



Rapid infant learning of syntactic–semantic links

Monica Barbir^{a,b,1} , Mireille J. Babineau^{a,c} , Anne-Caroline Fiévet^a , and Anne Christophe^a 

Edited by Barbara Landau, Johns Hopkins University, Baltimore, MD; received May 28, 2022; accepted November 7, 2022

In the second year of life, infants begin to rapidly acquire the lexicon of their native language. A key learning mechanism underlying this acceleration is syntactic bootstrapping: the use of hidden cues in grammar to facilitate vocabulary learning. How infants forge the syntactic–semantic links that underlie this mechanism, however, remains speculative. A hurdle for theories is identifying computationally light strategies that have high precision within the complexity of the linguistic signal. Here, we presented 20-mo-old infants with novel grammatical elements in a complex natural language environment and measured their resultant vocabulary expansion. We found that infants can learn and exploit a natural language syntactic–semantic link in less than 30 min. The rapid speed of acquisition of a new syntactic bootstrap indicates that even emergent syntactic–semantic links can accelerate language learning. The results suggest that infants employ a cognitive network of efficient learning strategies to self-supervise language development.

language acquisition | cognitive development | infant cognition | word learning | grammar learning

A fundamental feature of language is that word use is governed by rules, grammatical and morphological, among others. Infants are sensitive to the grammatical regularities of their native language early in life (1, 2). In the second year of life, we observe overt knowledge of syntactic–semantic links, when rule violations begin to elicit surprise (3). By 18 mo, infants actively use acquired links to self-supervise language learning (4). They use grammatical contexts to narrow the scope of possible word meanings and learn new elements of the lexicon, through a mechanism called ‘syntactic bootstrapping’ (5–7). How syntactic bootstraps are acquired remains unclear, in part because of the lack of direct observation of syntactic–semantic link acquisition in a natural language environment.

One experimental approach to syntactic–semantic link acquisition is ‘milestone charting’. Milestone charting identifies the developmental endpoints for each native language syntactic–semantic rule (e.g., refs. 4 and 8), demarcating when a capacity emerges. Milestones indicate a developmental trend from associations between elements (e.g., refs. 3 and 9) to bootstrapping some months later (e.g., refs. 4 and 10). They also highlight graded developmental milestones for each link. Category-level links are acquired before subcategory links: infants use syntactic–semantic links to infer whether a novel word is a noun or verb from 18 mo of age, but only start using them to determine whether a verb is transitive or intransitive from 25 mo of age (4, 11, 12).

A second approach is ‘toy grammar learning’. Toy grammar learning measures the acquisition of a simplified, artificial grammar targeting specific input features (e.g., refs. 13 and 14), identifying calculations infants perform. Toy grammars show that infants are rule learners, who extract and generalize abstract structures quickly and efficiently (e.g., refs. 13 and 15). Toy grammars mainly probe grammar learning. When they include semantics, the complexity of syntactic–semantic links is reduced to a key feature, such as word frequency (e.g., refs. 14 and 16). One toy grammar study, for example, demonstrated that infants can learn cross-element links, connecting phonological, distributional, and semantic cues (16). Toy grammars however face a scale-up challenge vis-à-vis natural languages. Infants’ internal mental representations of toy grammar may differ from that of natural language grammar; and the cognitive mechanisms involved in learning may be different for simplified artificial environments and noisy naturalistic ones. Simulating the multifaceted complexity of syntactic–semantic links in a toy grammar model is challenging (e.g., ref. 17). Milestone and toy grammar methods only provide indirect information on the cognitive mechanisms governing learning, by predicting but not describing learning trajectories.

In this preregistered study (DOI: [10.17605/OSF.IO/ZGBP8](https://doi.org/10.17605/OSF.IO/ZGBP8), [10.17605/OSF.IO/X8H3A](https://doi.org/10.17605/OSF.IO/X8H3A); (18)), we investigated whether infants can quickly harness novel grammatical elements in a native language environment to functionally expand their vocabulary (19). Infant acquisition

Significance

Infants start learning words at an incredible pace in their second year of life. One of the strategies they use to learn words so efficiently is to take advantage of clues hidden in grammar: ‘syntactic bootstrapping’. How infants with fledgling lexicons learn complex relationships between words and grammar is unknown. Using eye tracking, we demonstrate that 1 to 2-y-old infants can quickly learn a novel relationship between words and grammar from short videos and use it to learn new words. These results show that young language learners exploit links between language elements on the fly, suggesting that infants self-supervise learning through a network of efficient language-learning shortcuts.

Author contributions: M.Barbir, M.J.Babineau, and A.C. designed research; M.Barbir performed research; M.Barbir analyzed data; A.-C.F. contributed to provision of infant sample; recruitment of participants; and M.Barbir wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Copyright © 2022 the Author(s). Published by PNAS. This article is distributed under [Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0](https://creativecommons.org/licenses/by-nc-nd/4.0/) (CC BY-NC-ND).

¹To whom correspondence may be addressed. Email: monica.barbir@ircn.jp.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2209153119/-/DCSupplemental>.

Published December 27, 2022.

*We focus here on rules that must be acquired from the signal (language-specific rules), such as the rule that determiners precede nouns in English (e.g., the cat) but not in Japanese (e.g., *o neko* [ø cat]).

