Young children spontaneously recreate core properties of language in a new modality

Manuel Bohn[a,b,1,2], **Gregor Kachel**[c,1], **and Michael Tomasello**[d,e]

[a]Department of Psychology, Stanford University, Stanford, CA 94305; [b]Leipzig Research Center for Early Child Development, Leipzig University, 04109 Leipzig, Germany; [c]Faculty of Educational Sciences, Leipzig University, 04109 Leipzig, Germany; [d]Department of Developmental and Comparative Psychology, Max Planck Institute for Evolutionary Anthropology, 04103 Leipzig, Germany; and [e]Department of Psychology and Neuroscience, Duke University, Durham, NC 27708

Edited by Susan A. Gelman, University of Michigan, Ann Arbor, MI, and approved October 28, 2019 (received for review March 24, 2019)

How the world's 6,000+ natural languages have arisen is mostly unknown. Yet, new sign languages have emerged recently among deaf people brought together in a community, offering insights into the dynamics of language evolution. However, documenting the emergence of these languages has mostly consisted of studying the end product; the process by which ad hoc signs are transformed into a structured communication system has not been directly observed. Here we show how young children create new communication systems that exhibit core features of natural languages in less than 30 min. In a controlled setting, we blocked the possibility of using spoken language. In order to communicate novel messages, including abstract concepts, dyads of children spontaneously created novel gestural signs. Over usage, these signs became increasingly arbitrary and conventionalized. When confronted with the need to communicate more complex meanings, children began to grammatically structure their gestures. Together with previous work, these results suggest that children have the basic skills necessary, not only to acquire a natural language, but also to spontaneously create a new one. The speed with which children create these structured systems has profound implications for theorizing about language evolution, a process which is generally thought to span across many generations, if not millennia.

language | evolution | gesture | cognitive development

One of the great discoveries in the human sciences in recent decades is the finding that, when spatially isolated deaf persons are brought together into a community, they rapidly create a full-fledged conventional language in the gestural modality (1–3). Given that natural languages presumably took many millennia to evolve, it is startling that contemporary social groups, including groups of children, are able to do something similar in only a few generations. The original discovery focused on deaf children brought together in an educational setting and involved several cohorts over a number of years, but subsequent research has identified other instances in which such a community was formed in different ways (4). The fact that this process invariably takes place in the gestural modality—along with observations that reference is more readily established in gesture than vocalization (5, 6)—suggests that perhaps human linguistic communication originated in gesture (7).

The documentation of newly created sign systems has provided invaluable insights into the process of language emergence. First of all, language-like communication can arise without a language model. Deaf children born to hearing parents who do not receive any sign language instruction develop communication systems—called homesign—that share core features with spoken languages (8–11). Second, when these homesigners are brought together in a community, they rapidly converge on a conventional communication system (1, 3, 12). Finally, these systems change over time, and children who are exposed to them at a younger age end up with more complex and structured systems as adults (13, 14).

However, the study of newly emerging sign languages has mostly been a process of natural observation of the end product, and the actual process of creation has never been directly observed. Furthermore, naturalistic settings precluded experimental manipulation of the process. The latter point has been partially addressed by artificial language-learning studies. In experiments with speaking participants, researchers have studied how adults and children transmit artificial sign systems across generations (e.g., refs. 15 and 16). This approach has illuminated some of the dynamics of cultural language evolution. For example, languages become more compositionally structured over generations and, as a consequence, are easier to learn (15, 17). Somewhat at odds with the naturalistic observations described above, researchers in this tradition have found that children do not spontaneously structure artificial languages in a compositional way (16). Presumably, this is due to the high demands associated with learning an artificial language in the first place. More naturalistic paradigms, for example, using gesture instead of arbitrary symbols, might be more likely to evoke compositional structure in children (18). Yet, like the studies on newly emerging sign languages, studies on experimental language evolution do not answer the question of how (and how fast) a language-like communication system may spontaneously arise out of signals created ad hoc during social interaction.

Here we study the incipient steps of the emergence of novel communications systems in a comprehensive way. Our study design

Significance

Human social and cultural life rests on a unique set of communicative abilities. Newly emerging communication systems provide a window into the cognitive and interactional basis of these skills. In a muted video call setup, children were required to convey meaning to a partner while being unable to use spoken language. The gestural code systems that children created ad hoc within a 30-min test session exhibited core features of natural language and emergent sign languages: referential signals for objects, actions, and abstract concepts, conventional use of these signals, and grammatical structure. These results demonstrate that language-like communication systems can emerge rapidly out of social interaction.

Author contributions: M.B., G.K., and M.T. designed research; M.B. and G.K. performed research; M.B. analyzed data; and M.B., G.K., and M.T. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Data deposition: A version of *SI Appendix* is available at https://manuelbohn.github.io/ges3000/ges3000.html. Stimulus pictures, data files, and analysis scripts are available at https://github.com/manuelbohn/ges3000.

