Symbolic labeling in 5-month-old human infants

Claire Kabdebon[a,1] **and Ghislaine Dehaene-Lambertz**[a]

[a]Cognitive Neuroimaging Unit U992, Institut National de la Santé et de la Recherche Médicale, Commissariat à l'Énergie Atomique et aux Énergies Alternatives, Direction de la Recherche Fondamentale/Institut Joliot, Université Paris-Saclay, NeuroSpin Center, 91191, Gif/Yvette, France

Edited by Patricia K. Kuhl, University of Washington, Seattle, WA, and approved February 8, 2019 (received for review May 27, 2018)

Humans' ability to create and manipulate symbolic structures far exceeds that of other animals. We hypothesized that this ability rests on an early capacity to use arbitrary signs to represent any mental representation, even as abstract as an algebraic rule. In three experiments, we collected high-density EEG recordings while 150 5-month-old infants were presented with speech triplets characterized by their abstract syllabic structure—the location of syllable repetition—which predicted a following arbitrary label (e.g., ABA words were followed by a fish picture, AAB words by a lion). After a brief learning phase, EEG responses to novel words revealed that infants built expectations about the upcoming label based on the triplet structure and were surprised when it happened to be incongruent. Preverbal infants were thus able to recode the incoming triplets into abstract mental variables to which arbitrary labels were flexibly assigned. Importantly, infants also generalized to novel trials in which the pairing order was reversed (with the label preceding the auditory structure). Beyond conditioned associations, infants instantly inferred a bidirectional mapping between the abstract structures and the following label, a foundational operation for any symbolic system.

infants | language | learning | EEG | symbols

In many domains, humans use complex symbolic systems to describe abstract relations between sounds (music), quantities (arithmetic), commercial values (economics), individuals (uniforms), and so forth. Through symbolic representations, complex and multifaceted sensory inputs are transformed into discrete mental variables that are easier to handle, memorize, and communicate to others. The first and foremost of these systems is human language, in which a variety of sensory experiences are subsumed under the same arbitrary labels such as dog, thought, or repeat. But when and how do infants come to process arbitrary labels as symbolic representations? For language acquisition, it is classically assumed that infants undergo a first stage of slow associative learning during which they progressively stabilize the distributed connections relating events and labels before they become able to understand the symbolic value of words (1). This slow acquisition would parallel the extensive training required in nonhuman primates (2, 3), even chimpanzees (4), to acquire symbols. Only after this first stage would infants become able to recode the labeled events into mental variables and thus represent the underlying concept.

However, several observations suggest that the infant brain radically differs from the brain of other primates and that it might be equipped with powerful mechanisms for readily recoding perceptual input into abstract mental variables available for further operations, thereby increasing infants' learning power. First, even when animals acquire symbols, their success remains imperfect and imprecise (5), suggesting that the acquired representations might remain radically different from the sharpness of human symbolic representations. Second, the emergence of symbolic systems in humans has been related to the development of frontal areas and long-range tracts (5), allowing control operations on lower levels and integration between domains, whereas computations remain localized in domain-specific areas in macaques (6). Once thought to be poorly functional in infants, these regions are actually already active during the first months of life (7–9). Thus, infant cognition might benefit from their computational resources early on, despite their undeniable slowness due to immaturity (10). Third, an increasing number of studies show that infants learn associations between objects and words much earlier than originally presumed (11, 12) and that starting from the second semester of life, they are helped to categorize objects (13, 14) and sounds (15) and to track hidden objects (16) when a label, usually a word, is attached to the category/object. These results suggest that infants may use the label as an indication to look for a common referent in the different events.

Previous studies also explored preverbal abstraction abilities. Based on habituation/dishabituation paradigms, they revealed that 5- to 7-mo-old infants (17–19), and even neonates (20), react to a change in the localization of the repeated item in auditory as well as visual sequences (e.g., AAB words vs. ABA words: nonofe, gagalu, titina vs. rutaru, kemike, ladila). These results demonstrate that the developing brain captures some abstract attributes of the input. However, these experiments are limited to probing discrimination abilities for which a mere detector of immediate repetition, without an explicit representation of the sequence structure, is sufficient. It therefore remains unclear whether the abstracted information gets transformed into a unitary mental variable that is available for further computations or instead remains implicitly encoded in a mesh of cerebral connections. Both types of representations can translate into successful discrimination abilities. On that ground, the reported performance does not differ from that of many other animals, such as bees (21) and ducklings (22), which show remarkable abilities for detecting sameness between two elements. However, only information encoded as a unitary mental representation can be manipulated and passed along to further processing stages.

Significance

Humans naturally entertain complex representations of the world based on various symbolic systems, from natural language to mathematical or musical notation. They recode the input into abstract symbolic representations that can be internally manipulated and projected back onto the external world. We show that preverbal infants can redescribe complex percepts into abstract mental variables, which they can readily map onto arbitrary labels. Importantly, we show that, beyond associative learning, infants can readily infer a bidirectional relation between the abstracted representations and the associated labels, a capacity that animals do not spontaneously exhibit. Our findings buttress the hypothesis of symbolic representations in preverbal infants, which may serve as a foundation for our distinctively human learning abilities.

Author contributions: C.K. and G.D.-L. designed research; C.K. performed research; C.K. analyzed data; and C.K. and G.D.-L. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. Email: claire.kabdebon@gmail.com.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1809144116/-/DCSupplemental.

