes item-wise center-embedding, simpler mechanisms such as count-and-compare can be ruled out [see (6, 14) for this strategy].

Future studies need to test whether infants' attention toward nested relations depends on the presence of mirror structures, which, by their inherent symmetry, may influence salience or valence, or whether this ability extends to nested relations based on different structures. In this context, it is also worthwhile to evaluate different items or item categories. Furthermore, future studies will have to further pinpoint the ontogenetic trajectory of nested dependency processing after birth and specify the conditions under which nested dependencies can be acquired. Given that, by the nature of embedded structures, at least the outer dependencies are nonadjacent and involve more than one intervening element, we assume that the processing of nested dependencies requires infants' ability to process nonadjacent dependencies. The processing of adjacent dependencies has been observed from birth (10, 12, 13), and that of nonadjacent dependencies has been evidenced starting from the age of 3 to 4 months in AGL (9) and in natural grammar learning (11). This suggests a gradual development of processing abilities, from simple to more complex structures.

In addition to the ontogeny of embedding processing, the phylogenetic trajectory is also of interest. Our paradigm could be used to target nested dependency processing in nonhuman species, for which equivalents of human-like mismatch responses have been demonstrated, for example, in macaques (49), rats (50), and pigeons (51). Along those lines, a recent ERP study investigated primate precursors of complex auditory sequence processing in macaque monkeys (52). Macaques were presented with nonadjacent dependencies in trisyllabic sequences of human speech, while ERPs were recorded from the surface of their scalps. After a relatively long learning period, similar ERP responses were found in macaques as have previously been seen in human infants [see (9)]. The authors concluded that an important prerequisite for the processing of nested dependencies may already be in place in our primate relatives (52). However, to date, there is no compelling evidence of dependency learning beyond the level of simple, nonadjacent relations, such as nested dependencies, in nonhuman animals [see (53) for negative evidence and (54, 55) for positive evidence, which is controversially discussed in (56, 57)]. Similarly, cross-species comparisons in birds have not found grammars of a comparable complexity (57–59), although birdsong can include quite complex sequential regularities. Thus, future comparative studies using the current computationally minimized oddball paradigm could determine whether nonhuman primates or other species are able to process embedding and, if so, potentially serve to identify the phylogenetic roots of precursors of human higher cognition and language.

## MATERIALS AND METHODS

### Experimental design

#### Participants

All infants were recruited from the database of the Max Planck Institute for Human Cognitive and Brain Sciences in Germany. For experiment 1, 65 healthy, monolingual, German infants with no known history of hearing deficits or neurological conditions were invited for testing. Data from four infants had to be excluded because of short recording time as a result of noncompliance. A further 23 datasets

were excluded because of a high rate of movement or perspiration artifacts (inclusion criteria:  $\geq 25\%$  of deviant trials remaining after artifact rejection). All infants of the final group ( $n = 38$ ; 18 females) were born full term ( $M = 40.34$  gestation weeks,  $SD = 1.12$ ; all infants  $>38$  gestation weeks), with a normal birth weight ( $M = 3524.61$  g,  $SD = 417.13$ ), and had a mean age of 155.21 days, ranging from 145 to 165 days. The participants of experiment 2 were selected analogously to experiment 1, such that 74 different infants were invited for testing. Following the previous inclusion criteria, 7 datasets were excluded for short recording time and 28 datasets were excluded for high artifact rate. One dataset had to be excluded because of technical problems during recording, leaving a total of 38 datasets (16 females). All infants of this group were born full term ( $M = 40.68$  gestation weeks,  $SD = 1.12$ ), with a normal birth weight ( $M = 3623.32$  g,  $SD = 378.03$ ), and had a mean age of 155.84 days, ranging from 142 to 165 days. The experiments were approved by the Ethics Committee of the University of Leipzig and conducted in accordance with the Declaration of Helsinki (2008). Written informed consent was obtained from the infants' accompanying parents. Parents were reimbursed for their travel expenses and were invited to choose a toy gift for their infant.