[1]M.B. and G.K. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: manuel.bohn@uni-leipzig.de.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1904871116/-/DCSupplemental.

First published December 2, 2019.

unites the strength of previous approaches: In a controlled laboratory setting, participants played a reference game in which one child had to communicate the content of a picture to another child. While we blocked the possibility of using spoken language, children were free to use whatever form of communicating that they deemed suitable. By varying the complexity of the scenes depicted in the pictures that had to be communicated, we were able to study the influence of communicative pressures on the developing gestural code systems. In order to approximate the cognitive prerequisites for the creation of novel communication systems, we included children of different ages.

In the following described experiments, we used this approach to test if children recreate core properties of language: referentiality, conventionality, arbitrariness, and grammatical structure. As a first step, we tested how children establish reference in a new modality. Next, we asked if children could use their newly established mode of communication to communicate abstract concepts. Then we investigated the temporal dynamics of the process, that is, whether communicative partners converge on a set of conventional signals and whether signals become more arbitrary over time. Finally, we investigated when and how children use compositional structure to communicate more complex meanings.

Results

Initiation of Communication and Uptake of Signaling. Dyads of 4- and 6-y-old children engaged in a communication game in which a sender had to communicate the content of a picture to a recipient. Participants were situated in different rooms, interconnected by an audio-video channel (akin to a video call; Fig. 1). In each room, a set of 5 candidate pictures were placed on a wooden board (*SI Appendix*, Fig. S2). After a short training in which the children could convey the meaning verbally to establish common ground, the experimenter cut the audio connection and children continued playing the game without any further instructions. This situation (initiation phase) allowed us to assess whether children would spontaneously find a new way to express the content of the picture. Apart from one unsuccessful attempt to use lip reading, children used iconic gestures. If children did not spontaneously establish successful communication, experimenters provided stepwise prompts.

The majority of 6-y-olds spontaneously had already produced gestures in trial 1, more so compared to 4-y-olds (trial 1: $\beta = -4.01$, SE = 1.3, $P < 0.001$; trials 1 to 4: $\beta = -4.35$, SE = 1.14, $P < 0.001$) (Fig. 2*A*). Even though 4-y-olds initially needed prompts to produce gestures, they needed fewer prompts in later trials, showing that they transferred the idea of iconically depicting a referent to new pictures (main effect of trial: $\beta = -2.76$, SE = 1, $P = 0.002$). Thus, while both age groups successfully adopted a novel means of communication (iconic gestures) to coordinate, 6-y-olds did so independently. In a follow-up study, we tested 3- and 4-y-olds in the same setup but with an adult comprehension partner. Replicating the previous results, we found that 4-y-olds did not initially create iconic gestures but quickly transferred this mode of communication to new pictures. Three-year-olds, on the other hand, needed prompts from the experimenter for each new picture and independently produced gestures only when the pictures were repeated (see *SI Appendix* for detailed results). Taken together, we found a salient developmental pattern: Six-year-olds spontaneously invented novel referential signs. Four-year-olds quickly adopted a new mode of communication and used it productively to create novel gestures. Three-year-olds, however, imitated model solutions and had to be introduced to novel meanings in a piecemeal fashion.

Toward the end of the game, children switched roles. The sender became the recipient and vice versa (uptake phase). At that point, children of both ages spontaneously used iconic gestures from trial 1 onward (*SI Appendix*, Fig. S3*B*), often adopting the signs that their partner invented (see below). There was no difference between age groups at trial 1 ($P = 0.093$, Wilcoxon test). Furthermore, fewer 4-y-olds needed prompts in trial 1 during this phase as compared to the initial phase of the experiment ($\beta = -3.09$, SE = 1.21, $P = 0.002$). These results show that, once one child established a way of referring to the pictures, the other child easily picked it up.

Turning to comprehension, we found that gestures were understood at a very high rate (Fig. 2*B* and *SI Appendix*, Fig. S4). Comprehension of the gestures improved with trial ($\beta = 2.85$, SE = 1.29, $P = 0.003$) and was slightly better in uptake compared to initiation ($\beta = 2.16$, SE = 1.17, $P = 0.045$). We found no difference between age groups ($\beta = 0.79$, SE = 1.00, $P = 0.456$). This shows that children generally depicted the content of the pictures in a way that was comprehensible to their partner.

