

## COGNITIVE NEUROSCIENCE

# Versatile use of chimpanzee call combinations promotes meaning expansion

Cédric Girard-Buttoz<sup>1,2,3,4,\*</sup>, Christof Neumann<sup>5,6,7</sup>, Tatiana Bortolato<sup>2,3,4</sup>, Emiliano Zaccarella<sup>8</sup>, Angela D. Friederici<sup>8</sup>, Roman M. Wittig<sup>2,3,4</sup>, Catherine Crockford<sup>2,3,4\*</sup>

Language is a combinatorial communication system able to generate an infinite number of meanings. Nonhuman animals use several combinatorial mechanisms to expand meanings, but maximum one mechanism is reported per species, suggesting an evolutionary leap to human language. We tested whether chimpanzees use several meaning-expanding mechanisms. We recorded 4323 utterances in 53 wild chimpanzees and compared the events in which chimpanzees emitted two-call vocal combinations (bigrams) with those eliciting the component calls. Examining 16 bigrams, we found four combinatorial mechanisms whereby bigram meanings were or were not derived from the meaning of their parts—compositional or noncompositional combinations, respectively. Chimpanzees used each mechanism in several bigrams across a wide range of daily events. This combinatorial system allows encoding many more meanings than there are call types. Such a system in nonhuman animals has never been documented and may be transitional between rudimentary systems and open-ended systems like human language.

## INTRODUCTION

Language is a highly versatile communication system whereby a small repertoire of vocal elements, phonemes, is combined into meaningful words that are themselves combined into hierarchically structured sentences (1–4). Through this dual patterning, language is a powerful open-ended communication system, unique in the animal kingdom (4–6). This uniqueness poses a conundrum for identifying the origins of language (7–9). One particular feature of language is our ability to use numerous combinatorial mechanisms, that is, mechanisms through which the meaning of single vocal units, such as words or animal calls, is altered when combined into a sequence. For language, combinatorial mechanisms fall into two main categories. Semantically noncompositional combinations refer to those combinations where the overall meaning is not derived from the meaning of the composing units, such as forming idioms, like “go ape.” In contrast, in semantically compositional combinations, the meaning is derived from the meaning of the composing units such as in “the ape goes.” In this study, we use the term “meaning” to refer to the “signal meaning” facet of the multifaceted concept that is “meaning” (10), i.e., the information content of a signal.

Comparative studies aimed at retracing the evolutionary origins of language have focused on assessing the meaning and structure of non-human animal call combinations to determine their compositional

nature (11–15). We define call combinations as vocal sequences that combine meaning-bearing vocal units.

Both noncompositional and compositional structures have been reported in call combinations: Noncompositional combinations convey new meanings unrelated to the meaning of either composing unit, as in putty-nosed monkeys that combine two alarm calls into a vocal sequence that elicits travel (14). In contrast, compositional combinations convey meanings related to the meanings of the composing units. Two different compositional structures have been identified and demonstrated using rigorous playback experiments. First, affixation occurs when the meaning of the combination is similar to only one of the two units, with the other unit slightly modifying the meaning of the first unit. Campbell’s monkeys offer an example where the suffix “oo” is appended to alarm calls “krak” and “hok” to decrease the urgency of the message (15). Second, meaning maintenance occurs when the meaning of the combination retains the meaning of each of the units composing it and those meanings are added together as in Japanese tits and pied babblers that combine alarm and recruitment calls into an alarm + recruit sequence (12, 13).

A crucial feature of human language is syntax, which refers to the combination of words and the order in which they are arranged in a sentence, influencing meaning. In human language, syntax is characterized by a hierarchical structure where shorter sentences are often embedded in longer ones. While some species like chimpanzees are capable of producing hierarchically structured actions like when using tools (16), hierarchical structures have not yet been identified in animal vocal sequence productions. However, establishing whether the meaning of a combination is affected by simple ordering effects such as the order in which two units are combined is an essential step in investigating the evolutionary origins of syntax. To our knowledge, ordering effects on the meaning of combinations have only been found in two bird species (12, 13).

Animal call combinations thus feature several combinatorial mechanisms analogous to some mechanisms found in language. However, at most, one such mechanism has been identified per species, often in one or two call combinations only, and mostly

Copyright © 2025 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution NonCommercial License 4.0 (CC BY-NC).

<sup>1</sup>ENES Bioacoustics Research Laboratory, Centre de Recherche en Neurosciences de Lyon, CNRS, Inserm, University of Saint-Etienne, Saint-Etienne, France. <sup>2</sup>Institut des Sciences Cognitives Marc Jeannerod, CNRS, 67 Boulevard Pinel, 69675 BRON, Lyon, France. <sup>3</sup>Tai Chimpanzee Project, CSRS, Abidjan, Ivory Coast. <sup>4</sup>Department of Human Behaviour, Ecology and Culture, Max Planck Institute for Evolutionary Anthropology, 04103 Leipzig, Germany. <sup>5</sup>Cognitive Ethology Laboratory, German Primate Center - Leibniz Institute for Primate Research, Kellnerweg 4, 37077 Göttingen, Germany. <sup>6</sup>Department for Primate Cognition, Johann-Friedrich-Blumenbach Institute, Georg-August-Universität Göttingen, Kellnerweg 4, 37077 Göttingen, Germany. <sup>7</sup>Leibniz ScienceCampus, German Primate Center - Leibniz Institute for Primate Research, Kellnerweg 4, 37077 Göttingen, Germany. <sup>8</sup>Department of Neuropsychology, Max Planck Institute for Human Cognitive Sciences, 04103 Leipzig, Germany.

\*Corresponding author. Email: cedric\_girard@eva.mpg.de (C.G.-B); catherine.crockford@isc.cnrs.fr (C.C.)

constrained to events surrounding predator encounters. The question therefore remains how the evolutionary transition unfolded from the rudimentary combinatorial system of nonhuman animals to human language, which uses numerous combinatorial mechanisms to produce a virtually unlimited number of combinations to communicate about a wide variety of daily life events. Nowak *et al.* (17, 18) proposed a mathematical model for the evolution of communication systems that use numerous and diverse compound signals. They define compound signals as “signals that contain parts that have their own meaning.” Such a generalized combinatorial communication system can emerge if individual units occur in many different compound signals and the number of messages is substantially larger than the number of single units (17). Therefore, a vocal communication system in which (i) several constituent single vocal units are used across a broad range of vocal sequences and (ii) vocal sequences are used across a broad range of daily life events could offer a transitional vocal system bridging between the restricted use of vocal sequences found in animals and a complex and flexible combinatorial system, such as human language. We define “events” as the context or situation in which a vocalization was uttered, such as feed, nest, approach, aggress, or predator encounter. With respect to Nowak’s model, we postulate that such a transitional combinatorial system should also expand the range of messages that can be conveyed with single units by using combinatorial mechanisms that operate to modify the meanings of single units. Therefore, we add that this system would ideally (iii) use not one but several combinatorial mechanisms possibly overlapping with those found in language.

One potential candidate for such a transitional vocal system is the chimpanzee vocal repertoire. Chimpanzees have a vocal repertoire of 12 main single call types (single vocal units), with some variants for certain call types (19). Purely on a structural level, chimpanzees fulfill criterion (i) because they combine these 12 calls into hundreds of vocal sequences, and most single calls are used across a broad range of different vocal sequences (20). This system also fulfills criterion (ii) because the vocal sequences are used across a wide range of events and contain single calls, at least some of which are meaning bearing (21). Whether chimpanzees fulfill criterion (iii) remains to be broadly investigated because potential combinatorial mechanisms have only been assessed in a few vocal sequences. The extent of the capacity for meaning to be altered by single calls being incorporated into sequences thus remains to be assessed.