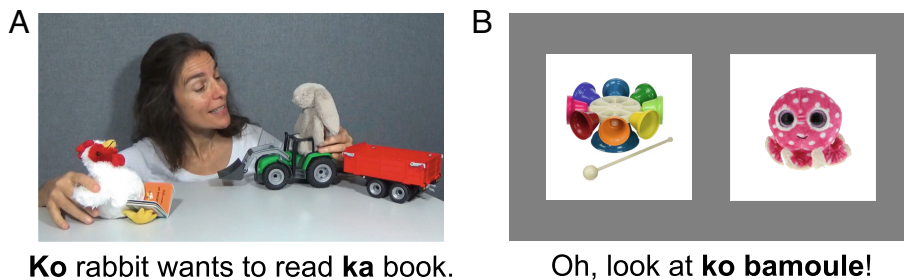


Fig. 1. Schematic of study design. (A) Infants first watched a training video with novel determiners (*ko* and *ka*), three times at home and once in the lab. (B) Then, they completed a test phase, during which they saw two novel toys (one animate and one inanimate) and heard a prompting sentence to look at one of the two toys. The prompting sentence contained a pseudonoun infants did not know (e.g., *bamoule*). The only way to solve the task was to use knowledge about the syntactic–semantic links associated with the animate and inanimate determiners.

of novel syntactic–semantic links in a complex ecological environment remains untested (see ref. (20) for an example with preschool children). Indirect evidence from native language milestones suggests that bootstrapping occurs months after syntactic–semantic links are formed; while evidence from simplified toy grammars suggests that bootstrapping can occur within minutes. If bootstrapping is used soon after learning a syntactic–semantic link, syntactic bootstrapping would be a more efficacious language-learning mechanism than previously thought.

We presented 20-mo-old infants with two novel subcategory-level determiners, ‘*ko*’ for animate objects and ‘*ka*’ for inanimate objects, inserted directly into their native language, French, as a replacement for existing determiners (‘*le*’, ‘*la*’, ‘*un*’, ‘*une*’: ‘the’ and ‘a’; *SI Appendix*, Fig. S1). The determiners represented a novel morphological distinction in an existing syntactic structure. This method ensured that infants would interpret the novel determiners as grammatical elements and not as content words (i.e., meaning-bearing words). The novel determiners functioned structurally as do French determiners; for instance, an adjective modifying a noun was inserted in the same place as in French, between the determiner and the noun (e.g., *le grand chien* = *ko grand chien*, **the big dog**).

The stimuli were selected to ensure alignment with existing evidence, ecological validity, and cross-language generality. The categorical distinction, animate–inanimate, is salient for infants (21–23) and grammatically common among the world’s languages (for review, ref. (24)), but is not a grammatical distinction in the native language (French) of our infant cohort. The contexts into which the grammatical elements were inserted were created to align with theories of syntactic–semantic link acquisition, while providing infants with the range of cues present in natural, everyday input.

A prominent idea in the literature is that syntactic–semantic links are inferred from co-occurring word pairings, from the moment the infant knows a handful of words (25). There is evidence that semi-supervised computational models tracking the co-occurrence of statistics between grammatical elements and just a few known words are capable of high-precision inference (26–28), see ref. (29) for an example with morphological elements). In our experimental design, infants could track the co-occurrences between the novel determiners and 12 common nouns – six animate and six inanimate (e.g., ‘*ka tractor*’, ‘*ka book*’, ‘*ko rabbit*’, ‘*ko chicken*’; Fig. 1A) – embedded in natural speech.

Infants could also track novel determiner distributions in broader sentence frames (e.g., ‘*Ko rabbit reads ka book*’). Animacy had a direct influence on the contextual distributions of the determiners. Animates and inanimates appeared on average in different sentence frames: animates were more likely to be the agents of actions and the subjects of a sentence, while inanimates were more

likely to be the receivers of actions and the objects of a sentence. The information present in word-pair co-occurrence and sentence frames reflected the multidimensional linguistic cues to syntactic–semantic links present in natural language.

In addition to linguistic cues, our design incorporates co-occurring perceptual and social cues, including gaze direction (30) and character/object agency (31). Vocabulary acquisition is thought to require a three-pronged mechanistic approach – perceptual, social, and linguistic (32) – and we extend this mechanistic framework to infant grammar acquisition. We postulate hitherto unconsidered cues that infants use beyond grammar-noun co-occurrences. We thus embedded novel determiners in a rich, naturalistic environment, so infants will have access to the wide array of cues, as well as noise, that they do on a daily basis.

The novel determiners were presented to infants in an ecologically valid workflow: infants watched a short video in which a woman acts out stories with toys (Fig. 1A). Infants watched the training video at home for three consecutive days before the laboratory visit, and once more at the laboratory. The total exposure time to the novel determiners was approximately 30 min, and the total period of use spanned 4 d. The extended exposure allowed for both a pragmatically sound introduction of the novel determiners and a prolonged timespan to process, or simply be familiarized with, the use of these linguistic elements (see *SI Appendix*, *SI Text* for preliminary study results).

After watching the training video at home, infants watched it again in the laboratory and proceeded to the test phase. During the test phase, infants were presented with novel toys and corresponding novel nouns. They saw two images on the screen, a novel animate toy and a novel inanimate toy, and heard a prompting sentence to look at one of the two images (e.g., Oh look at *ko bamoule!* Fig. 1B). The novel toys were unfamiliar to infants so they had no French label for them. Novel nouns were paired with a determiner from the training phrase. If infants inferred the syntactic–semantic links during training (*ko* + animate and *ka* + inanimate), they could use them to constrain and bootstrap the potential meaning of the novel nouns (e.g., *ko* + *bamoule*, *bamoule* = animate).

Results

Infants’ gaze was recorded with an eye-tracker to index their interpretation of the novel nouns. Gaze was analyzed via two statistical measures: the fine-grained ‘looking-while-listening’ measure that encodes the real-time evolution of gaze patterns during the trial, and the broad ‘preferential-looking’ measure that sums overall gaze time to a target during the trial (33). These two methods allowed for a complementary and comprehensive analysis of both specific information on when infants orient their gaze to the correct image, and their general preference in light of individual variation

(e.g., some infants look immediately to the target image, while others only do so toward the end of the trial).