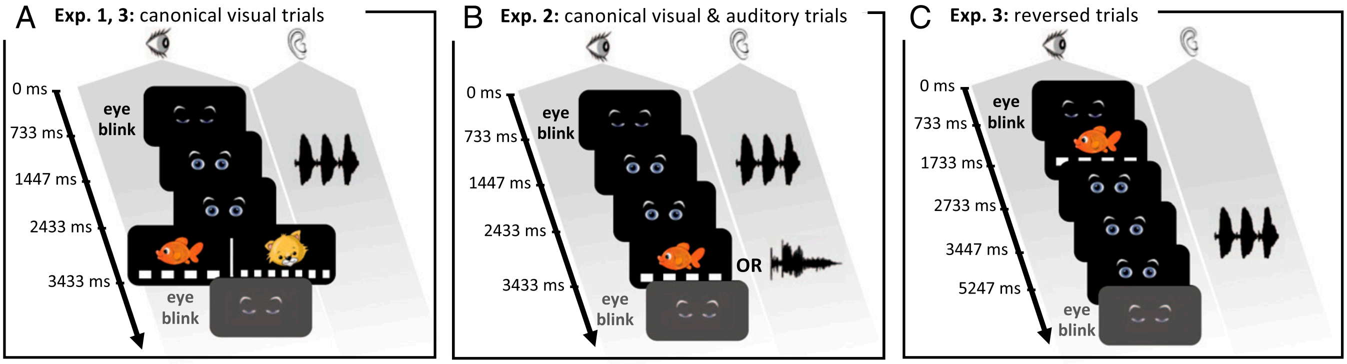

Fig. 1. Experimental paradigms. Trials began with blinking eyes. In canonical trials, a trisyllabic word was then presented, followed ∼1 s later by an image (a lion or a fish) presented on a flickering background (10 and 15 Hz) in experiments 1 and 3 (*A*), and by a fish or a word (schtroumpf) in experiment 2 (*B*). Infants first learned the association between two structures and their labels during a 36-trial exposure phase (e.g., ABA–lion and AAB–fish). During the subsequent test phase, incongruent structure–label pairs (e.g., ABA–fish and AAB–lion) and a new structure (e.g., ABB–fish and ABB–lion) were introduced in experiments 1 and 2. In experiment 3, the test phase consisted of an alternation of short blocks of six reversed trials (*C*) and longer blocks of 14 canonical trials (*A*). Canonical trials were all congruent, contrary to reversed trials, for which half were congruent and the other half incongruent.

Thus, the present study aims at clarifying the nature of preverbal representations in human infants. We reasoned that if the preverbal brain is equipped with a mechanism for recoding the incoming sensory information into abstract mental variables, it should be able to manipulate this abstract knowledge and concurrently entertain multiple representations. More specifically, in three associative learning tasks (conducted with 48, 57, and 45 infants), we tested 5-mo-old infants' abilities to extract and associate auditory abstract structures with arbitrary but stable images or sound. Infants were presented with trisyllabic words systematically followed by a label consisting of an image or a sound. Importantly, the structure of the words (AAB, ABA, or ABB) was predictive of the identity of the following label, and because words were instantiated with constantly varying syllables, the structure–label dependencies could not be discovered through basic associative-learning mechanisms. Infants had to process the abstract features of the triplets and recode each sequence into a unitary entity to detect the association with the following label. After testing infants' abilities to detect these higher-order associations, we explored the symbolic depth of the acquired association between the category-level structure and its associated label. We tested whether it was merely encoded as a temporal contingency or whether infants inferred a bidirectional mapping between the structure and the label. Symbolic mapping indeed entails a relation of symmetry: In contrast to unidirectional predictive relations, thoughts about symbols both elicit and are elicited by thoughts about their referents. Interestingly, while this symmetry relation is readily assumed by both human children and adults (23, 24), nonhuman animals repeatedly fail to infer the reverse relation after they have been trained with a unidirectional object–label relation (2–4).

Using high-density EEG, infants' learning abilities were assessed by exploring late event-related potential (ERP) brain responses to the introduction of rare incongruent structure–label pairings. Additionally, we used a frequency-tagging approach to reliably extract early sensory responses to visual labels and to explore whether this sensory activity was modulated by the pairing congruency. In experiments 1 and 2, infants' acquisition of the conditioned association was tested, preserving the temporal ordering of the structure–label pairing, while in experiment 3, the bidirectionality of the acquired association was assessed, reversing its temporal ordering for test items.

Results

Experiment 1. In experiment 1, after a short familiarization phase of 36 trials (Fig. 1*A*), learning was assessed during a test phase, when we introduced a small proportion of incongruent pairs (25%) and a novel neutral ABB structure, which was equally paired with both images (nonpredictive condition). Overall, for each infant, both images were presented with equal frequency. If infants successfully represented the abstract word structures and detected the associations, we predicted a late surprise response to incongruent pairs. Using nonparametric analyses, we observed a significant negative difference between congruent and incongruent images that slowly developed to become significant after 1 s on a cluster of right central electrodes (Monte Carlo $P_{\mathrm{corr}} = 0.036$ in 32 infants; Fig. 2*A*). This late response could reflect either the detection of semantic incompatibility between the preceding structure and the label (25, 26) or a more generic violation-of-expectations response (27). The case of neutral trials brings an additional insight: The neutral structure had no predictive value, so infants could not build any expectation or exhibit a violation-of-expectation response, but in terms of semantic relation, the ABB–image pairings were always incompatible and should be regarded as a semantic incongruency (25). Interestingly, the late ERP component was not present in response to these neutral trials. The recorded brain activity did not differ from the congruent condition [$t(31) = 1.06$, $P = 0.597$], but was significantly different from the incongruent condition [$t(31) = -2.84$, $P = 0.016$] (Fig. 2*A*). This late ERP component therefore reflects a violation-of-expectations response similar to the adult P300-type rather than an N400-type response.