Communicating Abstract Concepts. The pictures used in the first 4 trials all depicted concrete and familiar actions on objects. To see if children would be able to communicate more abstract concepts, we introduced a plain white picture ("nothing" or "empty") over the course of the session. Both 4- and 6-y-olds spontaneously produced intentional gestures for this item and did not differ in the respective rates (initiation: $P = 0.093$; uptake: $P = 0.217$; Wilcoxon test). Prototypic solutions were ostensibly standing still (doing "nothing") or shaking one's head. Comprehension partners understood these gestures at a rate above chance (all $P < 0.01$, except for 4-y-olds in the initiation phase, $P = 0.489$, Wilcoxon test), presumably by eliminating the other pictures as referents for the new gestures. Rate of comprehension was higher for 6-y-olds ($\beta = 2.94$, SE = 1.43, $P = 0.004$) and during the uptake part ($\beta = 3.51$, SE = 1.32, $P < 0.001$). Thus,

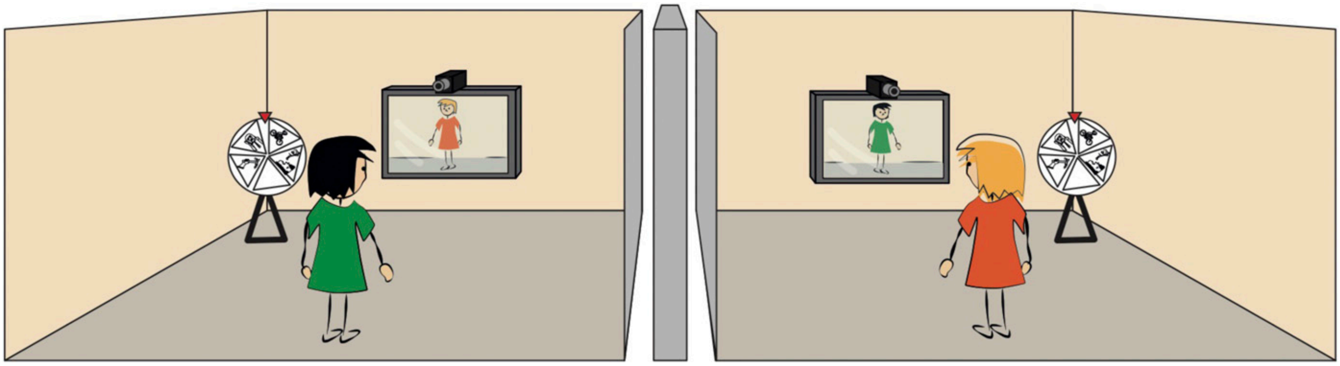

Fig. 1. Schematic drawing of the general setup. Children were placed in different rooms and communicated via a video channel. Candidate pictures were presented on a wooden board.

PSYCHOLOGICAL AND COGNITIVE SCIENCES

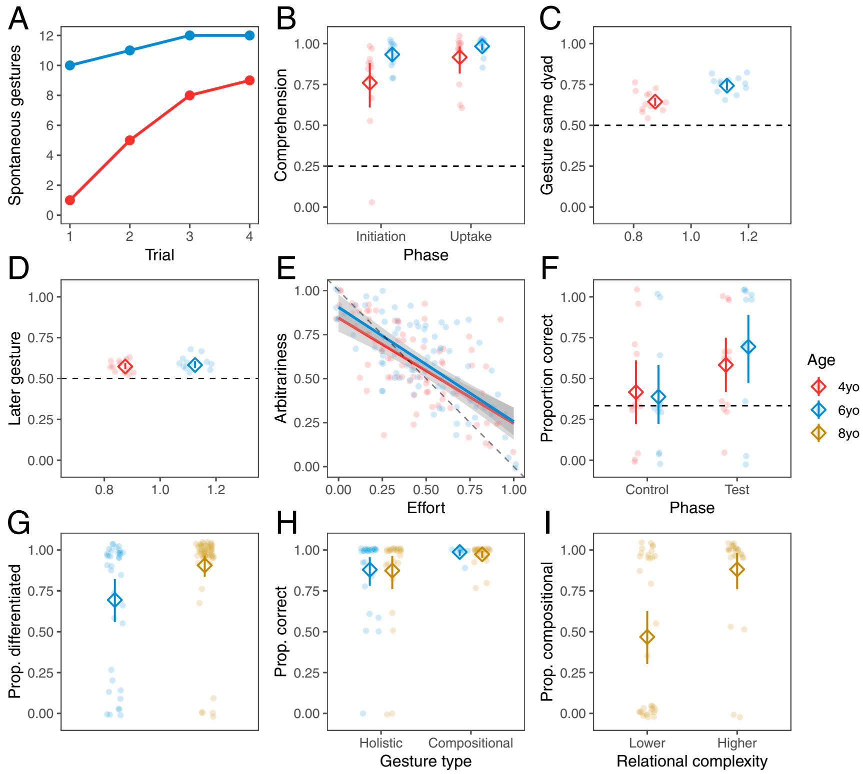

Fig. 2. Results for experiments on creation of novel communication systems (*A*–*F*) and use of grammatical structure (*G*–*I*). (*A*) Dyads who spontaneously produced gestures in the first 4 trials, each of which involved a different picture. (*B*) Comprehension of gestures in the initiation (before role switch) and uptake phase. (*C*) Rate at which naive raters judged gestures from the same dyad to be more similar compared to gestures from a different dyad. (*D*) Rate at which raters judged later gestures to be more arbitrary. (*E*) Relation between ratings of arbitrariness and effort. (*F*) Rate of comprehension of arbitrary gestures by children who witness the form drift to arbitrary (test) and children who see only the final, arbitrary gesture (control). (*G*) Rate of differentiation across conditions. (*H*) Comprehension for gestures composed of either a single holistic gesture or a compositional gesture sequence. (*I*) Proportion of compositional gesture sequences by relation complexity. "Lower" refers to conditions with one modifying predicate (movement, number, and size) and "Higher" to transitive actions between 2 agents (transitive 2 and 3). In *B*–*I*, jittered transparent dots represent individual means, diamonds show group means, and vertical lines show 95% CIs. Dashed lines in *B*, *D*, and *F* denote performance expected by chance.