Specifically, we analyzed the proportion of gaze time to the animate image, in trials where infants had heard a prompting sentence with *ko* vs. one with *ka* (within-subjects; Fig. 1*B*). If infants had inferred the syntactic–semantic links and were able to use them to narrow down the meanings of words they did not know, they should show distinct gaze patterns for novel nouns paired with *ko* and for those paired with *ka*. A cluster-based permutation analysis (34) revealed that infants looked more to the animate image on hearing *ko* than *ka* during the time-window from 300 to 2,440 ms after hearing the novel

determiners ($P = 0.01$, Fig. 2*A*). A mixed-effects regression analysis showed that infants looked longer overall, throughout the whole trial, to the animate image when they heard *ko* compared with *ka*, but this effect was moderate: $\beta = -0.09$, $SE = 0.04$; model comparison: $\chi^2(1) = 4.12$, $P = 0.04$; Cohen's $d = 0.53$ (means are displayed in Fig. 2*B*). The data thus suggest a robust early time-locked effect, followed by variability later in the trial. These results show that after brief, natural language exposure to two formal linguistic elements, infants were able to rapidly infer syntactic–semantic links and use this knowledge to constrain the set of potential meanings of unknown lexical words.

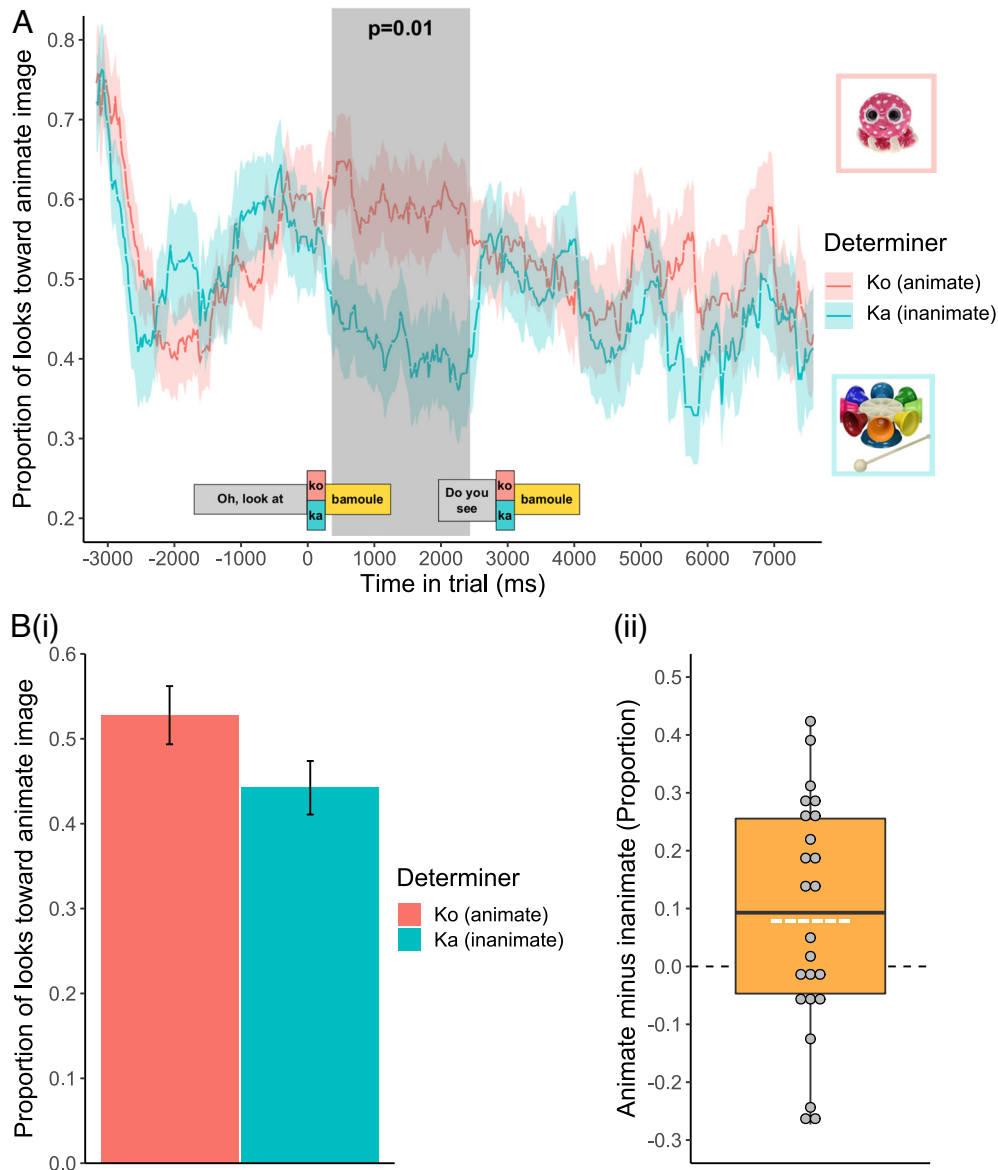


Fig. 2. Results. (A) Time-course. Proportion of looks toward the animate image at each point in time during the trial, when the infants ($n = 24$) heard the animate determiner *ko* paired with a novel noun (e.g., *ko bamoule*) in pink and when they heard the inanimate determiner *ka* (e.g., *ka pirdale*) in blue. Dark lines represent mean across participants, and light shading the 95% confidence intervals of the mean. Gray shading indicates the time-window during which the two conditions, animate determiner sentences vs. inanimate determiner sentences, diverge (300 to 2,440 ms, $P = 0.01$, cluster-based permutation analysis). (B) Overall looking preference. Proportion of looks toward animate image averaged over the whole trial window. Infants ($n = 24$) look significantly longer to the animate image when they heard the determiner *ko* than *ka* ($P = 0.04$, mixed-effects regression analysis). (i) Looking time means and SE. Mean proportion of looks toward the animate image, when the infants heard the animate determiner *ko* paired with a novel noun in pink and when then heard the inanimate determiner *ka* in blue. Error bars represent SEM. (ii) Difference between conditions per participant. The difference between the proportion of looks toward the animate image when the animate or inanimate determiner was heard, per participant. Dots indicate participants. Dashed white line indicates mean. Upper and lower regions of the box indicate the first and third quartiles (25th to 75th percentiles). The upper whisker represents the third quartile up to the 1.5 interquartile largest value, while the lower whisker the 1.5 interquartile smallest value to the first quartile. Dashed black line indicates no difference between proportion of looks the animate image, when infants heard the animate or inanimate determiner.