In addition, neutral trials allowed us to ask whether infants had learned the associations based on a representation of the entire trisyllabic structure or whether they were only detecting immediate repetitions. If infants were sensitive only to the presence/absence of an immediate repetition, then AAB and ABB words should share the same mental representation and thus predict the same image. In terms of EEG, this assumption implies a violation-of-expectation response when the image presented after an ABB word is not the image they learned to follow AAB words. On the contrary, if infants extract an exhaustive representation of each triplet structure, they should successfully entertain all three structures as distinct mental representations and realize that ABB words are not predictive of any specific image. When contrasting brain responses to the two images following ABB words, we found no violation of expectations [$t(31) < 1$], indicating that infants were indeed representing the entire word.

Beyond late violation of expectations, we wondered whether the expectations that infants derived from the abstracted structures might also prime early visual activity for receiving a precise input (27, 28). Because each image was presented over a background

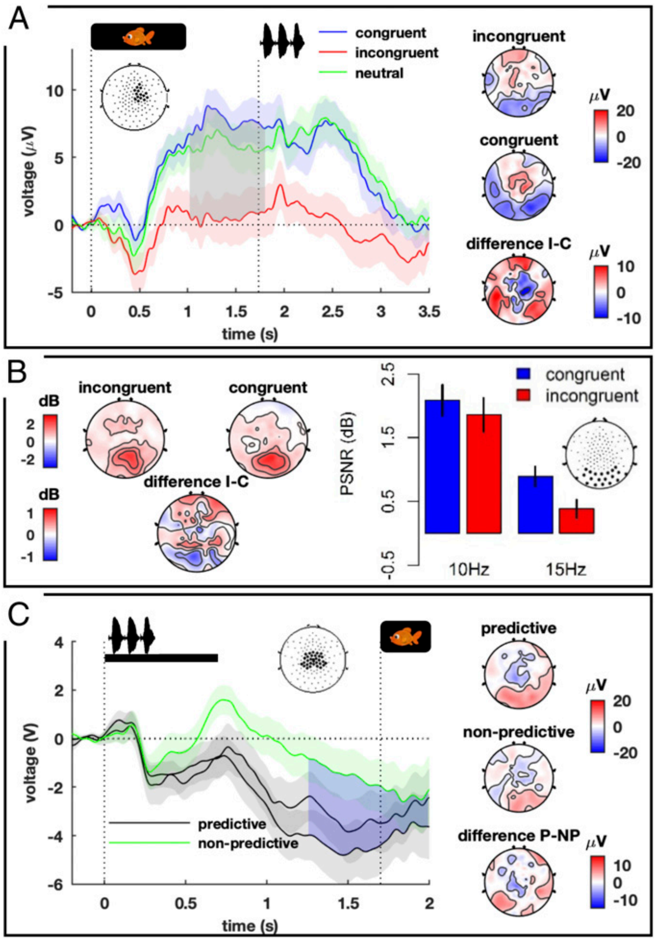

Fig. 2. Three neural signatures of learning in experiment 1. (*A*) Late violation of expectations. (*Left*) Grand average responses to congruent (blue), incongruent (red), and neutral (green) trials recorded from the significant cluster of electrodes presented on the graph. The vertical dotted line at 1.733 s indicates the onset of the next trial. (*Right*) ERP topographies for incongruent (I) and congruent (C) trials as well as their difference (I–C) averaged over the significant time window (gray area on the plot). Electrodes and time were identified using nonparametric analyses. (*B*) Early priming effect. (*Left*) PSNR topographies in response to the flickering background during label presentation for both conditions and their difference. (*Right*) Bars represent PSNR averaged over the occipital cluster in response to congruent (blue) and incongruent (red) trials for 10- and 15-Hz backgrounds. PSNR was significantly larger for congruent compared with incongruent trials. (*C*) CNV. (*Left*) Grand average responses to the two predictive structures (black) and the nonpredictive structure (green) recorded from central electrodes. The vertical dotted line at 1.700 s indicates the onset of the visual label. (*Right*) ERP topographies for the predictive (P) and nonpredictive (NP) trials as well as their difference (P–NP) averaged over the significant time window (blue area on the plot). The predictive structures (AAB and ABA) elicited a significantly larger CNV than the nonpredictive structure.

flickering at a specific frequency, we could inspect whether the strength of the entrained low-level visual activity was modulated by the congruency of the structure–image pair on a cluster of 20 occipitotemporal electrodes commonly used in the literature to study steady-state responses (29). A significant enhancement of cortical entrainment was observed for the expected frequency compared with the unexpected frequency [$F(1,31) = 7.97$, $P = 0.008$] (Fig. 2*B*; ERP analyses can be found in *SI Appendix*).

Furthermore, when observing brain responses before image onset, we observed a negative deflection over central electrodes that developed after the end of the triplet. This waveform corresponds to the contingent negative variation (CNV), commonly recorded after a cue predictive of a target stimulus in adults (30) as well as in infants (31). Its amplitude was affected by the triplet structure {39 infants; main effect of structure [$F(2,76) = 4.98$, $P = 0.009$]; Fig. 2*C*}: It was larger after the AAB and ABA words that were predictive of the following image than after the nonpredictive ABB words {AAB vs. ABB [$t(38) = -2.03$, $P = 0.049$]; ABA vs. ABB [$t(38) = -2.63$, $P = 0.012$]}, whereas there was no difference between AAB and ABA trials [$t(38) = -1.41$, $P = 0.168$].