### Study design

In experiment 1, we used a passive listening oddball paradigm with five-tone sequences as the standard and deviant elements. The frequently presented standard elements featured center-embedded rules (94.74% of trials; 12 different sequences, each presented 72 times). The deviant elements, appearing rarely (5.26% of trials; 12 different sequences, each presented four times), contained violations of the embedded rules of the standard elements. In each rule-conforming standard sequence, the center-embedding was realized such that the first tone predicted the fifth tone, the second tone predicted the fourth tone, while the third tone was an invariable center-marker at 1500 Hz (Fig. 2C). The rule-violating deviant sequences were defined by a positional exchange of the fourth and fifth tones. The frequency distances between the fourth and fifth (and accordingly, the first and second) tones were held constant at 700 Hz to maintain the discrimination difficulty between rule-defining tones across all sequences. Crucially, for each of the 12 standard sequences (see table S1 for a list of all standard sequences), a corresponding deviant sequence was built (see table S2 for a list of all deviant sequences), such that only the last two tones of a given sequence were informative of its rule-conforming or rule-violating character. Furthermore, for each standard sequence, we also used an inverted second form to ensure that each part of a deviant (including the positional violations) was also presented within a standard sequence. This was done to ensure that certain tone combinations were not informative about the rule per se.

The standard and deviant sequences were presented in a pseudo-randomized order, ensuring that the appearance of sequence type was optimally balanced throughout the experiment. We combined sequences into conceptual units of jittered lengths, which allowed us to keep deviants sufficiently far apart from each other. This maintained the establishment of the nested rules by the standards, while deviant occurrences remained unpredictable. Each unit contained 7 to 12 sequences and was led either by a standard (50% of cases) or a deviant (50% of cases). No more than three units of the same type (standard-first and deviant-first) were permitted to follow one another throughout the experiment. At the beginning of the experiment, three units containing only standard sequences were

presented to establish the nested rules. After that, no deviant sequence was directly preceded or followed by its corresponding standard sequence (see Fig. 5). The frequency of occurrence of individual sequences and their second forms was balanced within each unit and across all units. To balance sequences beginning with rising or falling tone pairs, no more than two first or second forms of different sequences were allowed to follow each other within each unit. Each five-tone sequence lasted for 740 ms, with an intersequence interval of 566 ms.

Experiment 2 was designed analogously to experiment 1, but using seven-tone sequences instead of five-tone sequences as elements. Accordingly, in each rule-conforming standard sequence (94.74% of trials; 24 different sequences, each presented 36 times), the first tone predicted the seventh tone, the second tone predicted the sixth tone, and the third tone predicted the fifth tone. The fourth tone was an invariable center-marker at 2100 Hz (see table S3 for a list of all standard sequences). Rule-violating deviant sequences (5.26% of trials; 24 different sequences, each presented two times) involved a positional exchange of the sixth and seventh tones (see table S4 for a list of all deviant sequences). Each seven-tone sequence lasted for 1060 ms, and the intersequence intervals were kept at 566 ms.

### Stimuli

For both experiments, the individual tones were created in Praat as pure sine tones at a digitization rate of 44.1 kHz, with an onset and offset rise and fall time of 5 ms and a duration of 100 ms (with an intertone interval of 60 ms). The tone set included 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, and 2100 Hz. The resulting tone differences of a minimum of 100 Hz were easily recognizable by infants [see (60)].

### Procedure

During the EEG recordings, infants were seated in an electrically shielded and sound-attenuated room on their parent's lap, facing the speakers. The stimulus material was presented using the presentation software package (Neurobehavioral Systems Inc.) through a pair of ELAC (Electroacoustic GmbH) speakers, positioned 90 cm from the infant, at a comfortable, constant intensity level. Whenever needed, infants' compliance was maintained by showing a muted infant movie or by engaging the infant in silent play.