Discussion

Our study demonstrates the remarkable speed with which infants acquire a novel syntactic–semantic link and use it to self-supervise vocabulary learning in their native language. In a complex linguistic environment, short exposure was sufficient for 20-mo-old infants to learn a new bootstrap. The ecologically valid testing method we used may leverage statistical and inferential relationships and fluidities between the basic components of language, like grammar and vocabulary. It evidences the mutual informativity between naturalistic and toy grammar protocols, akin to *in vivo* and *in vitro* studies in biology. These findings constrain the cognitive network that underlies the development of a syntactic–semantic matrix as a language-learning mechanism. They suggest that syntactic bootstrapping is a fundamental language-learning strategy, with only a brief naturalistic signal needed to construct each bootstrap.

The efficiency of bootstrap acquisition is likely based on a network of cognitive learning mechanisms. We assert that one such mechanism is the age-dependent deployment of information-heavy attention. Prior studies demonstrate that learners direct attention to highly informative cues, and that the relationship between attention and informativity can be formalized mechanistically (e.g., the competition model, (35)). When a child learns that certain linguistic elements, like grammar, provide rich information connected to other elements, their focused attention on those relationships, rather than a uniform allocation of cognitive resources, would enhance the speed of learning (36), see ref. (37) for parallel with attention to reliable cues). Morphological cues may be particularly informative for learners because attention to one location in the syntactic structure can provide multiple features for bootstrapping. Thus, the time to learn a syntactic bootstrap may depend on its information-heaviness.

A second mechanism we propose is infants' use of flexible hypotheses, where instead of learning rigid rules, infants use hypotheses and quickly revise them in light of divergent evidence. Prior studies on concept acquisition demonstrate that infants rapidly update concept scope in accordance with novel evidence (38–40). Current theories of language acquisition assume a one-way transition from not knowing to knowing, but the present study is designed to detect dynamic unstable transitions. For example, infants may stop using a familiar syntactic–semantic link to figure out plant classification upon hearing 'ko tree'. In this 'learning to learn' framework, the mere time to learn a linguistic element may be a coarse measurement. Direct observation of the dynamic statistical structure of the early grammar baseline and how it coevolves with vocabulary sheds light on the cognitive links that underlie acquisition of the complex network of language components.

Materials and Methods

This experiment is a preregistered study. All sample-sizes, exclusion criteria, materials, procedure, and analyses are as preregistered, unless otherwise stated. All materials, code, data, and analyses, as well as the preregistration, are available on the study's OSF page (DOI: [10.17605/OSF.IO/X8H3A](https://doi.org/10.17605/OSF.IO/X8H3A)) https://osf.io/zgpb8/?view_only=cd943c5166d0414d9a319c1b305c9730.

Participants. Infants were recruited from the lab database (voluntary response sample). All were monolingual French-learning infants, who heard less than 10% of another language. Infants were excluded because of attrition ($n = 2$) or intentionally in line with preregistered standards for not adhering to the experimental protocol ($n = 4$), not having at least two trials per condition with at least 50% of gaze to the screen ($n = 8$), excessive fussiness ($n = 2$), and technical error

($n = 3$). The remaining 24 infants were included in the analyses (mean: 19.25 mo; range: 19.15 to 20.11 mo; 16 girls, 8 boys). Sample size was chosen based on similar experiments with 20-mo-old infants (41) and developmental research sample size standards (42). Sample size was part of a three-pillar approach to statistical power: it was coupled with methodological choices (i.e., number of trials and trial length) that have been shown to increase measurement reliability (43, 44). Written informed consent was obtained from each child's parents prior to the experiment. All research was approved by the local ethical board: CER Paris Descartes 20140100001072.

Materials.

Novel determiners. The novel determiners *ko* (/ko/) and *ka* (/ka/) were created such that they were phonotactically possible in French and that they resembled in form existent singular determiners, which are monosyllabic and have roughly similar phonological forms for masculine and feminine variants. In French, determiners are marked for grammatical gender in the singular form:

Definite masculine: *le* /lə/

Definite feminine: *la* /la/

Indefinite masculine: *un* /ɛ̃/

Indefinite feminine: *une* /yn/

Training video. The training video consisted of live-action scenes in which a woman acted out stories with stuffed animals and toy objects, using child-directed speech (the script and video are available on the study's OSF page). The stories were entirely in French, except for the novel determiners.

The novel determiners were each presented 30 times during the training video ($ko \times 30$, $ka \times 30$). They were paired with six distinct animal nouns (*lapin* rabbit, *poule* chicken, *cochon* pig, *chien* dog, *chat* cat, and *souris* mouse) or object nouns (*livre* book, *tracteur* tractor, *biberon* bottle, *poussette* stroller, *voiture* car, *chaussure* shoe) that French-learning infants are likely to know by 20 mo of age based on previous French MacArthur–Bates Communicative Development Inventory (CDI; (45)) data gathered in our laboratory. All nouns began with a consonant so as to avoid cliticization that occurs with vowel-initial nouns (*le + avion = /lavion/*). The nouns were chosen such that half of the animal nouns were masculine and half feminine, and the same for object nouns. As such, the novel determiners could not also be marking grammatical gender.