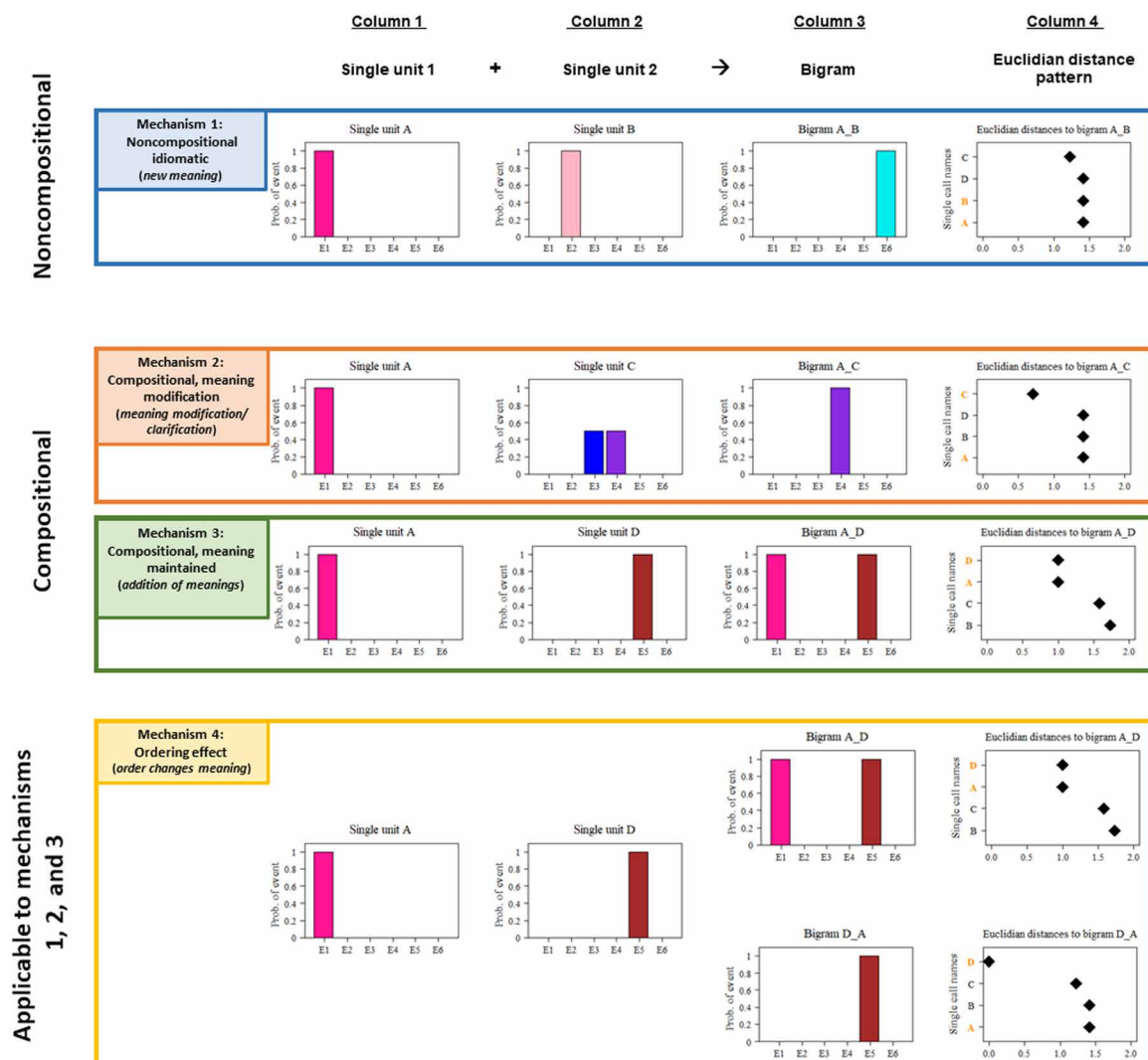
Several recent studies indicate considerable potential for chimpanzee vocal sequences to expand the range of meanings being conveyed. First, the diversity of sequences used is positively correlated with the diversity of events in which they are produced (21). Second, chimpanzees may utter compositional call combinations. Chimpanzees utter bigrams (two-call vocal sequences) that may combine information about self-identity and either food (pant-hoot + grunt sequence) (22) or subordination (panted-hoot + panted-grunt sequence) (23) or information about danger and recruitment (hoo + bark sequence) (11). Yet, while event specificity has been established for several of the call types when emitted singly (24–29), there is an overall lack of knowledge about event specificity and potential meanings of chimpanzee vocal sequences, beyond the three examples above. Such knowledge is critical for assessing versatility in the combinatorial mechanisms used [criterion (iii) above].

Our study aimed to determine whether and via which combinatorial mechanisms meaning can be expanded in chimpanzee vocal

sequences. We used the events during which single units and bigrams (vocal sequences composed of two different vocal units) were emitted as a proxy for the potential meaning of each utterance. It has been shown, in a broad range of taxa and across a large diversity of events, that single calls often convey event-specific information to conspecifics [e.g., predator presence (30–32), rest (24, 25, 27), aggression (33, 34), and copulation (35)]. Prior playback studies repeatedly demonstrated a good match between the event in which the call is typically emitted and the event-related information conveyed to receivers. However, there is an extensive debate that we will not address here about whether the signaler encodes the same meaning into a given utterance type that receivers extract from hearing it (36, 37). Nonetheless, it is reasonable to assume that the meaning of utterances, at least in part, relates to the events that elicit them. This is in keeping with a recent review on animal linguistics that defines meaning as “the set of features of circumstances that appear at a rate greater than chance across the signal’s occurrence” (38).

We considered four distinct combinatorial mechanisms that correspond to four scenarios via which the meaning of single units can be expanded (Fig. 1). These four mechanisms are based on the classification proposed for animal linguistics by Berthet *et al.* (38). We additionally provide a biological framework that focuses on the function of the combinatorial mechanisms. Accordingly, for each mechanism, we propose a terminology that explicitly states how the meaning of the single units is modified in the combination: (1) noncompositional idiomatic combinations (new meaning generation): The meaning of the call combination differs from the meaning of the composing units and of all other single vocal units in the repertoire; (2) compositional combinations with modification (meaning modification/clarification): The meaning of the call combination is similar to that of only one of the composing units with the second unit acting as a modifier/clarifier; (3) compositional combinations that maintain meaning (addition of meanings): The meaning of the vocal sequence reflects the meaning of each of the composing units; (4) ordering effect: The order in which the units are combined alters the meaning of the combination. Note that mechanism 4 can operate to alter the meaning of noncompositional and compositional combinations. Each of these mechanisms allows some form of expansion in the range of meanings that can be communicated from a limited repertoire of single calls. Mechanism 1 directly expands the range of meanings that can be conveyed. Mechanism 2 allows one to modify or nuance the meaning of the composing single units, such as to clarify socially complex situations [e.g., disambiguate signals and improve signal clarity (39, 40)]. Mechanism 3 enables communication about concurrently occurring events within a single utterance, such as communicating about self-identity and food (22), self-identity while greeting a conspecific (23), or warning conspecifics of a danger while concurrently recruiting them to the danger (11, 12). Mechanism 4, through changing unit order, allows further expansion of meanings that can be conveyed with the same vocal units.