children were able to successfully communicate more abstract concepts using newly created gestures. In later experiments (see section *Grammatical Constructions and Communicative Pressure*) children also created separate gestures for abstract relational concepts such as "big."

Conventionalization. A signature of mature languages is conventionality: speakers within a community use the same signs for a given referent. In emerging sign languages, conventionality emerges only once homesigners join in larger communities (3). Here, to assess whether children within a dyad used more similar gestures than children from different dyads, we asked naive adults to make similarity judgments for gestures from the same dyad compared to gestures from a different dyad. For both age groups, raters judged gestures from the same dyad to be more similar (both $P < 0.001$, Fig. 2*C*). In direct comparison, gestures were rated slightly more similar for 6- compared to 4-y-olds ($\beta = 0.78$, SE = 0.39, $P = 0.049$). Nevertheless, both age groups produced more similar gestures within than between dyads. (Fig. 2*C* and *SI Appendix*, Fig. S5). Partners adopted the sender's sign, copying even for subtle differences. As senders generally comprehended signs that they had just used, copying greatly increased successful communication.

Drift to the Arbitrary. Spoken languages involve mostly arbitrary mappings of form and meaning. Individual homesigners rely mostly on iconic signs (3). Signed languages also start out with a large number of iconic signs which become more arbitrary over time (19, 20). This reflects a general tendency for human communication systems to "drift to the arbitrary" (7). We investigated this process in 2 ways: For production, we asked whether gestures produced later in the experiment were more arbitrary compared to earlier gestures. For comprehension, we tested whether children would continue to recognize the meaning of a gesture when its form rapidly drifted from highly iconic to completely arbitrary.

To see if children's gestures become more arbitrary over time, we compared the first and the last gesture for a given picture from each child. Raters were instructed to select the gesture that they deemed more abstract (i.e., less evocative of the picture). An explanation for gestures drifting to the arbitrary over time is that iconic elements are dropped from the gesture to save effort as they become unnecessary for efficient communication. Therefore, we also asked raters to indicate which gesture was more effortful. As children participated in up to only 5 trials with each picture (4-y-olds in up to only 3 trials), we invited children back on a second day and repeated the procedure with fewer pictures (*SI Appendix*, Fig. S6) and more repetitions. Raters judged later gestures to be less transparent and more abstract with regard to their form (Fig. 2*D* and *SI Appendix*, Fig. S7*A*). We analyzed the influence of age and effort on abstractness ratings separately for each day. The later gestures of 6-y-olds were rated as more abstract more often compared to those of 4-y-olds on day 1 ($\beta = 0.56$, SE = 0.23, $P = 0.019$). Rater judgments of how abstract forms appeared to them were strongly influenced by ratings of perceived effort ($\beta = -2.30$, SE = 0.17, $P < 0.001$, Fig. 2*E*). On day 2, abstractness ratings were again strongly influenced by effort ($\beta = -2.69$, SE = 0.22, $P < 0.001$) but, in contrast to day 1, not by age ($\beta = 0.02$, SE = 0.30, $P = 0.95$). This suggests that arbitrariness of a gesture depends on how often it has been used rather than on the age of the gesturer (6-y-olds received more trials with each picture on day 1 but not on day 2). The effect for abstractness ratings was rather small, and later gestures were still highly iconic. Nevertheless, we found evidence that children's production shifted toward more arbitrary mappings between sign and referent over just a few iterations with the same partner.

To investigate comprehension, we paired children with a scripted adult partner. The adult started out with highly iconic gestures, but produced more and more arbitrary gestures over time until the gestures were completely arbitrary (see associated online repository for sample videos). The arbitrariness of the final gestures was validated in a control group of children who were presented with the same arbitrary signs but without having shared the interaction with the partner. Children who witnessed the gestures become arbitrary over time selected the correct picture above chance (Fig. 2*F* and *SI Appendix*, Table S3) and did so more often than children in the control group (β = 1.35, SE = 0.56, P = 0.013), with no difference between age groups (β = 0.20, SE = 0.53, P = 0.710). Children of both ages successfully retained the meaning of a gesture when its form gradually drifted from iconic to arbitrary. It is noteworthy that children's comprehension was consistently accurate across trials, which indicates that they reidentified gestures rather than interpreting each alteration of a gesture as a new sign and failing to map it to the intended referent (cf. ref. 21).