The novel determiners functioned in the same way structurally as existent determiners. For example, adjectives in French appear most often between the determiner and noun (e.g., *le joli chat*, the cute cat). This was thus replicated with the novel determiners: *ko joli chat*. Infants heard each determiner paired with each noun six times (e.g., *ka chaussure*), and one of the six times the pairing involved an adjective (e.g., *ka grande chaussure*). The adjective pairing also served to facilitate segmentation of the determiner–noun sequence, such that it would not be perceived as one new noun (e.g., *kachaussure*) or a proper noun.

To aid categorization, scenes were constructed to involve interactions between one animate and one inanimate (highlighting dissimilarity) or between two animates/two inanimates (highlighting similarity). In line with ambient distribution, animates were more likely to be agents or subjects of a sentence, while inanimates were more likely to be patients or objects of a sentence.

The training video included test item familiarization scenes, in which the woman telling the stories played with the test toys (two novel animal toys and two novel object toys) without naming them (e.g., "Look. This is my new toy. It has many colors."). These scenes contained a broad range of animacy cues (e.g., physical causality such self-propelled vs. caused motion and psychological causality such as goal-directed vs. without aim action, (22)). More generally, the scenes gave infants time to explore the toys visually before the test phase.

One scene from the video was refilmed using the real French determiners, and an acoustic analysis was performed comparing real and novel determiners, as well as on the first syllables of the following word (paired Student's *t* test). There were no significant differences in pitch (real determiner $M = 266.84$, $SD = 59.04$ vs. novel determiner $M = 297.64$, $SD = 69.2$; syllable following real determiner $M = 304.31$, $SD = 59.63$ vs. syllable following novel determiner $M = 290.04$, $SD = 61.03$) or length (real determiner $M = 0.16$, $SD = 0.05$ vs. novel determiner $M = 0.14$, $SD = 0.04$; syllable following real determiner $M = 0.26$, $SD = 0.09$ vs. syllable following novel determiner $M = 0.27$, $SD = 0.14$) of the determiners or first syllables (determiner pitch: $t(37) = -1.5$, $P = 0.14$, determiner length:

$t(37) = 1.59, P = 0.12$; first syllable pitch: $t(37) = 0.74, P = 0.46$; first syllable length: $t(37) = -0.28, P = 0.78$; *SI Appendix, Fig. S1*).

The video lasted for 6'49".

Novel nouns and items. During the test phase, novel determiners were presented with one of four novel nouns (*bamoule* /bamul/, *pardale* /pɛɪdal/, *doripe* /dɔɪp/, and *bradole* /bradɔl/). Novel nouns were created such that they are phonotactically possible in French. Each novel noun was paired with one novel original item, an animal or an object. Novel animals were a pink stuffed animal with a big head and many short feet, and a mouse-like animal with rabbit ears and an anteater's trunk; while novel items were a round colorful xylophone-like musical toy and a standing top. These novel nouns and novel items have been used in previous studies investigating vocabulary acquisition, and 20-mo-olds have been successful at learning the item-noun pairings (e.g., ref. (41)). Four pairings between the novel nouns and the items were constructed using a Latin-square design, so as to control for item effects. An equal number of children were assigned to each kind of pairing. Each novel noun appeared thus with *ko* for half the infants, and with *ka* for the other half.

A paired Student's *t* test was run on the acoustics of the test prompting sentences to ensure that there were no significant differences between the two determiners or between the novel nouns. The test revealed that there were no significant differences in pitch (animate determiner $M = 342.48, SD = 61.32$; inanimate determiner $M = 334.17, SD = 64.54$; $t(14) = -0.26, P = 0.8$), length (animate determiner $M = 0.13, SD = 0.01$; inanimate determiner $M = 0.13, SD = 0.03$; $t(14) = 0.46, P = 0.65$), or intensity (animate determiner $M = 69.9, SD = 3.68$; inanimate determiner $M = 70.04, SD = 3.57$; $t(14) = 0.08, P = 0.93$) of the two novel determiners; nor was there a significant difference in pitch (following animate determiner $M = 260.06, SD = 24.69$; following inanimate determiner $M = 270.42, SD = 11.09$; $t(14) = 1.08, P = 0.3$), length (following animate determiner $M = 0.22, SD = 0.03$; following inanimate determiner $M = 0.2, SD = 0.04$; $t(14) = -0.97, P = 0.35$), or intensity (following animate determiner $M = 68.07, SD = 2.14$; following inanimate determiner $M = 67.14, SD = 2.87$; $t(14) = -0.73, P = 0.48$) of the first syllable of the novel nouns.

Test items. To familiarize infants with the testing procedure, just prior to the test phase, infants saw two training trials, with words and the corresponding stuffed animals or toy objects seen during the training video (e.g., Infants saw an image of the rabbit and tractor from the video and heard *Oh regarde ko lapin ! Oh look at ko rabbit!*). During the test phase, there were test trials targeting novel nouns (with one novel animate image and one novel inanimate image) and filler trials targeting French nouns (with one familiar animate image and one familiar inanimate image). Each of the novel nouns was tested twice for a total of eight test trials. The number of test trials was maximized to domain standards, while balancing infant attentional constraints, to best capture each participant's true performance (43).

Eight filler trials were interspersed during the test phase so that an infant would not see more than two test trials in a row. The order of trials was pseudorandomized for each child. There were two kinds of filler trials, four of each: *seen* filler trials were nouns that were present in the training video, paired with a different visual exemplar of that noun (a different *souris* mouse, *cochon* pig, *biberon* bottle, *chaussure* shoe, from those seen in the training video); *known* filler trials were nouns that were not present during the training video, but likely to be known to 20-mo-old infants based on previous French CDI data gathered in our laboratory (*poisson* fish, *cheval* horse, *vélo* bike, *chapeau* hat). A cluster-based permutation analysis (34) revealed that infants correctly identified the target word in filler trials [920 to 2,840 ms ($P < 0.001$) and 3,200 to 4,680 ms ($P = 0.002$)]. If infants knew a filler trial noun, they could keep learning about the determiners' use during the test phase.