Together, these results demonstrate that infants successfully represented all three structures concurrently, without confusion, and in a format that allowed for a subsequent, immediate association with an arbitrary image. For infants to succeed in our experimental design, it was not sufficient to locally discriminate between two structures as in classical habituation/dishabituation paradigms, nor to rely on the surface features of the triplets to detect the association; rather, infants had to have access to an abstract summary of the sequence of syllable repetitions and changes in order to couple it with the subsequent image. In other words, they had to redescribe the auditory input into one mental structure that could be further manipulated. The predictive activity observed during the waiting period between the offset of the triplet and the onset of the image may reflect this manipulation stage, offering infants the possibility to prospectively infer the associated image. Overall, experiment 1 demonstrates that preverbal infants form mental representations that are not limited to the direct output of sensory processes but rather involve a representational redescription stage (32) that makes knowledge available to the brain for other mental operations, a fundamental feature of symbolic representations. In experiments 2 and 3, we inspected two aspects of the relation between the generated mental representations and the arbitrary labels: flexibility and bidirectionality.

Experiment 2. Since human adults are remarkably flexible in associating any mental representation with virtually any type of symbol (words, numbers, graphs, pictures, shapes, etc.), we asked, in experiment 2, whether 5-mo-old infants were just as versatile and might accept labels in different sensory modalities (Fig. 1*B*). The procedure was the same as in experiment 1, except that the lion image was replaced by an auditory word. Congruency effects were analyzed separately for the visual and auditory modalities. The visual effects computed on the same spatiotemporal cluster replicated the findings of experiment 1 {[$t(33) = 2.06$, $P = 0.024$], one tailed t test}. A 2 × 2 ANOVA, with Experiment as a between-subjects factor and Condition as a within-subject factor, revealed a main effect of Condition [$F(1,64) = 15.86$, $P < 0.001$], no effect of Experiment [$F(1,64) = 2.58$, $P = 0.11$], and no Condition × Experiment interaction [$F(1,64) = 1.31$, $P = 0.26$]. Combining visual responses in experiments 1 and 2 (nonparametric statistics with no a priori) confirmed the common violation-of-expectations response over a right central cluster (Monte Carlo $P_{\mathrm{corr}} = 0.0024$ in 66 infants overall). In response to the auditory labels, nonparametric statistics computed on the same late time window revealed a significant violation-of-expectations signal over a more posterior right central cluster, peaking later than 1 s after label onset (for the auditory label, Monte Carlo $P_{\mathrm{corr}} = 0.048$ in 32 infants; Fig. 3) (see further analyses in *SI Appendix*).

We thus replicated in experiment 2 the late surprise effect induced by an incongruent pairing, even when the label modalities were mixed. Like toddlers, who accept any arbitrary form (gestures, pictograms, or sounds) as a label if embedded in a naming routine (33), infants were thus similarly able to flexibly associate structures with labels in multiple modalities, suggesting that any artifact, provided it consistently followed the triplet, could be used to represent the given structure.

Experiment 3. Contrary to conditioned associative learning in which the temporal ordering of the paired elements is crucial, symbolic mapping implies a relation of symmetry, supporting

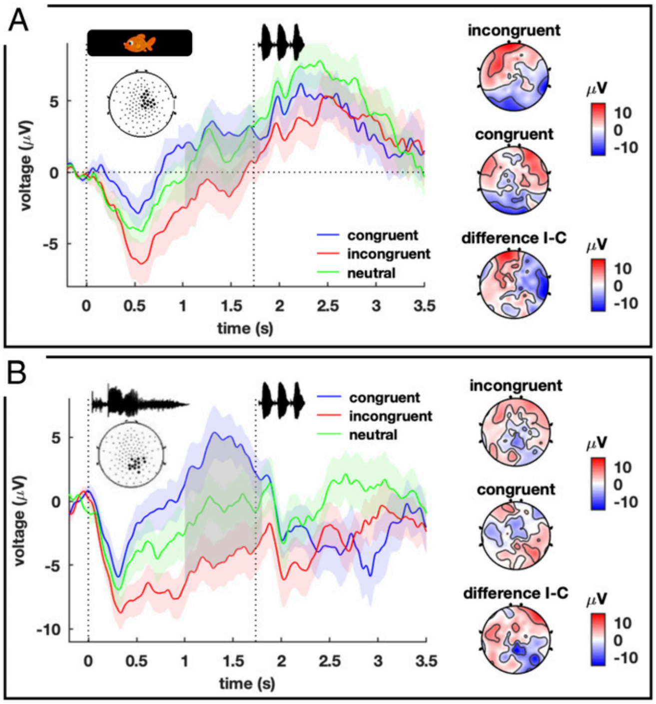

Fig. 3. Violation of expectations in experiment 2. (*A* and *B*, *Left*) Grand average for congruent (blue), incongruent (red), and neutral (green) trials recorded from the significant clusters of electrodes presented on the graphs for visual (*A*) and auditory labels (*B*). The vertical dotted line at 1.733 s indicates the onset of the trisyllabic word of the following trial. (*A* and *B*, *Right*) ERP topographies for incongruent (I) and congruent (C) trials as well as their difference (I–C) averaged over the significant time windows (gray area on the plots) for visual (*A*) and auditory labels (*B*). Electrodes and time were identified using nonparametric analyses.