Grammatical Constructions and Communicative Pressure. Next, we studied if and how 6- and 8-y-old children would impose grammatical structure onto their gestures when asked to communicate more complex meanings. Instead of differentiating between distinct referents, children were asked to disambiguate between depictions of the same referent with varying properties. We asked if children would spontaneously create compositional gesture sequences reminiscent of grammatical constructions. To qualify as a gesture sequence, the utterance had to be composed of separate meaningful gestures that could flexibly be recombined to communicate diverse meanings.

In a first step, we tested if children would produce sequences in which separate gestures are used to denote properties of referents. For example, when asked to contrast a big with a small duck, we wanted to see if children would produce a separate gesture for "big" and a separate gesture for "duck." We classified such sequences as compositional because the individual gestures denoting properties and referents could be recombined to communicate different meanings (the same gesture for "big" could be used to communicate about a big fork). Here we tested 3 conditions (*SI Appendix*, Figs. S9 and S12): In "movement," we contrasted moving vs. static referents. In "number," pictures differed in numerosity (one vs. many). In "size," we contrasted big and small depictions of the same referent. In a second step, we asked if children would further syntactically structure gesture sequences to communicate relations between referents. That is, would children use different structures when communicating "monkey chases cat" compared to "cat chases monkey." We conducted trials under 2 conditions, transitive 2 and transitive 3 (*SI Appendix*, Fig. S15): In transitive 2, we showed transitive actions (e.g., chasing) between 2 agents. Transitive 3 involved the same transitive actions between 2 agents but involved 3 possible interactants in total (see *SI Appendix* for details). In both steps, we contrasted the production of gesture sequences with the use of holistic gestures in which the property (or relation) was part of the depiction of the referent itself (e.g., marking a duck as big by extensively flapping the arms).

Across steps and conditions, we asked the following questions: 1) Do children differentiate between referents with different properties or relations and what affects the rate of differentiation? 2) When are compositional sequences used rather than holistic gestures? 3) How do compositional sequences affect comprehension? 4) Is the structure in children's gesture sequences influenced by the structure of their native language? Age effects were tested on the data from movement, number, and size. An additional analysis included data from 8-y-olds in all 5 conditions. The focus was on gesture production, and so we paired children with an adult comprehension partner.

First, we looked at the rate of differentiation across conditions. That is, we asked if children made any attempt to express the property of, or relation between, referents. In the transitive 2 and 3 conditions, differentiation also involved specifying the role of the characters involved, and children thus had to use syntactic structure to do so. Children differentiated in the majority of trials (Fig. 2*G* and *SI Appendix*, Fig. S18). The rate of differentiation did not differ across conditions (P = 0.190 both age groups and P = 0.139 for 8-y-olds) or between age groups (β = 0.97, SE = 0.63, P = 0.112). Results by condition can be found in *SI Appendix*.

Next we asked what influenced children's use of compositional gesture sequences as opposed to holistic gestures. Depicting a single property of a referent (its movement, number, or size) is possible by modifying an already existing gesture for the referent. On the other hand, expressing a transitive relation between 2 agents requires specifying the agent and patient, respectively. We expected children to first produce separate gestures for agent, patient, and predicate and then convert the transitive relation between agent and patient into a temporal one, thus specifying the interactional roles via the order in which they produce the gestures. Doing so results in a gesture sequence with syntactical structure, in which the same gestures were used in a different order to express different meanings. In order to create these sequences, children had to create separate gestures for subject, predicate, and object. As a consequence, we expected 8-y-olds to use more compositional gesture sequences in transitive 2 and 3 conditions compared to the other 3 conditions.

First, we analyzed the data for both age groups. Somewhat at odds with our prediction, children were more likely to use gesture sequences in size compared to movement (β = −1.97, SE = 0.74, P = 0.008) or number (β = −2.00, SE = 0.73, P = 0.006). Furthermore, older children were more likely to use gesture sequences (β = 1.36, SE = 0.62, P = 0.022). In *SI Appendix*, we offer a post hoc explanation of why size produced more compositional gesture sequences. We used the data from 8-y-olds to directly test if gestures for transitive actions were more likely to be communicated using a compositional gesture sequence than via holistic gestures. We classified conditions showing transitive actions (transitive 2 and 3: 2 arguments) as high relational complexity and conditions involving only a single modifying predicate (movement, number, and size: 1 argument) as low relational complexity. Gesture sequences were more frequent in conditions with high compared to low relational complexity (β = 2.05, SE = 0.71, P = 0.001, Fig. 2*I*). This pattern confirms our initial hypothesis that children use structure to cope with increased complexity (i.e., a larger number of agents and relations) in the scene to be depicted. Next we directly tested this compensatory relation by asking whether compositional gesture sequences reduced ambiguity and improved comprehension.

Both holistic and compositional gestures were understood at high rates (Fig. 2*H* and *SI Appendix*, Fig. S20). However, compositional gestures were better understood. This was the case when analyzing the data for both age groups in movement, number, and size (β = 3.64, SE = 1.1, $P < 0.001$) as well as when analyzing all 5 conditions for 8-y-olds (β = 1.99, SE = 0.8, P = 0.006). Compositional gesture sequences were almost always understood, which illustrates how structure reduced ambiguity.