Procedure. Before coming to the lab, infants watched the training video at home, once a day, for three consecutive days. For example, if the test session was on a Saturday, infants would watch the video at home on Wednesday, Thursday, and Friday. A 4-d training session was chosen to align with the distribution of determiners in natural languages and in response to pilot study results (*SI Appendix, Text S1*). It is possible that infants would have succeeded in the task with a shorter training phase. Parents received instructions to be as neutral and quiet as possible during the screenings and not to refer to the video after the screening (exact instructions are available on the study's OSF page).

The day before coming to the lab, parents filled out a short online questionnaire marking whether the infant had seen all three videos from beginning to end. Parents also filled out the French version of the MacArthur-Bates CDI (45).

Because parents had a lot to do before coming to the lab, they were not obliged to fill out the CDI. 17/24 parents filled it out. Receptive vocabulary ranged from 132 to 459 words (mean = 305 words). Mean looking time to the novel noun target at test and vocabulary size were not significantly correlated: $r(13) = 0.34, P = 0.18$.

At the lab, infants were seated on their parent's lap at a distance of 65 cm from a 42" screen. Parents wore headphones and listened to a neutral music-mix during the experiment. They could not hear the stimuli presented to the infant. The infant's gaze was recorded with an EyeLink 1000 eye-tracker at a frequency of 500 Hz. A five-point infant-friendly calibration was used. After the calibration, the experiment began. The experiment was coded in Python 3.5, using the Psychology 2.7 toolbox (all codes are available on the study's OSF page).

Infants first viewed the training video (for the fourth time). The test phase followed, in which infants were presented with two images on the screen, one on the left and one on the right (each about 30 cm × 30 cm). Each trial had one animate and one inanimate image. Presentation of animate and inanimate images, as well as the animacy and presentation side of the target were counterbalanced. Images were presented in silence for 2 s, then a prompting sentence began to play, asking the child to look at one of the two images. 1 s after the prompting sentence, the sentence was repeated. The trial ended 4 s after the end of the second repetition. Each trial lasted approximately 10 s. Trial length was maximized, while balancing total laboratory test time, to increase measurement reliability (44).

The test phase began with two training trials and then continued to a mix of test and filler trials. Test trials always included one novel animate image and one novel inanimate image; filler trials always included one familiar animate image and one familiar inanimate image. In the middle of the test phase, there was a short interlude video (~30 s), during which the woman played with toys but did not name them or use the novel determiners (video available on the study's OSF page). The interlude video served as a short but fun distraction.

The experiment lasted approximately 12 min. The experimenter was blinded to the test noun and item pairings.

Analyses. To determine whether infants looked more toward the animate image when they heard the animate determiner *ko* and a novel noun (e.g., *ko bamoule*), than when they heard the inanimate determiner *ka* and a novel noun (e.g., *ka pardale*), infants' gaze data were submitted to a cluster-based permutation analysis to investigate looking-while-listening patterns (46) and were submitted to a linear mixed-effects model to investigate overall looking preference. The two analyses were chosen to detect an effect across different processing patterns. The cluster-based permutation analysis provides fine-grained information about when participants look more to a target image, but is less appropriate when the effect varies in onset time across participants (e.g., when participants rely on different cognitive mechanisms, or when there is a lot of variability in decision time across participants); the linear mixed-effects model, on the other hand, sums looks across the trial: it thus provides coarse-grained information about general preference in light of individual variation but has blind spots when the effect is short with respect to the total trial time (e.g., when participants continue to explore during the remainder of the trial).

Trials where gaze away from the screen accounted for more than 50% of the total trial time were excluded from the analyses. This threshold was used to exclude trials where a low signal-to-noise ratio arose from insufficient data or infants' inattention. The analyses were all run from the point in time at which the determiner was heard to the end of the trial (0 to 7,500 ms). The analysis time-window was maximized to increase measurement reliability (44) and to avoid imbuing the analyses with assumptions. The analyses were computed using the eyetrackingR package (47) in R3.4.4.

For the cluster-based permutation analysis, data were down-sampled to 50 Hz, by averaging adjacent data-points into 20-ms bins. The analysis ran a *t* test on the arcsine-transformed proportion of looks toward the animate image at each time-point, when infants heard the animate determiner and when they heard the inanimate determiner. It grouped the adjacent time-points with a *t*-value greater than the predefined threshold of 1.5 into a cluster. A 1000 permutations were run. The cluster-based permutation analysis revealed a significant cluster ($P = 0.01$), between 300 and 2,440 ms, indicating that during that time-window infants looked more toward to animate image when they heard a prompting sentence with the animate determiner *ko* and a novel noun, than when they heard the inanimate determiner *ka* and a novel noun.

The linear mixed-effects regression analysis had overall looking time toward the animate image (0 to 7,500 ms) as the dependent variable, condition (animate or inanimate determiner) as the independent variable, and participant as a random intercept. The analysis was computed using the lme4 package (48) in R3.4.4. The mixed-effects regression revealed a moderate effect of condition, with infants looking longer to the animate image when they heard the animate determiner *ko* and a novel noun, than when they heard the inanimate determiner *ka* and a novel noun: $\beta = -0.09$, $SE = 0.04$; model comparison: $\chi^2(1) = 4.12$, $P = 0.04$; Cohen's $d = 0.53$. To verify that our preregistered mixed-effects regression was an appropriate analysis for the data, we checked the distribution of residuals graphically and with a Shapiro-Wilk test. Neither the graphical examination nor the Shapiro-Wilk test ($W = 0.99$, $P = 0.83$) showed non-normality of residuals.