### EEG recording

For both experiments, a continuous EEG was recorded with in-house QRefa Acquisition Software, Version 1.0 beta (Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany) using a Refa amplifier system (Twente Medical Systems International B.V.). For experiment 1, the EEG was sampled at a rate of 500 Hz from 15 Ag/AgCl electrodes held on an elastic cap (EasyCap GmbH, Herrsching, Germany), according to standard positions (International 10-20 system of Electrode Placement): FP1, FP2, AFz, Fz, FC1, FC2, Cz, CP1, CP2, and Pz. Cz served as the online reference and an electrode at POz as common ground. To control for eye movements, a vertical electrooculogram (EOG) was recorded from FP2, and a single electrode was attached below the right eye. A horizontal EOG was derived from F9 and F10, placed at the outer canthi of each eye. Electrode impedances were predominantly kept below 20 k $\Omega$  (and always below 50 k $\Omega$ ). Simultaneously with the continuous EEG, near-infrared spectroscopy data were recorded from 44 channels in two bilateral 3  $\times$  5 optode grids in experiment 1, which are not included in the analysis reported here. Experiment 2 only used EEG and followed the same EEG recording procedures, but with the following standard electrode positions: FP1, FP2, F7, F3, Fz, F4, F8, FC5, FC6,

T7, C3, Cz, C4, T8, CP5, CP6, P7, P3, Pz, P4, P8, O1, O2, and the common ground at CP1.

### ERP processing

For both experiments, EEG data were processed offline using EEP 3.2.1 software package (ANT Software B.V.; Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany). First, the data were algebraically rereferenced to linked mastoids and band-pass filtered at 0.3 to 20 Hz (−3 dB, cutoff frequencies of 0.38 and 19.92 Hz). Second, EEG epochs (so-called trials) from −200 to 600 ms relative to stimulus onset of the critical tone (fourth tone of experiment 1 and sixth tone of experiment 2) were derived. Third, a semiautomatic artifact treatment procedure was applied. Prototypical blinks and eye movements were individually identified for each infant and used as a correction template for trials containing blink and eye movement artifacts (correlation-based correction algorithm). Regarding all other artifacts, trials with an SD exceeding 70  $\mu$ V within a 500-ms sliding window were automatically rejected. After artifact treatment, the remaining rule-conforming standard trials and rule-violating deviant trials were each averaged across participants, resulting in means of 197.13 standard trials (SD = 91.85) and 17.97 deviant trials (SD = 6.18) for experiment 1 and 272.95 standard trials (SD = 97.22) and 19.26 deviant trials (SD = 5.99) for experiment 2. Two additional control analyses were performed relative to the onsets of the tones preceding the center-markers, following the same procedure but with slightly shorter epochs (−200 to 500 ms) from the onset of tone 2 (experiment 1) and tone 3 (experiment 2). This resulted in means of 189.39 standard trials (SD = 85.03) and 17.37 deviant trials (SD = 5.05) for experiment 1 and 254.11 standard trials (SD = 98.90) and 17.68 deviant trials (SD = 5.05) for experiment 2.

### Statistical analysis

For the statistical analysis of the ERP data, regions of interest comparable across experiments were defined involving anterior (FP1, FZ, and FP2), central (FC1, CZ, and FC2), and posterior (CP1, PZ, and CP2) regions for experiment 1 and anterior (FP1, FZ, and FP2), central (FC5, CZ, and FC6), and posterior (CP5, PZ, and CP6) regions for experiment 2. For both studies, RM-ANOVAs were conducted in SPSS Software Version 22 (IBM, Walldorf, Germany) with the factors Conditions (standards and deviants) and Region (anterior, central, and posterior) for consecutive 30-ms time windows covering trial lengths of 0 to 600 ms. To be considered relevant, an effect of Condition was required to be statistically significant (at  $P < 0.05$ ) in at least four adjacent time windows.

### SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at <http://advances.sciencemag.org/cgi/content/full/4/11/eaar8334/DC1>

Table S1. Standard sequences of experiment 1.

Table S2. Deviant sequences of experiment 1.

Table S3. Standard sequences of experiment 2.

Table S4. Deviant sequences of experiment 2.