Final analyses tested whether the internal structure that children impose on their gesture sequences was influenced by their native language (German). Regardless of condition, for each trial with a gesture sequence, we coded whether the word order was the same as in German. We used the active present tense as the reference. For movement, number, and size, the corresponding German word order was predicate–subject (PS) [movement: "hüpfender Elefant" (jumping elephant); number: "viele Hämmer" (many hammers); size: "kleine Gabel" (small fork)]. In the case of transitive 2 and 3, it was subject–predicate–object (SPO) ["Hase jagt Katze" (bunny chases cat)]. Because the majority of the world's natural languages put subject before object (22), we focused on the positioning of the predicate instead (see refs. 23–25 for the same approach). We found no substantial evidence for German word order (β = 0.5, SE = 0.31, P = 0.105) and no difference in the rate of German word order across conditions (P = 0.534). This corroborates previous work with adults showing that structure in gesture is independent of grammatical structure in spoken language (23–25). Overall, variation rather than preset principles characterized children's creation of gesture sequences showing that children spontaneously structured their gestures to solve a communicative problem.

Discussion

The studies described here demonstrate the resilience of children's communicative abilities (3). Children spontaneously used iconic gestures to initiate communication. They also invented gestures for abstract concepts. All gestures were readily understood by comprehension partners, and when it was their turn to communicate, they used the same gestures that their partner had introduced. Over time, children produced more arbitrary gestures. When confronted with gestures that rapidly drifted from highly iconic to arbitrary, children retained the meaning. When asked to communicate more complex meanings, children began to impose compositional structure on their signals. This structure was independent from their native language.

We found that the reduction in iconicity over time was largely explained by a decrease in production effort. This suggests that efficiency and usability shape language change. In support of this idea, iterated learning studies found that children reduce complexity when transmitting random dot patterns across generations (26). The results were patterns that were easier to use and to learn. Similarly, studies with adults show that, over time, adults produce signals that are easier to use even at the expense of clarity (27). While the introduction of ambiguity over time might seem counterintuitive, ambiguity can be seen as a feature of efficient communication systems (28). Signals do not have to be tightly coupled with a given referent (easing production and allowing reuse) when contextual information is available and social aspects of the interaction compensate for an increase in ambiguity. We showed that, when presented with rapid changes in a set of gestures from highly iconic to arbitrary, children are able to follow these changes and still comprehend the altered symbols at high rates. The meaning of the gesture is not just a function of the signal that is being used but also of the conversational history (common ground) (29, 30). Ambiguity is reduced by "outsourcing" information to common ground. In line with this idea, we saw that children used the same gestures for a referent and thereby established a local convention. Conventions are another way of leveraging the accumulation of shared information over the course of a social interaction. Natural languages can be thought of as local conventions that have been lifted to the group level (7).

Children in our study spontaneously imposed structure onto their gesture sequences. This finding echoes earlier work in which children were asked to use gestures to describe motion events (18). However, compositional gesture sequences were not a default but rather a consequence of an increase in relational complexity in the scene to be communicated. Holistic signals are constrained in their expressive power, especially when it comes to expressing relations between objects and agents. Importantly, the structure that children imposed was unrelated to their native language, suggesting that it was not the by-product of a mediating linguistic representation (23–25). These results show that compositional structure in communication not only emerges to increase fidelity and learnability, but also can be intentionally used as a tool to reduce ambiguity in the signal.

There are several ways in which the work should be extended. First of all, a greater variation of the stimulus material and the context should be used to test the generalizability of the results. Even though we found that children would spontaneously structure their gestures, we do not know if and how this structure is further modified when the communication system is transmitted across generations and/or used in larger groups. Based on prior work we would predict that structure would increase even further (13, 15, 17, 31). Assuming a ceiling effect on the degree to which a system can be compositionally structured, we would expect that emerging communication systems that start out in a structured way are less prone to change over time compared to systems that rely on holistic signals.

Across experiments, we found that 6-y-old children were able to quickly recreate core properties of a conventional, abstract, and structured communication system. Even though 8-y-olds produced more compositional gestures, the performance of 6-y-olds followed similar patterns. From a cognitive perspective, we may speculate that second-order theory of mind abilities (e.g., discerning what others believe about my beliefs) (32), as well as metalinguistic abilities (33), both of which are in place by that age, play crucial roles in the transition from being able to learn a language to being able to spontaneously create one. At the same age children also grasp the coordinative function of conventions as evident in their propensity to spontaneously create novel game rules (34). However, the extensive work with homesigners suggests that such complex systems might be created at even younger ages (8–11). Because these systems are created over a longer period of time, different cognitive processes might be at work. Studying the direct link between cognitive abilities and the creation of structured communication systems will be a fruitful avenue for future research.