The same two analyses were conducted on the gaze data from the filler trials, to ensure that infants were able to correctly identify words they know when they were paired with novel determiners. A cluster-based permutation analysis with the same parameters as for test trials revealed two significant clusters, between 920 and 2,840 ms ($P < 0.001$) and 3,200 to 4,680 ms ($P = 0.002$). It confirmed that infants were able to correctly identify words they know when paired with the novel determiners *ko* and *ka*. A mixed-effects regression analysis with the same parameters as for test trials, confirmed a strong effect of condition, with infants correctly identifying the image corresponding to words they know (looks toward animate image when hearing animate sentences: ($M = 0.596$, $SD = 0.11$); vs. inanimate sentences: ($M = 0.438$, $SD = 0.123$): $\beta = -0.16$, $SE = 0.03$; model comparison: $\chi^2(1) = 18.96$, $P < 0.001$; Cohen's $d = 1.36$).

Data, Materials, and Software Availability. Eye tracking data (ASCII format) have been deposited on OSF: (DOI: [10.17605/OSF.IO/X8H3A](https://doi.org/10.17605/OSF.IO/X8H3A)) https://osf.io/zgpb8/?view_only=cd943c5166d0414d9a319c1b305c9730 (18).

ACKNOWLEDGMENTS. We thank participating families and Miollis public preschool; M. Dutat and V. Ul for technical help; S. Recht for help with the experiment code; Z. Wang and the Eyelink team for help with the code for the infant-friendly calibration; K. Sivakumar for help during the piloting phase; C. Crook for help during testing; E. Rolland and M. McCune for help coding stimuli; C. Yokoyama, B. Strickland, J. Gervain, N. Sebastián-Gallés, S. Tsuji, members of the Laboratoire de Sciences Cognitives et Psycholinguistique and members of the Tsuji Laboratory for helpful comments. Portions of this work were developed from the doctoral dissertation of M.B. This work was supported by the World Premier International Research Center Initiative (WPI), MEXT, Japan. Ecole Normale Supérieure PhD fellowship grant (M.Barbir). Fondation Fyssen postdoctoral study grant (M.Barbir, M.J.Babineau). Japan Society for the Promotion of Science post-doctoral fellowship grant P20722 (M.Barbir). Marie Skłodowska-Curie Actions postdoctoral fellowship grant 799380 (M.J.Babineau). Agence Nationale de la Recherche ANR-13-APPR-0012 LangLearn (A.C.). Agence Nationale de la Recherche ANR-17-CE28-0007-01 LangAge (A.C.). Agence Nationale de la Recherche ANR-17-EURE-0017 FrontCog (A.C.).

Author affiliations: ^aLaboratoire de Sciences Cognitives et Psycholinguistique, Ecole Normale Supérieure, Paris Sciences et Lettres, 75005 Paris, France; ^bInternational Research Center for Neurointelligence (World Premier International Research Center Initiative; WPI-IRCN), Institutes for Advanced Study (UTIAS), The University of Tokyo, Tokyo 113-0033, Japan; and ^cDepartment of Psychology, University of Toronto, Toronto, ON M5S 3G3, Canada