### REFERENCES AND NOTES

1. A. D. Friederici, Evolution of the neural language network. *Psychon. Bull. Rev.* **24**, 41–47 (2017).
2. A. D. Friederici, The brain basis of language processing: From structure to function. *Physiol. Rev.* **91**, 1357–1392 (2011).
3. S. Koelsch, M. Rohmeier, R. Torrecuso, S. Jentschke, Processing of hierarchical syntactic structure in music. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 15443–15448 (2013).
4. R. Friedrich, A. D. Friederici, Mathematical logic in the human brain: Syntax. *PLOS ONE* **4**, e5599 (2009).
5. W. T. Fitch, Toward a computational framework for cognitive biology: Unifying approaches from cognitive neuroscience and comparative cognition. *Phys. Life Rev.* **11**, 329–364 (2014).
6. W. T. Fitch, A. D. Friederici, Artificial grammar learning meets formal language theory: An overview. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **367**, 1933–1955 (2012).
7. W. T. Fitch, Empirical approaches to the study of language evolution. *Psychon. Bull. Rev.* **24**, 3–33 (2017).
8. N. Chomsky, *Syntactic Structures* (Mouton Publishers, 1957).
9. 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).
10. H. Bulf, S. P. Johnson, E. Valenza, Visual statistical learning in the newborn infant. *Cognition* **121**, 127–132 (2011).
11. 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).
12. N. Kudo, Y. Nonaka, N. Mizuno, K. Mizuno, K. Okanoya, On-line statistical segmentation of a non-speech auditory stream in neonates as demonstrated by event-related brain potentials. *Dev. Sci.* **14**, 1100–1106 (2011).
13. T. Teinonen, V. Fellman, R. Näätänen, P. Alku, M. Huotilainen, Statistical language learning in neonates revealed by event-related brain potentials. *BMC Neurosci.* **10**, 21 (2009).
14. G. Jäger, R. Rogers, Formal language theory: Refining the Chomsky hierarchy. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **367**, 1956–1970 (2012).
15. L.-L. Nishibayashi, L. Goyet, T. Nazzi, Early speech segmentation in French-learning infants: Monosyllabic words versus embedded syllables. *Lang. Speech* **58**, 334–350 (2015).
16. P. W. Jusczyk, D. M. Houston, M. Newsome, The beginnings of word segmentation in English-learning infants. *Cogn. Psychol.* **39**, 159–207 (1999).
17. T. Nazzi, D. G. Kemler-Nelson, P. W. Jusczyk, A. M. Jusczyk, Six-month-olds' detection of clauses embedded in continuous speech: Effects of prosodic well-formedness. *Infancy* **1**, 123–147 (2000).
18. A. Seidl, Infants' use and weighting of prosodic cues in clause segmentation. *J. Mem. Lang.* **57**, 24–48 (2007).
19. M. Soderstrom, D. G. K. Nelson, P. W. Jusczyk, Six-month-olds recognize clauses embedded in different passages of fluent speech. *Infant Behav. Dev.* **28**, 87–94 (2005).
20. J. R. Saffran, E. K. Johnson, R. N. Aslin, E. L. Newport, Statistical learning of tone sequences by human infants and adults. *Cognition* **70**, 27–52 (1999).
21. J. R. Saffran, E. L. Newport, R. N. Aslin, Word segmentation: The role of distributional cues. *J. Mem. Lang.* **35**, 606–621 (1996).
22. R. L. Gómez, J. Maye, The developmental trajectory of nonadjacent dependency learning. *Infancy* **7**, 183–206 (2005).
23. J. Lany, R. L. Gómez, Twelve-month-old infants benefit from prior experience in statistical learning. *Psychol. Sci.* **19**, 1247–1252 (2008).
24. R. L. Gómez, Variability and detection of invariant structure. *Psychol. Sci.* **13**, 431–436 (2002).
25. A. Kerkhoff, E. De Bree, M. De Klerk, F. Wijnen, Non-adjacent dependency learning in infants at familial risk of dyslexia. *J. Child Lang.* **40**, 11–28 (2013).
26. E. Kidd, E. L. Bavin, English-speaking children's comprehension of relative clauses: Evidence for general-cognitive and language-specific constraints on development. *J. Psycholinguist. Res.* **31**, 599–617 (2002).
27. A. Fengler, L. Meyer, A. D. Friederici, How the brain attunes to sentence processing: Relating behavior, structure, and function. *Neuroimage* **129**, 268–278 (2016).
28. F. Karlsson, Constraints on multiple center-embedding of clauses. *J. Linguist.* **43**, 365–392 (2007).
29. M. S. Blauergs, M. D. Braine, Short-term memory limitations on decoding self-embedded sentences. *J. Exp. Psychol.* **102**, 745–748 (1974).
30. M. H. de Vries, P. Monaghan, S. Knecht, P. Zwitserlood, Syntactic structure and artificial grammar learning: The learnability of embedded hierarchical structures. *Cognition* **107**, 763–774 (2008).
31. J.-R. Hochmann, M. Azadpour, J. Mehler, Do humans really learn  $A^nB^n$  artificial grammars from exemplars? *Cogn. Sci.* **32**, 1021–1036 (2008).
32. P. Perruchet, A. Rey, Does the mastery of center-embedded linguistic structures distinguish humans from nonhuman primates? *Psychon. Bull. Rev.* **12**, 307–313 (2005).
33. J. Lai, F. H. Poletiek, The impact of adjacent-dependencies and staged-input on the learnability of center-embedded hierarchical structures. *Cognition* **118**, 265–273 (2011).
34. J. L. Mueller, J. Bahlmann, A. D. Friederici, Learnability of embedded syntactic structures depends on prosodic cues. *Cogn. Sci.* **34**, 338–349 (2010).
35. M. C. Frank, E. Gibson, Overcoming memory limitations in rule learning. *Lang. Learn. Dev.* **7**, 130–148 (2011).
36. A. Bendixen, E. Schröger, W. Ritter, I. Winkler, Regularity extraction from non-adjacent sounds. *Front. Psychol.* **3**, 143 (2012).
37. E. Schröger, A. Bendixen, N. J. Trujillo-Barreto, U. Roeber, Processing of abstract rule learning in audition. *PLOS ONE* **2**, e1131 (2007).
38. R. Näätänen, P. Paavilainen, T. Rinne, K. Alho, The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clin. Neurophysiol.* **118**, 2544–2590 (2007).