We evaluated which combinatorial mechanisms could apply to chimpanzee vocal sequences by assessing the potential meaning of 11 of the 12 single vocal units (one unit was never produced alone in this sample) and of 16 commonly used bigrams. Our set of 16 bigrams included four pairs of bigrams in which each single unit occurred in either order (e.g., hoo + grunt versus grunt + hoo), allowing us to investigate ordering effects in relation to the potential meaning of bigrams. We focused on bigrams in our study for two reasons. (i) Unlike longer sequences, they occur frequently enough



**Fig. 1. Predictions for vocal bigram usage detailing scenarios for the four combinatorial mechanisms considered and the expected Euclidean distance patterns.**

This figure depicts the event distribution of four theoretical bigrams and the respective composing units when emitted alone, illustrating the four combinatorial mechanisms considered. E1 to E6 on the x axis refer to the six different events in which these theoretical vocal utterances could have been uttered. Each row depicts the events in which single units are uttered (columns 1 and 2) and the events in which the bigrams are uttered when constructed out of the two respective composing units (column 3). Column 4 represents the Euclidean distance measuring the difference in the event distribution between bigrams and their composing vocal units when emitted alone. Short distances (low values) indicate similarity in the events in which the bigram and the single units are uttered. Detailed descriptions of the theoretical examples can be found in the text. Example 1 in blue illustrates mechanism 1, a noncompositional idiomatic bigram that conveys a new meaning. Example 2 in orange illustrates mechanism 2, a compositional bigram with meaning modification that modifies/clarifies the meaning of one of the composing units. Example 3 in green illustrates mechanism 3, a compositional bigram with meaning maintained that adds the meaning of the two single units. Example 4, in yellow, illustrates mechanism 4, ordering effects, that is a meaning shift created by the bigram when the order of the composing units is reversed. Note that what is depicted here is only one possible example, but ordering effects can operate shifts between any of the three other mechanisms and also within the same mechanism (e.g., A\_B can create a new meaning and B\_A can create another new meaning or A\_B can disambiguate A while B\_A disambiguates B).

to quantitatively evaluate the events in which they are uttered, and (ii) most combinatorial mechanisms described in animals are examples of bigrams, which simplify drawing parallels between this and previous studies.

We assessed the similarity and differences of each bigram's potential meaning with the potential meaning of each single unit in the repertoire by calculating the Euclidean distances between the proportion of occasions in which a given bigram was uttered in each event and the proportion of occasions in which each single unit was uttered in each event. For each combinatorial mechanism, we predicted

different patterns of Euclidean distances between single units and bigrams (detailed in Fig. 1). Euclidean distances are conventionally used in experimental linguistics to evaluate the phonological similarity between word pairs (41, 42).

## RESULTS

### General protocol

We conducted our study on 53 wild chimpanzees above the age of 10 years from three communities (East, North, and South) at the Tai

National Park, Côte d'Ivoire (5°45'N, 7°07'W) within the Taï Chimpanzee project (43). Using continuous focal animal sampling, T.B. audio recorded vocalizations during 6-hour, half-day focals [see details in (20)], including those emitted by the focal subject or produced by individuals visible around the focal animal for whom the identity of the caller and the events associated with the call could be identified with certainty *ad libitum* (44). For each vocalization recorded, T.B. noted the events related to each vocalization following a preestablished ethogram [see details in Materials and Methods and in (21)]. T.B. assigned events to utterances by noting what was occurring during vocal emission, specifically the signaler's current (i) activity (feed, rest, and travel), (ii) social interaction (e.g., greet, play, groom, affiliate, aggress, and receive aggression), and (iii) changes in the environment (e.g., animal encounter, third-party aggression, and interparty communication).

In total, using these criteria, we identified 22 different events linked to vocalizations in this sample. Twenty of these events have been previously described as eliciting chimpanzee vocalizations [table S1, reviewed in (19)], and 12 events have been previously noted to elicit event-specific vocalizations (21) (see table S1 and Materials and Methods for a complete list of events and their definitions). Note that, as utterances can be emitted during more than one event unfolding sequentially or simultaneously (21), each respective event is associated with the respective utterance, such as affiliation during fusion that would result in two events: affiliation + fusion.

### Vocalizations studied

For this study, we focused on bigrams because they are the vocal sequences for which we have a large sample size, enabling us to match events to bigrams accurately. Furthermore, most combinatorial mechanisms in nonhuman animals have been described for bigrams, allowing us to draw direct parallels to those found in other taxa. Given that we are interested in assessing meaning shifts generated by chimpanzee vocal sequences that were more likely generic than idiosyncratic, we considered all bigrams that were produced by not one but several individuals. To balance this criterion with the need to obtain suitable sample sizes for analysis, we took a reasonable but arguably arbitrary threshold, including in our analysis all bigrams that were produced by at least 10 individuals. Note that the number of individuals effectively having each of the 16 bigrams in their repertoire is likely to be much larger than 10 (see Supplementary Methods and fig. S1). In a previous study, we applied randomization procedures and Bayesian binomial tests to assess the structure of chimpanzee bigrams in the same population (20). Several bigrams were produced above chance, single units showed positional bias within bigrams, and bigrams demonstrated some transitional bias between first and second units. This previous study shows that chimpanzee bigrams are structured and not random juxtaposition of calls. Randomization tests function to weed out signal sequences that are unlikely to demonstrate structured changes in meaning. Such tests used in (20) become irrelevant when contextual shifts are directly assessed, as they are here. With our cutoff of 10 individuals, we also ensured that we analyze bigrams that were present in the vocal communication of several individual chimpanzees, including bigrams recorded rarely that would not be detected as "above chance" not because they are random but because they are associated with rare events (e.g., animal encounter) or because they are associated with events during which recording vocalizations from individually identified chimpanzees is challenging (e.g., nesting). The 16 bigrams

studied together with their sample size are presented in Fig. 2A and were GR\_BK, GR\_HO, GR\_PG, GR\_PN, HO\_GR, HO\_PG, HO\_PH, HO\_PN, PG\_GR, PG\_PN, PH\_HO, PH\_PB, PH\_PG, PH\_PS, PH\_SC, and PN\_PG. Here and throughout the manuscript, the abbreviations are as follows: BK, bark; GR, grunt; HO, hoo; PN, pant; PB, panted bark; PG, panted grunt; PH, panted hoo; PS, panted scream; SC, scream. Therefore, we used the audio recordings in which chimpanzees emitted either 1 of these 16 bigrams or any of the 12 single call types (referred to as single units) comprised in the chimpanzee repertoire [see (45) and Materials and Methods for details] to compare the events in which single units and bigrams were emitted. This resulted in a dataset of 4323 vocal utterances comprising 3589 calls emitted singly (single units) and 734 bigrams. Because the single unit panted roar (PR) did not occur in an utterance alone, our final dataset comprised only 11 single units.

### Event distribution comparisons

The goal of our analysis was to systematize the relation between vocal utterance and event. We quantified similarities in the event distributions between different utterance types. We use the term "event distribution" to describe the probability for each event to be associated with a given utterance, either a single vocal unit or a bigram (e.g., in Fig. 1, single units A and B have different event distributions, while single unit A and the bigram D\_A have identical event distributions).

We modeled the observed event distribution for all single units and bigrams using a single multinomial model, which allows for a given observation (i.e., utterance) to be associated with one or more events (e.g., in mechanism 3 in Fig. 1, bigram A\_D is always associated with the same two events, while the single unit D is always associated with only one event). In this example, A\_D and D are two different utterance types. We fitted our model with utterance type and caller ID as varying intercepts. From this model, we extracted the estimated probability vectors, i.e., the estimated "event distributions," for all utterance types. From these estimated probability vectors, we calculated Euclidean distances as our measure of similarity of event distribution between all possible pairs of a target bigram with all single units. When we describe hereafter utterance types being far from or close to each other, we refer to these Euclidean distances (see specific examples below). In an extreme case, two utterances will have a distance of 0 if their estimated probability vectors are identical, i.e., the two utterance types have identical event distribution (see, e.g., mechanism 4 in Fig. 1).

The Euclidean distances were derived from posterior distributions of model parameters and can thus be represented as a posterior distribution. This allowed us to calculate the proportion of posterior samples in which the event distribution of a given single unit was the closest to that of a target bigram (from here on referred to as "P.support.single"). A single unit's event distribution is the closest to that of a bigram if its probability vector is the most similar to the probability vector of the bigram. We also calculated the proportion of posterior samples where both composing units are the two closest units to a target bigram in terms of Euclidean distances (from here on referred to as "P.support.pair").