- J. Gervain, M. Nespior, R. Mazuka, R. Horie, J. Mehler, Bootstrapping word order in prelexical infants: A Japanese-Italian cross-linguistic study. *Cogn. Psychol.* **57**, 56–74 (2008).
- R. Shi, M. Lepage, The effect of functional morphemes on word segmentation in preverbal infants. *Dev. Sci.* **11**, 407–413 (2008).
- R. Shi, A. Melançon, Syntactic categorization in French-learning infants. *Infancy* **15**, 517–533 (2010).
- A. X. He, J. Lidz, Verb learning in 14- and 18-month-old English-learning infants. *Lang. Learn. Dev.* **13**, 335–356 (2017).
- J. Gleitman, The structural sources of verb meanings. *Lang. Acquis.* **1**, 3–55 (1990).
- L. Naigles, Children use syntax to learn verb meanings. *J. Child Lang.* **17**, 357–374 (1990).
- R. W. Brown, Linguistic determinism and the part of speech. *J. Abnorm. Soc. Psychol.* **55**, 1–5 (1957).
- A. Seidl, G. Hollich, P. W. Juszczyk, Early understanding of subject and object wh-questions. *Infancy* **4**, 423–436 (2003).
- M. Babineau, R. Shi, A. Christophe, 14-month-olds exploit verbs' syntactic contexts to build expectations about novel words. *Infancy* **25**, 719–733 (2020).
- Y. Gertner, C. Fisher, J. Eisengart, Learning words and rules: Abstract knowledge of word order in early sentence comprehension. *Psychol. Sci.* **17**, 684–691 (2006).
- C. Fisher, Structural limits on verb mapping: The role of abstract structure in 2.5-year-olds' interpretations of novel verbs. *Dev. Sci.* **5**, 55–64 (2002).
- S. Yuan, C. Fisher, "Really? She blinked the baby?" Two-year-olds learn combinatorial facts about verbs by listening. *Psychol. Sci.* **20**, 619–626 (2009).
- L. Gerken, A. Bollt, Three exemplars allow at least some linguistic generalizations: Implications for generalization mechanisms and constraints. *Lang. Learn. Dev.* **4**, 228–248 (2008).
- C. Marino, C. Bernard, J. Gervain, Word frequency is a cue to lexical category for 8-month-old infants. *Curr. Biol.* **30**, 1380–1386 (2020).
- G. F. Marcus, S. Vijayan, S. Bandi Rao, P. M. Vishton, Rule learning by seven-month-old infants. *Science* **283**, 77–80 (1999).
- J. Lany, J. R. Saffran, From statistics to meaning: Infants' acquisition of lexical categories. *Psychol. Sci.* **21**, 284–291 (2010).
- C. L. Hudson Kam, E. L. Newport, Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Lang. Learn. Dev.* **1**, 151–195 (2005).
- M. Barbir, M. Babineau, A.-C. Fievet, A. Christophe, Can infants learn new determiners, classifying nouns based on their animacy feature? *Open Science Framework* (2022, October 6). <https://doi.org/10.17605/OSF.IO/ZG8P8>.
- M. Barbir, "The way we learn" (Ecole Normale Supérieure, Paris Sciences et Lettres, Paris, France, 2019).
- M. Babineau, A. de Carvalho, J. Trueswell, A. Christophe, Familiar words can serve as a semantic seed for syntactic bootstrapping. *Dev. Sci.* **24**, e13010 (2021).
- J. M. Mandler, L. McDonough, Concept formation in infancy. *Cogn. Dev.* **8**, 291–318 (1993).
- D. H. Rakison, D. Poulin-Dubois, Developmental origin of the animate-inanimate distinction. *Psychol. Bull.* **127**, 209–228 (2001).
- B. Ferguson, E. Graf, S. R. Waxman, Infants use known verbs to learn novel nouns: Evidence from 15- and 19-month-olds. *Cognition* **131**, 139–146 (2014).
- B. Strickland, Language reflects "core" cognition: A new theory about the origin of cross-linguistic regularities. *Cogn. Sci.* **41**, 70–101 (2017).
- L. R. Gleitman, K. Cassidy, R. Nappa, A. Papafragou, J. C. Trueswell, Hard words. *Lang. Learn. Dev.* **1**, 23–64 (2005).
- T. H. Mintz, Frequent frames as a cue for grammatical categories in child directed speech. *Cognition* **90**, 91–117 (2003).
- A. Gutman, I. Dautriche, B. Crabbé, A. Christophe, Bootstrapping the syntactic bootstrapper: Probabilistic labeling of prosodic phrases. *Lang. Acquis.* **22**, 285–309 (2014).
- P. Brusini, O. Seminc, P. Amsili, A. Christophe, The acquisition of noun and verb categories by bootstrapping from a few known words: A computational model. *Front. Psychol.* **12**, 661479 (2021).
- A. E. Ural, D. Yuret, F. N. Ketrez, D. Kocbas, A. C. Kuntay, Morphological cues vs. number of nominals in learning verb types in Turkish: The syntactic bootstrapping mechanism revisited. *Lang. Cogn. Neurosci.* **24**, 1393–1405 (2009).
- D. A. Baldwin, Infants' contribution to the achievement of joint reference. *Child Dev.* **62**, 875–890 (1991).
- P. Setoh, D. Wu, R. Baillargeon, R. Gelman, Young infants have biological expectations about animals. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 15937–15942 (2013).
- P. Bloom, Language capacities: Is grammar special? *Curr. Biol.* **9**, R127–R128 (1999).
- A. Fernald, R. Zangl, A. L. Portillo, V. A. Marchman "Looking while listening: Using eye movements to monitor spoken language" in *Language Acquisition & Language Disorders. Developmental Psycholinguistics: On-Line Methods in Children's Language Processing*, I. A. Sekerina, E. M. Fernandes, H. Clahsen, Eds. (John Benjamins, 2008), vol. **44**, pp. 97–135.
- E. Maris, R. Oostenveld, Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* **164**, 177–190 (2007).
- E. Bates, B. MacWhinney, "Functionalism and the competition model" in *The Crosslinguistic Study of Sentence Processing*, B. MacWhinney, E. Bates, Eds. (Cambridge University Press, 1989), pp. 3–76.
- J. Trueswell, L. Gleitman, Children's eye movements during listening: Developmental evidence for a constraint-based theory of sentence processing in *The Interface of Language, Vision, and Action*, J. M. Henderson, F. Ferreira, Eds. (Psychology Press, 2004), pp. 319–346.
- K. S. Tummelshammer, D. Mareschal, N. Z. Kirkham, Infants' selective attention to reliable visual cues in the presence of salient distractors. *Child Dev.* **85**, 1981–1994 (2014).
- C. A. Best, H. Yim, V. M. Sloutsky, The cost of selective attention in category learning: Developmental differences between adults and infants. *J. Exp. Child Psychol.* **116**, 105–119 (2013).
- A. Gopnik et al., Changes in cognitive flexibility and hypothesis search across human life history from childhood to adolescence to adulthood. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 7892–7899 (2017).
- N. J. Blanco, V. M. Sloutsky, Adaptive flexibility in category learning? Young children exhibit smaller costs of selective attention than adults. *Dev. Psychol.* **55**, 2060–2076 (2019).
- I. Dautriche, L. Fibla, A. C. Fievet, A. Christophe, Learning homophones in context: Easy cases are favored in the lexicon of natural languages. *Cogn. Psychol.* **104**, 83–105 (2018).
- L. M. Oakes, Sample size, statistical power, and false conclusions in infant looking-time research. *Infancy* **22**, 436–469 (2017).
- M. C. DeBolt, M. Rhemtulla, L. M. Oakes, Robust data and power in infant research: A case study of the effect of number of infants and number of trials in visual preference procedures. *Infancy* **25**, 393–419 (2020).
- M. Zettersten et al., Peekbank, Exploring children's word recognition through an open, large-scale repository for developmental eye-tracking data. *Cognit. Sci.* **43**, 2950–2956 (2021).
- S. Kern, Lexicon development in French-speaking infants. *First Lang.* **27**, 227–250 (2007).
- C. D. Luche, S. Durrant, S. Poltrock, C. Floccia, A methodological investigation of the intermodal preferential looking paradigm: Methods of analyses, picture selection and data rejection criteria. *Infant Behav. Dev.* **40**, 151–172 (2015).
- J. Dink, B. Ferguson, eyetrackingR. The Comprehensive R Archive Network (R package version 0.1.6, Vienna, Austria, 2016). <https://cran.r-project.org/web/packages/eyetrackingR/index.html>.
- D. Bates, M. Maechler, B. Bolker, S. Walker, Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).