39. C. He, L. Hotson, T. J. Trainor, Mismatch responses to pitch changes in early infancy. *J. Cogn. Neurosci.* **19**, 878–892 (2007).
40. I. Winkler, E. Kushnirenko, J. Horváth, R. Čeponienė, V. Fellman, M. Huotilainen, R. Näätänen, E. Sussman, Newborn infants can organize the auditory world. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 11812–11815 (2003).
41. M. Cheour-Luhtanen, K. Alho, T. Kujala, K. Sainio, K. Reinikainen, M. Renlund, O. Aaltonen, O. Eerola, R. Näätänen, Mismatch negativity indicates vowel discrimination in newborns. *Hear. Res.* **82**, 53–58 (1995).
42. E. Maris, R. Oostenveld, Nonparametric statistical testing of EEG- and MEG- data. *J. Neurosci. Methods* **164**, 177–190 (2007).
43. P. Paavilainen, The mismatch-negativity (MMN) component of the auditory event-related potential to violations of abstract regularities: A review. *Int. J. Psychophysiol.* **88**, 109–123 (2013).
44. G. P. Novak, W. Ritter, H. G. Vaughan Jr., M. L. Wiznitzer, Differentiation of negative event-related potentials in an auditory discrimination task. *Electroencephalogr. Clin. Neurophysiol.* **75**, 255–275 (1990).
45. B. Höhle, 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 (2006).
46. A. Hupbach, R. L. Gómez, R. R. Bootzin, L. Nadel, Nap-dependent learning in infants. *Dev. Sci.* **12**, 1007–1012 (2009).
47. R. L. Gómez, R. R. Bootzin, L. Nadel, Naps promote abstraction in language-learning infants. *Psychol. Sci.* **17**, 670–674 (2006).
48. A. L. Ferry, A. Fló, P. Brusini, L. Cattarossi, F. Macagno, M. Nespor, J. Mehler, On the edge of language acquisition: Inherent constraints on encoding multisyllabic sequences in the neonate brain. *Dev. Sci.* **19**, 488–503 (2016).
49. D. C. Javitt, C. E. Schroeder, M. Steinschneider, J. C. Arezzo, H. G. Vaughan, Demonstration of mismatch negativity in the monkey. *Electroencephalogr. Clin. Neurophysiol.* **83**, 87–90 (1992).
50. L. Harms, W. R. Fulham, J. Todd, T. W. Budd, M. Hunter, C. Meehan, M. Penttonen, U. Schall, K. Zavitsanou, D. M. Hodgson, P. T. Mitchie, Mismatch negativity (MMN) in freely-moving rats with several experimental controls. *PLOS ONE* **9**, e110892 (2014).
51. U. Schall, B. W. Müller, C. Kärger, O. Güntürkün, Electrophysiological mismatch response recorded in awake pigeons from the avian functional equivalent of the primary auditory cortex. *Neuroreport* **26**, 239–244 (2015).
52. A. E. Milne, J. L. Mueller, C. Männel, A. Attaheri, A. D. Friederici, C. I. Petkov, Evolutionary origins of non-adjacent sequence processing in primate brain potentials. *Sci. Rep.* **6**, 36259 (2016).
53. W. T. Fitch, M. D. Hauser, Computational constraints on syntactic processing in a nonhuman primate. *Science* **303**, 377–380 (2004).
54. T. Q. Gentner, K. M. Fenn, D. Margoliash, H. C. Nusbaum, Recursive syntactic pattern learning by songbirds. *Nature* **440**, 1204–1207 (2006).
55. K. Abe, D. Watanabe, Songbirds possess the spontaneous ability to discriminate syntactic rules. *Nat. Neurosci.* **14**, 1067–1074 (2011).
56. W. Zuidema, Context-freeness revisited, in *Proceedings of the 35th Annual Conference on Cognitive Science Society*, M. Knauff, M. Pauen, N. Sebanz, I. Wachsmuth, Eds. (Cognitive Science Society, 2013), pp. 1664–1669.
57. G. J. L. Beckers, J. J. Bolhuis, K. Okanoya, R. C. Berwick, Birdsong neurolinguistics: Songbird context-free grammar claim is premature. *Neuroreport* **23**, 139–145 (2011).
58. C. ten Cate, K. Okanoya, Revisiting the syntactic abilities of non-human animals: Natural vocalizations and artificial grammar learning. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **367**, 1984–1994 (2012).
59. R. C. Berwick, A. D. Friederici, N. Chomsky, J. J. Bolhuis, Evolution, brain, and the nature of language. *Trends Cogn. Sci.* **17**, 89–98 (2013).
60. L. W. Olsho, Infant frequency discrimination. *Infant Behav. Dev.* **7**, 27–35 (1984).