bidirectional predictions. The ability to appreciate bidirectional mappings is readily observed in both human adults and children but appears to be particularly challenging for nonverbal species (2, 4, 23, 24, 34). This property was tested in experiment 3 (Fig. 1*C*). After the familiarization phase, infants were presented with short blocks of reversed trials (image–word) embedded within long blocks of congruent canonical trials (word–image). Half of the reversed trials were congruent and the other half incongruent, preventing any associative learning within reversed trials. Furthermore, words used different syllables in reverse and canonical trials. If infants possess a symbolic mapping mechanism that can operate over abstract representations, they should spontaneously transfer the structure–image association from canonical to reversed trials, and we would then expect different brain responses to congruent and incongruent trials. Indeed, ERPs showed a significant congruency effect over a late central dipolar component culminating around 1.7 s after word onset (Monte Carlo $P_{\mathrm{corr}} = 0.044$ in 34 infants; Fig. 4). Because the surprise effect was elicited by an auditory stimulus as in experiment 2, we additionally inspected the spatiotemporal cluster defined in experiment 2 for the auditory modality and found a significant difference {$[t(33) = 1.73, P = 0.046]$, one-tailed t test}. When both experiments were analyzed together on this cluster, the ANOVA revealed a main effect of Condition $[F(1,64) = 12.97, P < 0.001]$, a weak trend for an Experiment effect $[F(1,64) = 2.99, P = 0.09]$, and no Condition × Experiment interaction $[F(1,64) = 2.52, P = 0.12]$. Nonparametric statistics in the 66 infants of both experiments 2 and 3 (with no a priori) confirmed a common violation-of-expectations response over a central cluster of 15 electrodes (Monte Carlo $P_{\mathrm{corr}} = 0.014$). Therefore, not only did 5-mo-olds expect to see the appropriate image after having recovered the abstract word structure in canonical blocks, but they also expected to experience the appropriate structure after seeing the image in reversed blocks, without any additional training.

Discussion

This series of three experiments highlights the symbolic depth of mental representations in preverbal infants with two important findings. First, we show that well before they start talking, infants can combine a series of abstract features (here, a sequence of syllable repetitions and changes) into an operable mental variable that is available for subsequent associative operations. While classical habituation/dishabituation paradigms merely address information encoded in the infant brain as distributed implicit network states, our paradigm allows one to pinpoint abstract, redescribed representations that are available to the infant brain. Contrary to previous speculations that infants merely use labels as mnemonic cues (35), attentional grabbers (36), or essence placeholders (37), we propose here that labels drive infants into representational redescription and help them organize their inner mental space through a mapping between abstract representations and their associated labels. Importantly, these results show that language proficiency is not a prerequisite for combining different abstract features (here, the presence of a repetition and its ordinal position) into a unitary mental representation.

Second, we demonstrate that these preverbal yet sophisticated mental representations are flexibly associated with arbitrary labels in multiple modalities and, more importantly, through a powerful bidirectional mapping. Whereas nonhuman animals fail to spontaneously generalize conditioned associations to the reverse order (2, 4, 34), humans are particularly prone to disregarding temporal orderings and to inferring symmetry (23), sometimes at the expense of logical reasoning (24). It was therefore commonly acknowledged that language experience supported symmetry. Our study, which reports preverbal abilities for symmetry, questions this assumption. Together with previous studies showing that even primates with previous linguistic experience fail to infer bidirectional mappings when learning symbols using conditioned associations (4), this empirical finding suggests that symmetry might be the early marker of an inner representational system that is prone to attributing a mental variable to an external category; in other words, reducing it to a sign.

With intensive training, nonhuman animals also come to use arbitrary labels to represent objects, colors, or quantities (38, 39). Monkeys may even learn to combine abstract signs in formal operations such as addition (40). However our findings, showing that infants rapidly and spontaneously infer a powerful bidirectional mapping between a set of abstract representations and a set of labels, together with the abundant evidence that nonhuman animals have difficulties in processing bidirectional associations, suggest a clear discontinuity between human and nonhuman behavior, which might be a building block for human cognition and notably language development.

Methods

All experiments were approved by the regional committees for biomedical research [Comité de Protection des Personnes (CPP) Kremlin Bicêtre and CPP Île de France III]. Both parents were informed and provided their written consent before the experiment.

Stimuli. The triplets were generated as the concatenation of synthesized consonant/vowel (CV) syllables to conform to AAB, ABA, and ABB structures. Two different sets of 15 CV syllables were used to build two separate vocabularies of 120 distinct words. One vocabulary was made from the consonants b, t, and k and the French vowel sounds a, u, ou, in, and e; and the other vocabulary was made from consonants p, d, and g and the French vowel sounds an, eu, i, o, and on. The syllables were generated with a duration fixed at 240 ms and flat intonation using the MBROLA text-to-speech software (41) with French diphones and digitized at 22,050 Hz. The syllables were concatenated