Conclusion

We have documented in a controlled setting what might plausibly be the incipient steps in the emergence of a new language-like communicative system, which has allowed us to speculate about the general dynamics of this process. Reference is spontaneously established using iconic signs that are grounded in the interlocutors' shared experience. Through reciprocal imitation, signs and referents become linked in a conventional way. The initially transparent relations between iconic sign and referent recede due to an effort to streamline expressions (assuming knowledgeable partners), as well as a refinement of the referential scope of a signal. Finally, grammatical structure comes in a multitude of forms, emerging when more fine-grained distinctions are needed. Here, all of these processes occurred spontaneously in interactions of naive human children and in less than 30 min.

Materials and Methods

A version of *SI Appendix* is available at https://manuelbohn.github.io/ges3000/ges3000.html.

Participants. All experiments were approved by an internal ethics committee at the Max Planck Institute for Evolutionary Anthropology (Leipzig, Germany). All children ($n = 198$) came from an ethnically homogeneous, midsized German city, were mostly monolingual, and had mixed socioeconomic backgrounds. They were recruited from a database of children whose parents volunteered to take part in studies of child development. Parents accompanying their children were introduced to the procedure and gave informed consent on behalf of their children.

Setup and Procedure. We established an audio-video connection between television sets connecting 2 rooms in a child laboratory (Fig. 1). The direct digital visual interface (DVI) connection between camera and TV made sure that participants could interact in a smooth and contingent way. One of the rooms served as a Production Room (PR) and the other as a Comprehension Room (CR). Both rooms were located on the same hallway allowing experimenters (E1 and E2) to coordinate easily when stepping out. In each room, we installed a picture board. A picture was selected by placing a red arrow next to it. Pictures were not visible for partners and varied across experiments.

In each trial, E1 was with the child in PR [Production Child (PC)] and E2 with the child in CR [Comprehension Child (CC)]. Trials in all experiments were conducted as follows: E1 selected a picture in PR. On a visual signal from E1, both experimenters said "ready, set, go" and left the room. The task of the child in PR was to communicate the content of the selected picture to the child in CR.

The child in CR made a choice by selecting a picture on her board. Children received feedback about whether or not CC selected the correct picture.

All experiments started with a training phase in which the video as well as the audio connection was still intact and children could talk to coordinate. This familiarized children with the game. During training, each picture was selected once. After the last training trial, children and experimenters left the room and met in the hallway. One experimenter went into each room and cut the audio connection. Children were told that they would play the same game in a different way. Gestures were never mentioned as an alternative means of communication. Next, children went into their respective room again. In each test trial, children had 1 min to establish communication. If they remained passive, E1 entered the PR and prompted PC to use iconic gestures in 3 steps (see *SI Appendix* for details). Beyond the information given in the main text, *SI Appendix* lists gives detailed information about each experiment, including participant age, stimuli, procedure, number of trials, analysis, coding, and reliability.

Analysis. All data and analysis scripts are available at https://github.com/manuelbohn/ges3000. Analysis code is also included in *SI Appendix*. For all analysis, we used R (35), and for models we used the functions *glm* and *glmer* of the *lme4* package (version 1.1–17) (36). For mixed models, we used a maximal random effect structure. *P* values for fixed effects are based on likelihood ratio tests (37), which were computed via the function *drop1*.

Data and Materials Availability. Stimulus pictures, data files, and analysis scripts are available at https://github.com/manuelbohn/ges3000.

ACKNOWLEDGMENTS. We thank Tünde Malomvölgyi, Anna Seidel, Tom Hausmann, Anja Dreves, Claudia Salomo, Matthias Allritz, Ailis Cournane, and Cristina Zickert. M.B. received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement no. 749229.