**Acknowledgments:** We thank all the participating infants and parents for supporting our research and the EEG laboratory team for help with data acquisition. We are thankful to M. Grigutsch for running the cluster-based permutation tests and J. Grant for proofreading and editing service. **Funding:** The study was funded by the Max Planck Society and the German Research Foundation (project MA 6897/2-1 awarded to C.M.). **Author contributions:** Study design: M.W., J.L.M., C.M., and A.D.F. Data collection: M.W. and C.M. Data analysis: M.W. and C.M. Writing of the manuscript: M.W., C.M., J.L.M., and A.D.F. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.

Submitted 22 December 2017

Accepted 24 October 2018

Published 21 November 2018

10.1126/sciadv.aar8334

**Citation:** M. Winkler, J. L. Mueller, A. D. Friederici, C. Männel, Infant cognition includes the potentially human-unique ability to encode embedding. *Sci. Adv.* **4**, eaar8334 (2018).



## Infant cognition includes the potentially human-unique ability to encode embedding

M. Winkler, J. L. Mueller, A. D. Friederici and C. Männel

*Sci Adv* 4 (11), eaar8334.  
DOI: 10.1126/sciadv.aar8334

### ARTICLE TOOLS

<http://advances.sciencemag.org/content/4/11/eaar8334>

### SUPPLEMENTARY MATERIALS

<http://advances.sciencemag.org/content/suppl/2018/11/16/4.11.eaar8334.DC1>

### REFERENCES

This article cites 58 articles, 7 of which you can access for free  
<http://advances.sciencemag.org/content/4/11/eaar8334#BIBL>

### PERMISSIONS

<http://www.sciencemag.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of Service](#)

---

*Science Advances* (ISSN 2375-2548) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. 2017 © The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. The title *Science Advances* is a registered trademark of AAAS.