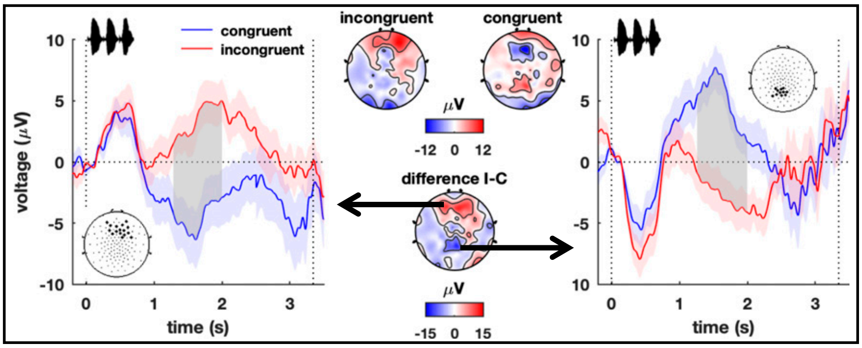

Fig. 4. Reversed trials in experiment 3. (*Left* and *Right*) Grand average responses to congruent (blue) and incongruent (red) trials recorded from the positive (*Left*) and negative (*Right*) significant clusters of electrodes presented on the graph. (*Middle*) ERP topographies for incongruent (I) and congruent (C) conditions and their difference (I−C) over the significant time window (gray area on the plots). Electrodes and time were identified using nonparametric analyses. Infants were able to generalize the structure–label pairings from the trained canonical order to the untrained reverse order, demonstrating a bidirectional mapping between the auditory structures and the arbitrary visual labels. The vertical dotted line at 3.247 s indicates the onset of the following trial.

to form trisyllabic words (duration, 720 ms), ensuring that the CV from syllable A were systematically different from CV from syllable B. As visual labels, we used two cartoon pictures (a red fish and a yellow lion); as auditory label, we used the monosyllabic word "schtroumpf" (duration, 1 s) recorded by a female native French speaker in an infant-directed speech register.

Protocol. Brain activity was recorded using a high-density EEG net (128 channels; Electrical Geodesics, Inc.) while the infants were seated on their caregivers' laps. A typical canonical trial consisted of the presentation of a word, followed 1 s later by a label (Fig. 1). In experiments 1 and 3, we used the two images as labels, each presented during 1 s over a background flickering at a specific frequency (10 or 15 Hz). In experiment 2, we used the red fish as one label and the auditory word schtroumpf as the other label. During the familiarization phase, infants were presented with a series of 36 trials in which the two labeled structures were consistently paired with their respective labels. The pairings were counterbalanced across participants.

For experiment 1, 47 healthy 5-mo-old infants were recruited to participate; 15 participants were excluded from analyses of the visual labels (final group: 14 girls and 18 boys 21 ± 2 wk old; range, 17 to 26 wk), and eight infants were excluded from analyses of the CNV after the triplets (final group: 16 girls and 23 boys 21 ± 2 wk old; range, 17 to 26 wk). During the familiarization phase, infants were exposed to AAB and ABA words followed by their visual labels. In the test phase, we used the second vocabulary and introduced the third structure, ABB, which was equally paired with the two images. Lastly, in 25% of AAB and ABA trials, the visual labels were swapped such that the associations were incongruent with the learned pairings. The two visual labels and the three structures were presented with equal frequency.

For experiment 2, 57 healthy 5-mo-old infants were recruited. The visual and auditory labels were analyzed separately. We report on 34 infants for the visual label analyses (20 girls and 14 boys 20 ± 2 wk old; range, 18 to 24 wk) and on 32 infants for the auditory label analyses (19 girls and 13 boys 20 ± 2 wk old; range, 18 to 24 wk). Experiment 2 was similar to experiment 1, except that the labels were either an image or a word, and the ABA and ABB structures were labeled, keeping AAB as the third structure equally paired with each label. In this experiment, both vocabularies were used during testing, and we set aside the generalization vocabulary for a subset of congruent, incongruent, and neutral trials used for the statistical analyses.

For experiment 3, 44 healthy 5-mo-old infants were recruited to participate; 10 participants were excluded from subsequent analyses (final group: 11 girls and 23 boys 20 ± 1 wk old; range, 18 to 23 wk). The familiarization phase was identical to that in experiment 1. During testing, small blocks of six reversed trials were interspersed between larger blocks of 14 canonical trials that were always congruent. In reversed trials, the image was presented first during 1 s, followed 1 s later by the trisyllabic word drawn from the generalization vocabulary. Half of these reversed trials were incongruent with the learned associations so that infants could not learn any label–structure associations from the reversed trials.

Data Processing. The EEG was continuously digitized at 250 Hz from 128 scalp electrodes referenced to the vertex. For each channel, we rejected epochs with fast average amplitude exceeding 250 µV or when deviations between fast and slow running averages exceeded 150 µV. Participants with fewer than 10 trials in one of the experimental conditions were rejected from subsequent analyses. The remaining trials were averaged, locked to stimulus onset (either labels or triplets, depending on the analysis), digitally transformed to an average reference, band-pass filtered (0.2 to 15 Hz for ERP analyses, 0.2 to 40 Hz for frequency-tagging analyses), and corrected for baseline either over a 200-ms window before stimulus onset (similar results were observed with the raw, unfiltered data) or over the first 200 ms after trisyllabic word onset in experiment 3, where we could capitalize on the signal alignment, which was easier with the sharp auditory response.

Statistical Analyses. Based on the literature, we inspected late surprise effects to incongruent labels and early priming effects amplifying the early sensory responses to the congruent labels (16, 17, 29). While some brain responses are well characterized in time and space, the understanding and description of high-level ERP components in infants are still incomplete and imprecise. Accordingly, we analyzed well-described brain responses (e.g., CNV or visual entrainment) based on specific clusters of electrodes and time windows determined from the literature. However, for late responses whose topographies and latencies are not well described and depend on various experimental parameters, we adopted a data-driven approach. To control for the risk of false positives due to the numerous measures, we used nonparametric statistics, combining a clustering and randomization procedure (42).