1. M. J. Kegl, A. Senghas, M. Coppola, "Creation through contact: Sign language emergence and sign language change in Nicaragua" in *Language Creation and Language Change: Creolization, Diachrony, and Development*, M. DeGraff, Ed. (MIT Press, Cambridge MA, 1999), pp. 179–237.
2. A. Senghas, M. Coppola, Children creating language: How Nicaraguan sign language acquired a spatial grammar. *Psychol. Sci.* **12**, 323–328 (2001).
3. D. Brentari, S. Goldin-Meadow, Language emergence. *Annu. Rev. Linguist.* **3**, 363–388 (2017).
4. W. Sandler, I. Meir, C. Padden, M. Aronoff, The emergence of grammar: Systematic structure in a new language. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 2661–2665 (2005).
5. N. Fay, C. J. Lister, T. M. Ellison, S. Goldin-Meadow, Creating a communication system from scratch: Gesture beats vocalization hands down. *Front. Psychol.* **5**, 354 (2014).
6. M. Bohn, J. Call, M. Tomasello, Natural reference: A phylo- and ontogenetic perspective on the comprehension of iconic gestures and vocalizations. *Dev. Sci.* **22**, e12757 (2019).
7. M. Tomasello, *Origins of Human Communication* (MIT Press, Cambridge, MA, 2010).
8. S. Goldin-Meadow, H. Feldman, The development of language-like communication without a language model. *Science* **197**, 401–403 (1977).
9. D. Hunsicker, S. Goldin-Meadow, Hierarchical structure in a self-created communication system: Building nominal constituents in homesign. *Language (Baltim)* **88**, 732–763 (2012).
10. S. Goldin-Meadow, *The Resilience of Language: What Gesture Creation in Deaf Children Can Tell Us about How All Children Learn Language* (Psychology Press, New York, 2005).
11. L. Rissman, S. Goldin-Meadow, The development of causal structure without a language model. *Lang. Learn. Dev.* **13**, 286–299 (2017).
12. S. Goldin-Meadow, D. Brentari, M. Coppola, L. Horton, A. Senghas, Watching language grow in the manual modality: Nominals, predicates, and handshapes. *Cognition* **136**, 381–395 (2015).
13. A. Senghas, S. Kita, A. Ozyürek, Children creating core properties of language: Evidence from an emerging sign language in Nicaragua. *Science* **305**, 1779–1782 (2004).
14. J. L. Singleton, E. L. Newport, When learners surpass their models: The acquisition of American Sign Language from inconsistent input. *Cognit. Psychol.* **49**, 370–407 (2004).
15. S. Kirby, H. Cornish, K. Smith, Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 10681–10686 (2008).
16. L. Raviv, I. Arnon, Systematicity, but not compositionality: Examining the emergence of linguistic structure in children and adults using iterated learning. *Cognition* **181**, 160–173 (2018).
17. S. Kirby, M. Tamariz, H. Cornish, K. Smith, Compression and communication in the cultural evolution of linguistic structure. *Cognition* **141**, 87–102 (2015).
18. Z. Clay, S. Pople, B. Hood, S. Kita, Young children make their gestural communication systems more language-like: Segmentation and linearization of semantic elements in motion events. *Psychol. Sci.* **25**, 1518–1525 (2014).
19. N. Frishberg, Arbitrariness and iconicity: Historical change in American sign language. *Language* **51**, 696–719 (1975).
20. N. Fay, B. Walker, N. Swoboda, S. Garrod, How to create shared symbols. *Cogn. Sci.* **42** (suppl. 1), 241–269 (2018).
21. S. C. Creel, Phonological similarity and mutual exclusivity: On-line recognition of atypical pronunciations in 3–5-year-olds. *Dev. Sci.* **15**, 697–713 (2012).
22. J. H. Greenberg, *Language Universals: With Special Reference to Feature Hierarchies* (Walter de Gruyter, 2010).
23. E. Gibson *et al.*, A noisy-channel account of crosslinguistic word-order variation. *Psychol. Sci.* **24**, 1079–1088 (2013).
24. S. Goldin-Meadow, W. C. So, A. Ozyürek, C. Mylander, The natural order of events: How speakers of different languages represent events nonverbally. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 9163–9168 (2008).
25. A. Kocab, H. Lam, J. Snedeker, When cars hit trucks and girls hug boys: The effect of animacy on word order in gestural language creation. *Cogn. Sci.* **42**, 918–938 (2018).
26. V. Kempe, N. Gauvrit, D. Forsyth, Structure emerges faster during cultural transmission in children than in adults. *Cognition* **136**, 247–254 (2015).
27. N. Fay, T. M. Ellison, The cultural evolution of human communication systems in different sized populations: Usability trumps learnability. *PLoS One* **8**, e71781 (2013).
28. S. T. Piantadosi, H. Tily, E. Gibson, The communicative function of ambiguity in language. *Cognition* **122**, 280–291 (2012).
29. M. Bohn, B. Köymen, Common ground and development. *Child Dev. Perspect.* **12**, 104–108 (2018).
30. H. H. Clark, *Using Language* (Cambridge University Press, Cambridge, MA, 1996).
31. S. Goldin-Meadow, The impact of time on predicate forms in the manual modality: Signers, homesigners, and silent gesturers. *Top. Cogn. Sci.* **7**, 169–184 (2015).
32. J. Perner, H. Wimmer, "John thinks that Mary thinks that…" attribution of second-order beliefs by 5- to 10-year-old children. *J. Exp. Child Psychol.* **39**, 437–471 (1985).
33. B. Köymen, M. Tomasello, Children's meta-talk in their collaborative decision making with peers. *J. Exp. Child Psychol.* **166**, 549–566 (2018).
34. S. Göckeritz, M. F. Schmidt, M. Tomasello, Young children's creation and transmission of social norms. *Cogn. Dev.* **30**, 81–95 (2014).
35. R Core Team, R: A Language and Environment for Statistical Computing (Version 3.5.1, R Foundation for Statistical Computing, Vienna, 2017). https://www.R-project.org/.
36. D. Bates, M. Mächler, B. Bolker, S. Walker, Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).
37. A. J. Dobson, A. Barnett, *An Introduction to Generalized Linear Models* (CRC Press, Boca Raton, FL 2008).