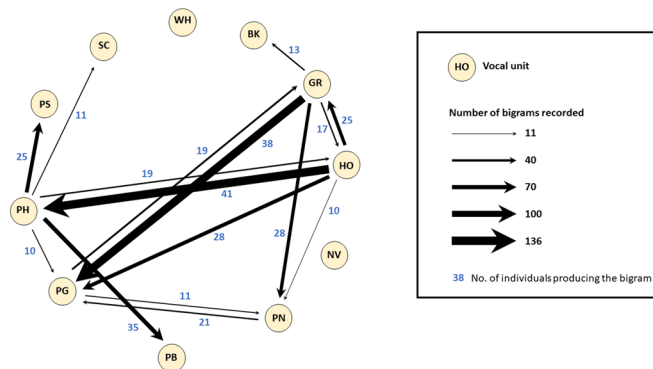
Shifts in event distributions from those of the composing units to those of the 16 bigrams allow us to infer what potential combinatorial mechanisms apply to each bigram. Figure 1 represents theoretical examples of the four combinatorial mechanisms we considered. It also shows the expected Euclidean distance pattern between bigrams



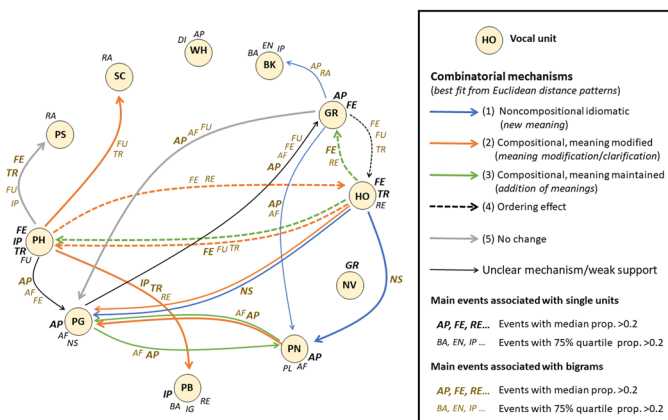
and their respective composing single units given each combinatorial mechanism.

The first mechanism is noncompositional idiomatic combinations (new meaning generation). Here, unit A (only uttered in event

## A Bigrams studied



## B Combinatorial mechanisms



**Fig. 2. Sixteen chimpanzee vocal bigrams studied and four combinatorial mechanisms that shape meaning changes.** (A) Sixteen bigrams studied with sample size indicated by the size of arrows. Blue numbers indicate the number of individuals producing each bigram. For both (A) and (B), the circles depict the different single vocal units of the chimpanzee vocal repertoire included in our sample. The arrows link the first and second vocal units of each bigram. (B) Combinatorial mechanisms that were best supported by the Euclidean distance patterns: mechanism 1 (blue), noncompositional combinations (new meaning); mechanism 2 (orange), compositional combinations with meaning modification (meaning modification/clarification); mechanism 3 (green), compositional combinations with meaning maintained (addition of meanings); mechanism 4 (dashed arrows), pairs of bigrams demonstrating ordering effect and no change scenario (gray). Thin black lines depict bigrams for which none of the mechanisms had a strong fit. Thin colored lines represent bigrams for which the mechanism was not strongly supported by Euclidean distance patterns. Two letter codes next to the circles in black indicate the events frequently associated with each single unit (i.e., with 75% quartile proportion >0.2). Bold indicates events that are strongly associated with the single unit (i.e., with a median proportion >0.2). Likewise, two-letter codes in brown along the arrows indicate the events frequently associated with the bigrams. Event codes: affiliation (AF), approach (AP), bystander aggression (BA), distress (DI), encounter (EN), feed (FE), fusion (FU), grooming (GR), intergroup encounter (IG), interparty communication (IP), nest (NS), play (PL), receive aggression (RA), rest (RE), and travel (TR).

E1) is combined with unit B (only uttered in event E2) to form a bigram A\_B that is only uttered in event 6 (an event during which single units are rarely uttered). For this mechanism, P.support.single and P.support.pair should be low because the Euclidean distance between A\_B and all single units is large and calls A and B are not closer to bigram A\_B than any other single units.

The second mechanism is compositional combinations with meaning modification (meaning modification/clarification of one of the single units). Here, the meaning of the single unit C is ambiguous because it is emitted sometimes in event E3 and sometimes in event E4. Unit A is combined with unit C to form the bigram A\_C, which is typically uttered during event E4. Here, the meaning of C is disambiguated by the bigram A\_C. For this mechanism, P.support.single should be high because the Euclidean distance between A\_C and one of the composing units (C) is smaller than the distance between A\_C and any of the other single units.

The third mechanism is compositional combinations that maintain meaning (addition of the meaning of the composing units). Here, the bigram A\_D is uttered in concomitant events E1 and E5, which are also the events in which the single units A and D are uttered, respectively. For this mechanism, P.support.pair should be high because the Euclidean distances between A\_D and each of the composing units (A and D) is smaller than the distance between A\_D and any other single units.

The fourth mechanism is the ordering effect (the order in which the units are combined in the bigram impacts meaning). Accordingly, the Euclidean distance patterns and event distribution patterns should differ between a bigram, like A\_D, and its reverse, D\_A. This mechanism differs from the other three in that it operates across bigrams and is thus depicted differently in Fig. 1. For example, while bigram A\_D is always uttered in concomitant events E1 and E5, the reversed bigram D\_A is only uttered in event E5. Accordingly, D\_A is much closer to D, while A\_D is equidistant from A and D. Note that ordering effects can produce a variety of shifts in event distributions between a bigram and its reverse, such as a partial change in the event distribution as illustrated here or a complete change. Ordering effects can apply to both noncompositional and compositional bigrams, and mechanism 4 can operate in combination with any of the other three mechanisms (mechanisms 1, 2, and 3).

Figure S2 shows a case of a “no-change” scenario, which is expected if no combinatorial mechanism operates. Here, the addition of unit A after unit D to form the bigram D\_A does not modify the meaning of D. In this case, the distance between D and the bigram D\_A is expected to be negligible and P.support.single to be very high.

## Meaning of the bigrams studied

Across the 16 bigrams studied, we found clear support for all the four mechanisms considered. Mechanisms 1, 2, and 4 were found in at least one bigram each for which, in each case, the Euclidean distance and the event distribution matched the expected patterns (summarized in Table 1). Mechanism 3, compositional combinations that maintain meaning (addition of meaning), was supported by Euclidean distances but with a less clear event distribution pattern. Hybrid mechanisms were evident for seven bigrams where the Euclidean and/or the event distribution patterns fitted both mechanisms 2 and 3, i.e., both types of compositional combinations, such that the meanings of the composing single units were both added and clarified within the bigram (Table 1; details below and in the Supplementary Materials). The combinatorial mechanisms best supported by

**Table 1. Summary of the best-supported combinatorial mechanisms for each chimpanzee vocal bigram in terms of Euclidean distance patterns and event distribution patterns.** Distance 1st and distance 2nd indicate the Euclidean distance between the event distributions of each bigram and the event distribution of the closest call and second closest call, respectively. Prop. support indicates the proportion of posterior distribution supporting the closest (column 3) and second closest (column 6) calls as the two closest calls to each bigram. Prop. support both composing calls are closest (column 8) indicates the proportion of posterior support for both composing calls to be the two closest calls. Euclidean distance pattern support (column 9) for the five combinatorial mechanisms is indicated with numbers corresponding to each mechanism in brackets, with bold indicating strong support and italics indicating weak support. Likewise, numbers in brackets (columns 3, 4, 6, 7, and 8) indicate strong (in bold) and weak (in italics) support for the following mechanisms: (1) mechanism 1, noncompositional idiomatic (new meaning); (2) mechanism 2, compositional with affixation (meaning modification/clarification); (3) mechanism 3, compositional, meaning maintained (addition of meanings); mechanism 4, ordering effect (this mechanism operates across bigrams, and paired bigrams with ordering effects on their potential meanings are shown with symbols & and §); (5) no change scenario. Cells with no numbers in brackets show no clear support for any mechanism. The matches to each mechanism are made on the basis of Euclidean distance patterns, as derived from the theoretical examples presented in Fig. 1. In the last two columns, we summarize separately the best-fitting mechanisms in terms of Euclidean distances and the best fitting mechanisms in terms of event distribution patterns based on inspection of the distribution of the proportion of the events associated with each single unit and each bigram.