Late surprise effects. The late responses were studied by comparing congruent and incongruent trials using nonparametric statistics with an alpha threshold set to 0.1, a minimal cluster size of three electrodes, and 5,000 permutations over a 1- to 1.8-s time window. This procedure allowed for the detection of positive, negative, and dipolar components, with a stringent control on false positives (31). We tested whether the identified response to incongruent visual labels was replicated from experiment 1 to experiment 2 by running an ANOVA, with Condition as a within-subjects factor and Experiment as a between-subjects factor over the spatiotemporal cluster identified in experiment 1. We additionally tested whether the late response to the incongruent triplet structure in experiment 3 could be considered a replication of the response to the incongruent auditory label in experiment 2. Voltages averaged over the spatiotemporal cluster identified in experiment 2 were submitted to an ANOVA, with Condition as a within-subjects factor and Experiment as a between-subjects factor. In experiments 1 and 2, given that the proportion of congruent and incongruent trials was strongly imbalanced (75% vs. 25%, respectively), a subset of congruent trials was defined to match the number of incongruent trials. In experiment 1, for each incongruent trial, we selected the closest congruent trial, controlling for the identity of the image (e.g., for each fish-incongruent trial, we selected the closest fish-congruent trial; after this procedure, we report on 22 artifact-free trials per subject in the congruent/incongruent conditions on average). In experiment 2, the subset was defined upstream with the use of a new vocabulary (on average, 15 and 16 artifact-free trials per subject in the congruent and incongruent visual conditions, respectively; 16 and 16 artifact-free trials per subject in the congruent and incongruent auditory conditions, respectively). In experiment 3, the proportion of incongruent and congruent trials was perfectly balanced (on average, 16 and 15 artifact-free trials per subject in the congruent and incongruent trials, respectively).

Early priming effects. The analysis and results of early sensory ERP components is described in *SI Appendix*. Visual labels were additionally presented over a flickering background to pinpoint early visual activity in a robust and predictive way. The rhythmic stimulation elicited an entrained occipital response during label presentation, which we could easily inspect for congruency modulations. Entrainment was measured over an occipital cluster of electrodes (20 electrodes encompassing electrodes Oz, T5, and T6) during picture presentation, delayed by 200 ms to compensate for the conduction time of the stimulation to the visual cortices (i.e., 200 to 1,200 ms after stimulus onset). For each subject, each electrode, and each trial, we computed the power spectral density (PSD) of the signal over this time window using the fast Fourier transform algorithm as implemented in MATLAB and averaged the values across trials for each condition of interest separately for each stimulation frequency (10 and 15 Hz). Next, we fitted each average PSD with a power-law function (i.e., in the log space, an affine function using MATLAB polyfit) and then measured the peak signal-to-noise ratio (PSNR) as the ratio between the deviation of the signal power from power law at the tagged frequency and the root-mean-square deviation from the power law over neighboring frequencies, in decibels. PSNR was averaged over the cluster in each subject and submitted to an ANOVA, with Condition (congruent vs. incongruent) and Stimulation Frequency (10 vs. 15 Hz) as within-subject factors.

CNV. Because the CNV typically develops between the offset of a cueing stimulus and the onset of the following target stimulus over central recording sites (29), we inspected brain responses during the silence period before label onset in experiment 1 and defined a central cluster of electrodes (32 electrodes around Cz) based on the main response with all trials averaged together. Brain activity was averaged over this cluster for each condition, between 1 and 1.5 s after word onset and was submitted to a three-way ANOVA, with Structure (AAB, ABA, or ABB) as a within-subject factor. On average, we obtained 42 artifact-free trials per subject and per condition (41, 42, and 42 for AAB, ABA, and ABB, respectively).

ACKNOWLEDGMENTS. We thank all the infants and their parents who participated in this study, as well as Antoinette Jobert, Giovanna Santoro, the medical team of Unité de Recherche en NeuroImagerie Applicative Clinique et Translationnelle at NeuroSpin, and the Laboratoire de Sciences Cognitives et Psycholinguistique BabyLab in Paris, who helped in carrying out the experiments. This research was supported by the Fondation Bettencourt Schueller and the European Research Council Advanced Grant Babylearn.