Bigram	Closest call	Distance 1st	Prop. sup- port	Second closest	Distance 2nd	Prop. sup- port	Prop. sup- port both composing calls closest	Supported mechanisms	
								By Euclidean distances	By event distribution
HO_PN	PG	0.529 <b>(1)</b>	0.847 <b>(1)</b>	HO	0.635 <b>(1)</b>	0.034 <b>(1)</b>	0 <b>(1)</b>	<b>1</b>	<b>1</b>
GR_BK	HO	0.38 <i>(1)</i>	0.261 <i>(1)</i>	SC	0.387 <i>(1)</i>	0.31 <i>(1)</i>	0.002 <i>(1)</i>	<i>1</i>	<i>1</i>
GR_PN	PG	0.251	0.56 <i>(1)</i>	PN	0.264	0.439 <i>(1)</i>	0 <i>(1)</i>	<i>1</i>	2
HO_PG	PG	0.567 <b>(1)</b>	0.985 <b>(2)</b>	HO	0.687 <b>(1)</b>	0.004	0.682	<b>1 and 2</b>	<b>1 and 2</b>
PH_SC	PH	0.345 <i>(2)</i>	0.867 <i>(2)</i>	HO	0.411 <b>(2)</b>	0.099 <b>(2)</b>	0.02 <b>(2)</b>	2	<b>2</b>
PH_PB	PH	0.238 <i>(2)</i>	0.84 <i>(2)</i>	PB	0.322 <b>(2)</b>	0.142 <b>(2)</b>	0.575 <b>(2)</b>	2	<b>2</b>
PH_HO&	HO	0.252 <i>(2)</i>	0.821 <i>(2)</i>	PH	0.328 <b>(2)</b>	0.125 <b>(2)</b>	0.659 <b>(2)</b>	2	2 and 3
HO_PH&	PH	0.221 <b>(2)</b>	0.954 <b>(2)</b>	HO	0.292 <b>(2)</b>	0.046 <b>(2)</b>	0.994 <b>(3)</b>	<b>2 and 3</b>	2 and 3
PN_PG	PG	0.228 <b>(2)</b>	0.974 <b>(2)</b>	PN	0.377 <b>(2)</b>	0.024 <b>(2)</b>	0.928 <b>(3)</b>	<b>2 and 3</b>	<b>2 and 3</b>
HO_GR§	HO	0.24 <b>(3)</b>	0.672 <b>(3)</b>	GR	0.274 <b>(3)</b>	0.322 <b>(3)</b>	0.947 <b>(3)</b>	<b>3</b>	2 and 3
PG_PN	PN	0.332 <b>(3)</b>	0.482 <b>(3)</b>	PG	0.334 <b>(3)</b>	0.503 <b>(3)</b>	0.947 <b>(3)</b>	<b>3</b>	3
GR_PG	PG	0.174 <b>(5)</b>	0.999 <b>(5)</b>	PN	0.357 <b>(5)</b>	0 <b>(5)</b>	0.157 <b>(5)</b>	<b>5</b>	<b>5</b>
PH_PS	PH	0.169 <b>(5)</b>	1 <b>(5)</b>	HO	0.348 <b>(5)</b>	0 <b>(5)</b>	0.006 <b>(5)</b>	<b>5</b>	<b>5</b>
GR_HO§	HO	0.271	0.726	PH	0.315	0.244	0.174		2 and 3
PG_GR	PG	0.329	0.67	GR	0.37	0.296	0.629		2 and 3
PH_PG	PG	0.287	0.775	GR	0.396	0.147	0.017		2 and 3

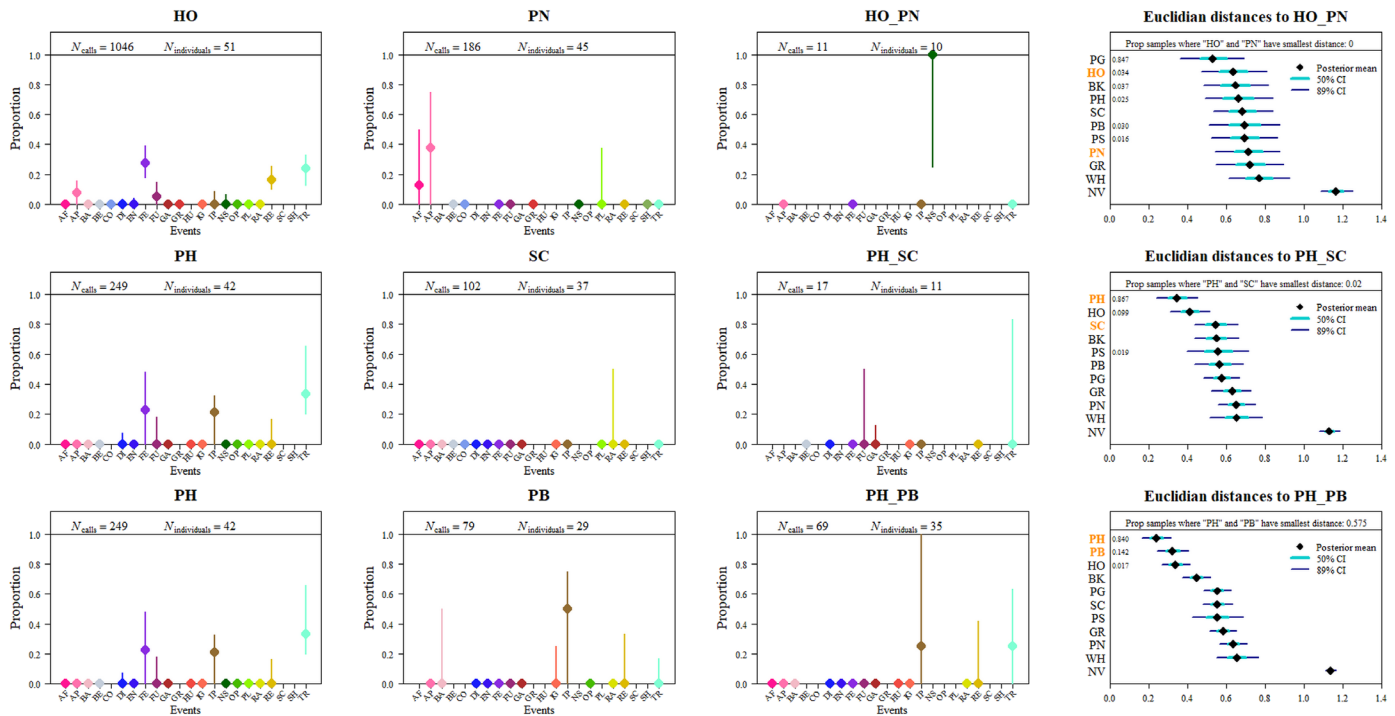
Euclidean distance patterns for each bigram and the main events in which each single unit and bigram were produced are summarized in Fig. 2B. In Table 2, we formalize the combinatorial rules governing changes in events between single calls and bigrams corresponding to the four combinatorial mechanisms that we found. To facilitate visibility, Figs. 3 and 4 depict only the empirical event (i.e., observed) proportion for each utterance type and do not depict the results from our statistical model. Estimated proportions from the statistical model are shown in the Supplementary Materials (figs. S7 to S9). We describe only one detailed example of one bigram per combinatorial mechanism in the main text and provide detailed descriptions for the other bigrams in the Supplementary Materials. For Figs. 3 and 4, we depict the median and 25 and 75% quartiles of the empirical proportion at which each utterance type was uttered. To simplify the narrative, in the rest of Results and Discussion, we considered an utterance to be frequently emitted during a given event if the 75% quartile was above 0.2. All the events frequently associated with each utterance type are indicated in Fig. 2B, and events in bold are those strongly associated with a given utterance (i.e., the median proportion is above 0.2). The median proportion of each event associated with each utterance in our dataset is 0.02, and we set the threshold of

0.2 at 10 times the median proportion. It is important to note however that, despite this arbitrary threshold, the proportion of each event for each utterance type and each individual chimpanzee was modeled and entered into our calculation of Euclidean distances and associated uncertainty.