1. Nazzi T, Bertoncini J (2003) Before and after the vocabulary spurt: Two modes of word acquisition? *Dev Sci* 6:136–142.
2. Medam T, Marzouki Y, Montant M, Fagot J (2016) Categorization does not promote symmetry in Guinea baboons (*Papio papio*). *Anim Cogn* 19:987–998.
3. Srihasam K, Mandeville JB, Morocz IA, Sullivan KJ, Livingstone MS (2012) Behavioral and anatomical consequences of early versus late symbol training in macaques. *Neuron* 73:608–619.
4. Kojima T (1984) Generalization between productive use and receptive discrimination of names in an artificial visual language by a chimpanzee. *Int J Primatol* 5:161–182.
5. Nieder A (2009) Prefrontal cortex and the evolution of symbolic reference. *Curr Opin Neurobiol* 19:99–108.
6. Wang L, Uhrig L, Jarraya B, Dehaene S (2015) Representation of numerical and sequential patterns in macaque and human brains. *Curr Biol* 25:1966–1974.
7. Tincoff R, Jusczyk PW (2012) Six-month-olds comprehend words that refer to parts of the body. *Infancy* 17:432–444.
8. Tincoff R, Jusczyk PW (1999) Some beginnings of word comprehension in 6-month-olds. *Psychol Sci* 10:172–175.
9. Bergelson E, Swingley D (2012) At 6-9 months, human infants know the meanings of many common nouns. *Proc Natl Acad Sci USA* 109:3253–3258.
10. Dehaene-Lambertz G, Spelke ES (2015) The infancy of the human brain. *Neuron* 88: 93–109.
11. Dehaene-Lambertz G, et al. (2006) Functional organization of perisylvian activation during presentation of sentences in preverbal infants. *Proc Natl Acad Sci USA* 103: 14240–14245.
12. Dehaene-Lambertz G, Dehaene S, Hertz-Pannier L (2002) Functional neuroimaging of speech perception in infants. *Science* 298:2013–2015.
13. Waxman SR, Braun I (2005) Consistent (but not variable) names as invitations to form object categories: New evidence from 12-month-old infants. *Cognition* 95:B59–B68.
14. Ferry AL, Hespos SJ, Waxman SR (2010) Categorization in 3- and 4-month-old infants: An advantage of words over tones. *Child Dev* 81:472–479.
15. Yeung HH, Werker JF (2009) Learning words' sounds before learning how words sound: 9-month-olds use distinct objects as cues to categorize speech information. *Cognition* 113:234–243.
16. Xu F, Cote M, Baker A (2005) Labeling guides object individuation in 12-month-old infants. *Psychol Sci* 16:372–377.
17. Marcus GF, Vijayan S, Bandi Rao S, Vishton PM (1999) Rule learning by seven-month-old infants. *Science* 283:77–80.
18. Marcus GF, Fernandes KJ, Johnson SP (2007) Infant rule learning facilitated by speech. *Psychol Sci* 18:387–391.
19. Frank MC, Slemmer JA, Marcus GF, Johnson SP (2009) Information from multiple modalities helps 5-month-olds learn abstract rules. *Dev Sci* 12:504–509.
20. Gervain J, Macagno F, Cogoi S, Peña M, Mehler J (2008) The neonate brain detects speech structure. *Proc Natl Acad Sci USA* 105:14222–14227.
21. Giurfa M, Zhang S, Jenett A, Menzel R, Srinivasan MV (2001) The concepts of 'sameness' and 'difference' in an insect. *Nature* 410:930–933.
22. Martinho A, 3rd, Kacelnik A (2016) Ducklings imprint on the relational concept of "same or different". *Science* 353:286–288.
23. Sidman M, et al. (1982) A search for symmetry in the conditional discriminations of rhesus monkeys, baboons, and children. *J Exp Anal Behav* 37:23–44.
24. Ogawa A, Yamazaki Y, Ueno K, Cheng K, Iriki A (2010) Neural correlates of species-typical illogical cognitive bias in human inference. *J Cogn Neurosci* 22:2120–2130.
25. Friedrich M, Friederici AD (2011) Word learning in 6-month-olds: Fast encoding-weak retention. *J Cogn Neurosci* 23:3228–3240.
26. Friedrich M, Wilhelm I, Mölle M, Born J, Friederici AD (2017) The sleeping infant brain anticipates development. *Curr Biol* 27:2374–2380.e3.
27. Kouider S, et al. (2015) Neural dynamics of prediction and surprise in infants. *Nat Commun* 6:8537.
28. Emberson LL, Richards JE, Aslin RN (2015) Top-down modulation in the infant brain: Learning-induced expectations rapidly affect the sensory cortex at 6 months. *Proc Natl Acad Sci USA* 112:9585–9590.
29. Norcia AM, Appelbaum LG, Ales JM, Cottereau BR, Rossion B (2015) The steady-state visual evoked potential in vision research: A review. *J Vis*, 10.1167/15.6.4.
30. Walter WG, Cooper R, Aldridge VJ, McCallum WC, Winter AL (1964) Contingent negative variation : An electric sign of sensori-motor association and expectancy in the human brain. *Nature* 203:380–384.
31. Mento G, Valenza E (2016) Spatiotemporal neurodynamics of automatic temporal expectancy in 9-month old infants. *Sci Rep* 6:36525.
32. Karmiloff-Smith A (1995) *Beyond Modularity: A Developmental Perspective on Cognitive Science* (MIT Press, Cambridge, MA).
33. Namy LL (2001) What's in a name when it isn't a word? 17-Month-olds' mapping of nonverbal symbols to object categories. *Infancy* 2:73–86.
34. Lionello-DeNolf KM (2009) The search for symmetry: 25 years in review. *Learn Behav* 37:188–203.
35. Needham A, Baillargeon R (2000) Infants' use of featural and experiential information in segregating and individuating objects: A reply to Xu, Carey and Welch (2000). *Cognition* 74:255–284.
36. Sloutsky VM (2003) The role of similarity in the development of categorization. *Trends Cogn Sci* 7:246–251.
37. Xu F (2002) The role of language in acquiring object kind concepts in infancy. *Cognition* 85:223–250.
38. Matsuzawa T (1985) Use of numbers by a chimpanzee. *Nature* 315:57–59.
39. Pepperberg IM (2009) *The Alex Studies: Cognitive and Communicative Abilities of Grey Parrots* (Harvard Univ Press, Cambridge, MA).
40. Livingstone MS, Srihasam K, Morocz IA (2010) The benefit of symbols: Monkeys show linear, human-like, accuracy when using symbols to represent scalar value. *Anim Cogn* 13:711–719.
41. Dutoit T (1997) An *I*ntroduction to *T*ext-to-*Speech Synthesis*, Text, Speech and Language Technology, eds Ide N, Véronis J (Springer Science+Business Media, Dordrecht, The Netherlands), Vol 3.
42. Maris E, Oostenveld R (2007) Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Methods* 164:177–190.