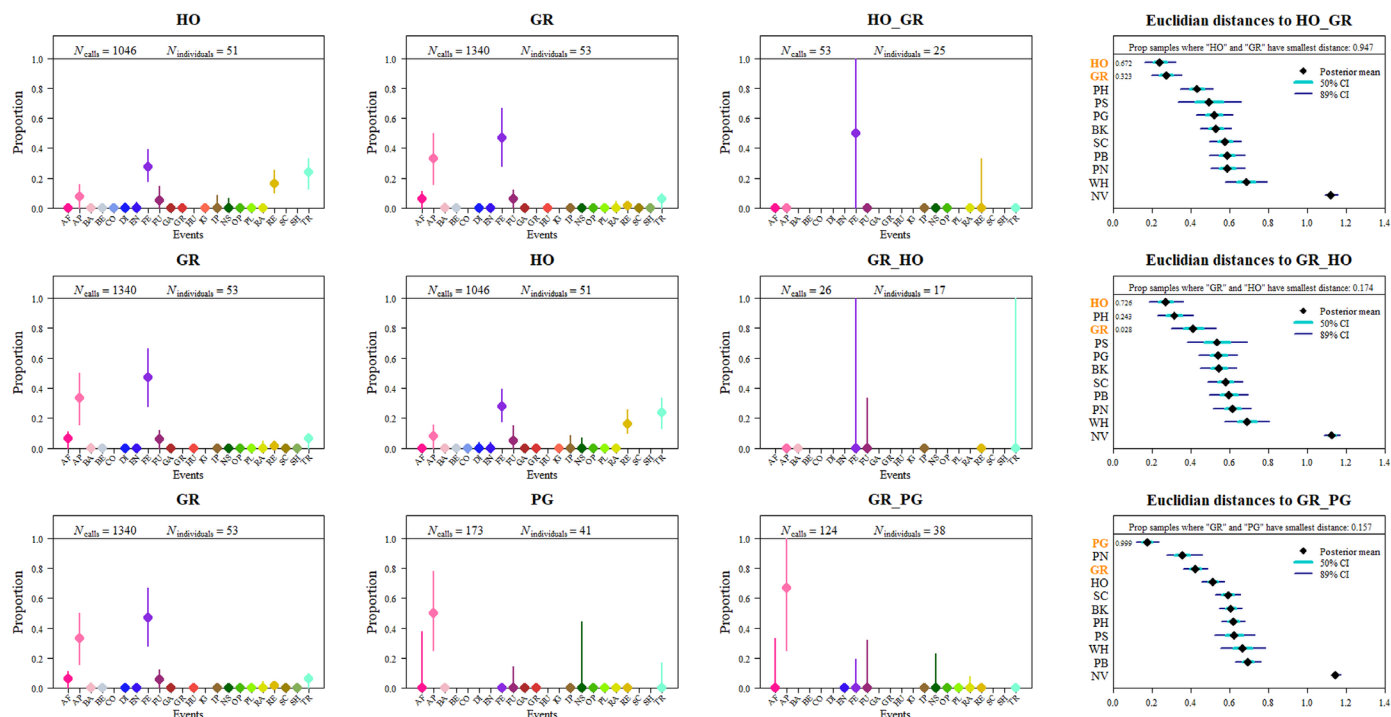
**Bigrams fitting mechanism 1: Noncompositional idiomatic combinations (new meaning)**

The bigram HO\_PN fulfilled all criteria for noncompositional idiomatic combinations that convey new meaning both in terms of event distribution and Euclidean distance patterns. The single unit HO was frequently uttered during feed and travel and occasionally during rest (Fig. 3 and fig. S7C), and the single unit PN was mostly uttered during approach, affiliation, and play (Fig. 3 and fig. S7H). In stark contrast, HO\_PN was overwhelmingly emitted during a single event, nest (Fig. 3 and fig. S8A). Single units of any type, including the composing single units, were rarely uttered during nest events. Accordingly, the Euclidean distance between HO\_PN and all other single units was large (>0.5; Table 1 and Fig. 3) and neither of the two composing single units was the closest unit to the bigram. Three other bigrams, HO\_PG, GR\_BK, and GR\_PN, also partially

Table 2. Combinatorial rules corresponding to each combinatorial mechanism in chimpanzee vocal bigrams, inferred from shifts in event usage. A, B, and C denote the events in which bigram A + B is emitted. A <sub>A</sub> , subset of events in which A is emitted.				
Combinatorial mechanisms	Rule for bigram A + B event usage	Refuting alternative 1: "only 1 expression"	Refuting alternative 2: "separate utterances"	Example bigrams
(1) Noncompositional/ idiomatic (new meaning)	A + B = C	A and B emitted as independent meaning-bearing units	C ≠ A, B, A + B	HO_PN, GR_BK
(2) Compositional: meaning clarification of A only	A + B = A <sub>A</sub>	As above	Meaning of B not found in A + B	PH_SC
(2 + 3) Compositional: meaning clarification of A & B	A + B = A <sub>A</sub> B <sub>B</sub>	As above	meaning A + B ≠ meaning B + A	PH-HO
(3) Compositional: maintained meaning of A & B	A + B = A + B	As above	meaning A + B ≠ meaning B + A	HO_GR
No change scenario	A + B = A or B			GR_HO



**Fig. 3. Chimpanzee vocal bigrams fitting mechanism 1, new meaning (HO\_PN), and mechanism 2, disambiguation (PH\_SC and PH\_PB).** Each row depicts the results for a given bigram. Columns 1 to 3 depict the event distributions of the first and second composing vocal units (columns 1 and 2) and of the bigram (column 3). Column 4 depicts the Euclidean distances measuring the difference in the event distribution between the bigrams and their composing vocal units when emitted alone. For the event distributions (columns 1 to 3), the colored dots depict the observed median individual proportion of occurrence of each call in each event and the vertical colored lines the 25 and 75% quartiles. A 25% quartile at a proportion *P* indicates that 75% of the individuals uttered that utterance in that event in a proportion  $\geq P$ . Likewise, a median at a proportion *P* indicates that half the individuals uttered that utterance in that event in a proportion  $\geq P$  and a 75% quartile at a proportion *P* indicates that 25% of the individuals uttered that utterance in that event in a proportion  $\geq P$ . For the Euclidean distances (column 4), the mean and 50 and 89% credible intervals for the distance of each bigram to each single call are depicted. The single units composing each bigram are indicated in orange. The values next to the single unit's name indicate the proportion of the sample from the posterior distribution of the model where each unit is the closest to the considered bigram (P.support.single). Only values above 0.01 (i.e., >1% support) are indicated. Indicated at the top of each Euclidean distance graph is the proportion of the sample from the posterior where the composing single units are the two closest single units to the respective bigram (P.support.pair).



**Fig. 4. Chimpanzee vocal bigrams fitting mechanism 3, combined meaning (HO\_GR), 4, ordering effect (HO\_GR versus GR\_HO), and 5, no change (GR\_PG).** Each row depicts the results for a given bigram. Columns 1 to 3 depict the event distributions of the first and second composing vocal units (columns 1 and 2) and of the bigram (column 3). Column 4 depicts the Euclidean distances measuring the difference in the event distribution between the bigrams and their composing vocal units when emitted alone. For the event distributions (columns 1 to 3), the colored dots depict the observed median individual proportion of occurrence of each call in each event and the vertical colored lines the 25 and 75% quartiles. A 25% quartile at a proportion  $P$  indicates that 75% of the individuals uttered that utterance in that event in a proportion  $\geq P$ . Likewise, a median at a proportion  $P$  indicates that half the individuals uttered that utterance in that event in a proportion  $\geq P$  and a 75% quartile at a proportion  $P$  indicates that 25% of the individuals uttered that utterance in that event in a proportion  $\geq P$ . For the Euclidean distances (column 4), the mean and 50 and 89% credible intervals for the distance of each bigram to each single call are depicted. The single units composing each bigram are indicated in orange. The values next to the single unit's name indicate the proportion of the sample from the posterior distribution of the model where each unit is the closest to the considered bigram ( $P$ .support.single). Only values above 0.01 (i.e.,  $>1\%$  support) are indicated. Indicated at the top of each Euclidean distance graph is the proportion of the sample from the posterior where the composing single units are the two closest single units to the respective bigram ( $P$ .support.pair).

fitted the criteria for noncompositional idiomatic combinations (Table 1; details in the Supplementary Materials and figs. S3 and S8).

### Bigrams fitting mechanism 2: Compositional combinations with modification (meaning modification/clarification)

The bigram PH\_SC offers a potential case of clarification of a single unit, PH. PH\_SC was predominantly emitted during fusion and travel, two events frequently associated with PH alone, but was rarely emitted during interparty communication or feed (Fig. 3 and fig. S7E), two other prominent events associated with PH alone (Fig. 3 and fig. S7G). SC, in contrast, was frequently emitted alone in receive aggression (Fig. 3 and fig. S7J), an event in which PH\_SC never occurred in our sample (Fig. 3 and fig. S7E). These patterns indicate that adding SC to PH shifted and limited the event usage of PH, such that the bigram PH\_SC predominantly emitted in travel events may convey to receivers that the signaler is traveling. In support of mechanism 2, the composing unit PH was the closest unit to the bigram PH\_SC with moderate support from the posterior ( $P$ .support.single = 0.867; Table 1 and Fig. 3), but the distance between PH and PH\_SC remained large (0.345; Table 1).

Two bigrams fitted a pattern indicating clarification of not only one but both single units in a hybrid compositional structure where

the meaning of both units is maintained but also clarified. For instance, PH\_PB presented Euclidean distance patterns that match the expected pattern for mechanism 2, clarification, because PH was the closest single unit to the bigram PH\_PB, with moderate posterior support ( $P$ .support.single = 0.840; Table 1 and Fig. 3), and PB was the second closest unit to PH\_PB (but with weak posterior support for both units being the closest,  $P$ .support.pair = 0.575). However, PH\_PB may clarify not one but both single units. This bigram was predominantly uttered during interparty communication, rest, and travel (Fig. 3 and fig. S8G). Two of these events are frequently associated with both composing single units PH and PB (travel and interparty communication; Fig. 3 and fig. S7). PH alone and PB alone were also often emitted in events virtually never associated with the bigram PH\_PB such as feed and fusion for PH and bystander aggression and intergroup encounter for PB (Fig. 3 and fig. S7G). The set of events in which PH\_PB was uttered was thus reduced in comparison to the respective set of either PH alone or PB alone. Thus, while the bigram PH\_PB is associated with certain events also associated with each composing unit, it also likely clarified the meanings of both.

We found another bigram that matched (at least partly) expected patterns for the disambiguation of both single units' scenario:



PH\_HO (Table 1; see details in the Supplementary Materials and fig. S4).

Previous studies have highlighted how event-related information can be encoded in the acoustic structure of each unit composing a vocal sequence, such as sequences containing an identity marker vocalization (46, 47). Here, we expand previous findings by showing that adding different single units to the chimpanzee identity marker, PH, to form bigrams also altered event specificity and, therefore, the potential meaning of the bigram. As already suggested for the bigram PH\_GR (22), other bigrams containing PH noted in this study could represent cases of predication where self-identity, communicated by the PH, is combined with other meaning-bearing units. Predication has likewise been suggested for bigrams that may function similarly in the banded mongoose and Diana monkeys (4, 22).

### Bigrams partially fitting mechanism 3: Compositional combinations that maintain meaning (addition of meaning)

Two bigrams fulfilled the criteria for compositional combinations that maintain meaning (addition of meaning) in terms of Euclidean distance patterns: HO\_GR and PG\_PN. For HO\_GR, the two composing single units (HO and GR) were clearly the two closest units to the respective bigram ( $P_{\text{support.pair}} = 0.947$ ; Fig. 4 and Table 1). However, HO\_GR was emitted during some but not all of the events frequently associated with its two composing units. While HO was frequently produced during feed, rest, and travel and GR during approach and feed (Fig. 4 and fig. S7), HO\_GR was mostly produced during one event associated with both composing units (feed) and, to a lesser extent, also with one event associated only with one composing unit (rest) (Fig. 4 and fig. S8C). HO\_GR may thus add and clarify the meaning of both composing units. Hoos emitted alone during rest prelude longer resting periods for both signalers and receivers (27), and chimpanzee grunts are associated with feed (48, 49). Further research should determine whether the bigram HO\_GR may similarly function to prelude longer stays in food patches than when grunts are emitted alone. We describe a second candidate bigram for mechanism 3, PG\_PN, in the Supplementary Materials and figs. S3 and S9 (see also Table 1).

### Bigrams fitting both mechanisms 2 and 3: Both forms of compositional combinations (meanings added and clarified)

Two combinatorial mechanisms may operate within a single bigram, such as in the case of HO\_PH and PN\_PG. These bigrams had Euclidean distance patterns that matched mechanism 2; as in each case, one of the composing units was clearly the closest unit to the respective bigram (PH for HO\_PH and PG for PN\_PG, both  $P_{\text{support.pair}} > 0.95$ ; Table 1). Each bigram also matched mechanism 3, given that both composing units were the closest to each respective bigram ( $P_{\text{support.pair}} > 0.92$  for both; Table 1). This co-occurrence of combinatorial mechanisms is also reflected in the event distribution (Table 1; detailed in the Supplementary Materials and figs. S4 and S9).

### Bigrams fitting mechanism 4: Ordering effect

Of the four pairs of bigrams in this sample that were emitted in both directions (A\_B and B\_A), two pairs demonstrated clear ordering effects such that the order in which the single vocal units were uttered within the bigram changed the events in which the bigram was emitted: the pair HO\_GR and GR\_HO and the pair HO\_PH and PH\_HO. Compared to HO\_GR (described above), GR\_HO was more often emitted during travel and fusion and less often during

rest (Fig. 4 and fig. S9D). Hence, GR\_HO was closer in distance to only one of the two composing units, HO, and not equidistant to GR and HO like HO\_GR. The distance to the other composing unit, GR, came in third position across all single units (Table 1 and Fig. 4). Furthermore, GR was the closest unit to GR\_HO in a very small proportion of the posterior ( $P_{\text{support.single}} = 0.028$ ; Table 1), which contrasts with the pattern of HO\_GR where HO and GR were similar to each other with respect to being the closest unit to the respective bigram (Table 1). For a detailed description of the ordering effect in the pair HO\_PH and PH\_HO (Table 1), see the Supplementary Materials and figs. S5 and S9.

### Bigrams showing no change in meaning

The bigram GR\_PG (Fig. 4) had a very similar event distribution to that of one of its composing units, PG. Both GR\_PG and PG were mostly produced during approach and to a lesser extent in affiliation and nest (Fig. 4 and fig. S9F). In addition, GR\_PG was also produced during fusion. Approach is an event frequently associated with the other single composing unit GR (Fig. 4 and fig. S7B). Yet, GR is also very frequently associated with feed, an event that is much less prominent in GR\_PG (Fig. 4 and fig. S9F). Consequently, GR\_PG was substantially closer to PG than to all other units ( $P_{\text{support.single}} = 0.999$ ; Table 1 and fig. S4) and the distance between GR\_PG and PG was small (0.173). Another bigram potentially conveys the same information as one of the single units: PH\_PS (Table 1; see details in the Supplementary Materials and figs. S5 and S9).

## DISCUSSION

Human language is a highly powerful communication system, not least through using compositional and noncompositional mechanisms to form combinations that expand the meaning capacity of the composing elements. Such mechanisms, together with hierarchically organized syntax, allow communication of a virtually infinite number of meanings based on a limited set of phonemes and words (9). Some nonhuman species have demonstrated the use of either one compositional or one noncompositional mechanism to produce meaningful call combinations but not more than one. Most examples relate to a single event, namely predator presence (11–13, 50). Hence, it remains puzzling how versatile combinatorial communication systems that are used across daily life events, like language, evolved. Our study changes this and highlights the potential for noncompositional and compositional combinations to be present within the communication system of a single nonhuman species, the chimpanzee, and to be used across a wide range of daily events. This theoretically offers a system with two meaning construction channels, such that when single calls are combined into bigrams, they can either generate a new meaning by creating noncompositional combinations or retain or modify their meaning, creating compositional combinations.

When mapping the events in which each call type (single unit) was uttered, we found a strong match with the call event specificity demonstrated in other studies. Examples include studies using playback experiments showing that some hoos are processed as rest calls (25) and observational studies showing that, like in our study, hoos are frequently emitted during rest and travel (24, 27, 51), panted grunts (PG) are uttered while approaching a conspecific (23, 28, 29), and grunts (GR) are predominantly uttered during feed (48, 49) or approaches (52) (see fig. S7 for further details). Our method thus

allows us to draw conclusions across the vocal repertoire about the potential meaning of single units and bigrams based on the events in which they are emitted.

When comparing the events in which bigrams and their composing units were uttered, we found bigrams that fitted the criteria for all four combinatorial mechanisms we considered in terms of Euclidean distance patterns (Fig. 2B), namely (1) noncompositional idiomatic (new meaning), (2) compositional sequences with modification (meaning modification/clarification), (3) compositional sequences that combine meaning (meaning combination), and (4) ordering effects. We also found several cases demonstrating a hybrid of compositional mechanisms 2 and 3 whereby the meanings of the composing units were both combined and clarified. Examples of each mechanism were evident in more than one bigram and involved at least three different single units (call types; Fig. 2B). Likewise, some single units were used in several mechanisms (Fig. 2B). These patterns suggest a versatile interactional system of mechanisms and call types across the chimpanzee repertoire. Last, we found two examples that fitted the null scenario, no meaning change. Together, our results indicate that most but not all bigrams potentially alter and/or combine the meanings of single units.

### Noncompositional bigrams

We found evidence of noncompositional bigrams fitting a pattern consistent with mechanism 1, conveying new meanings. For example, the bigram HO\_PN was almost exclusively produced during nest, an event that was rarely associated with either composing unit. This finding mirrors patterns reported for forest monkeys that combine two alarm calls to produce a call combination that initiates travel (14). While alarm and travel are very dissimilar meanings, one could argue that nesting consists of a long resting period and thus is close in meaning to the event rest that is frequently associated with the single unit HO. However, single HOs were almost never emitted in nest events, suggesting a contextual difference between rest and nest. Unlike rest, nesting is characterized by the building of a nest to spend the night, out of reach of night predators, like leopards (53). Furthermore, natural observations suggest that the production of the bigram HO\_PN elicits very distinct behavioral responses in receivers compared with the single unit HO (climbing a tree and building a nest rather than staying put on the ground). Accordingly, HO\_PN may have a very specific meaning that differs from that of single HOs.

### Compositional bigrams

For compositional bigrams, we found evidence for mechanism 2 that modification or clarification of the meaning of single units may occur in two ways. First, clarification can operate on only one of the composing units, such that the respective bigram was emitted in an overlapping but narrower set of events than the one of the composing units. This may be useful for units that are emitted in several contexts such as PH or GR.

Second, clarification can also operate not only on one but on both units within a bigram. In this case, the respective bigram was more likely emitted in a narrowed set of events shared by both composing units than in events pertaining to only one of the composing units (PH\_HO and PH\_PB).

The two clarification mechanisms of either one or both units in chimpanzee bigrams may be analogous to examples of affixation in other animals [e.g., (50)]. Yet, chimpanzee compositional structures

with modification appear to differ in three key aspects. First, the unit that acts to clarify the meaning of the other unit in the bigram is also a meaning-bearing unit when produced alone, whereas in other species like the Campbell's monkey "krak + oo" example (50), the affix "oo" is never produced alone. We suggest that, in chimpanzees, the same call type may have a dual function either as a meaning-bearing unit or as an affix. This idea would need to be tested. Second, whereas in the "krak-oo" example, the affix "oo" modifies the urgency of an already relatively specific alarm call (krak), in chimpanzees, the addition of an extra unit to form a bigram reduces the range of possible meanings of a given single unit, thereby clarifying its meaning. Such a system might be particularly adaptive in populations like Tai chimpanzees living in dense, low-visibility habitats. Chimpanzees use different variants of the same call type in different events (24), but the subtle difference in acoustic properties of the variants might be lost over distance. Thus, constructing a bigram that contains a clarifying component might provide useful contextual precision. Third, as far as we know, we report a new phenomenon where the meaning of both units can be clarified concurrently. We suggest that, in this case, each unit acts as a modifier for the other (Table 2).

In support of mechanism 3, we found two compositional bigrams that maintain the meanings of the composing units when emitted alone. In support of mechanism 4, ordering effects on event distributions were evident for two pairs of bigrams. In particular, HO\_GR demonstrated both features of combined meaning and ordering effects, making it a potential candidate for what has been labeled in other studies as "compositional syntax," whereby the meaning of a combination is determined by both the meaning of the composing units and the order in which they are combined. Yet, this needs to be formally demonstrated by future studies either using playback experiments or taking advantage of modern eye tracking (54) or touch screen (55) techniques, such as by designing violation of expectation paradigms that assess the information content of HO\_GR versus GR\_HO conveyed to receivers.

Clear cases of compositional syntax have only been reported so far for two bird species (12, 13), and only in alarm events, by combining an alarm call and a recruitment call into an "alarm-recruit" sequence. Likewise, the chimpanzee hoo\_bark (HO\_BK) combination is a suggested case of compositional syntax, but ordering effects here have not been tested (11). In our study, we do not strictly demonstrate that bigrams are compositional or noncompositional structures but rather show that chimpanzee call combinations likely encode meaning by using these mechanisms. Compositionality has been suggested but not tested for another chimpanzee bigram, the panted hoo (PH) + grunt (GR) (22) emitted to recruit conspecifics at food patches. Our study opens the possibility that chimpanzees liberally use both compositional and noncompositional call combinations in numerous daily life situations, not only in alarm events.

### Alternative explanations, implications, and limitations

Recently, Schlenker *et al.* (56) suggested that, besides the classic example of the Japanese tits (12), compositionality has not been shown in nonhuman animals (56). Schlenker *et al.* (56) propose two alternative "simpler" mechanisms, which could explain vocal combinations that appear compositional. First, the "only one expression" hypothesis suggests that the vocal sequence AB is a single stand-alone expression with its own meaning if A and B are never emitted independently. Second, the "separate utterance" hypothesis posits

that A and B are separate utterances that are neither syntactically nor semantically combined but simply randomly occur together. Our results demonstrate that the “only one expression” hypothesis cannot account for chimpanzee combinatorial mechanisms. Most single calls (single units) are meaning bearing when emitted alone and can be flexibly combined with several different call types into several bigrams. These bigrams in turn engage different combinatorial mechanisms leading to a change in meaning (see Table 2 for details). Likewise, the “separate utterance” hypothesis may fall short when looking at pairs of bigrams (PH\_HO versus HO\_PH and HO\_GR versus GR\_HO) where the order in which the calls are combined alters the event distribution, and therefore, the potential meanings, of the bigrams (Table 2). If composing units were randomly associated entities treated as separate utterances, the order in which they are combined would not affect the meaning of the bigram.

Together, our study showcases the potential for high versatility in combinatorial mechanisms used in chimpanzee bigrams and, in particular, four ways meaning can be modified. The chimpanzee vocal system is likely more complex than exemplified here because we only studied 16 bigrams, and chimpanzees produce hundreds of different vocal sequences, combining up to eight different vocal units within the same sequence (20). Longer sequences than bigrams are frequently produced when more than two events are concomitant (21). Thus, even though sequences with three or more units are rarer than bigrams, they are likely to further expand the communication potential of chimpanzees, especially because the number of single and concomitant daily life scenarios that elicit vocalizations [131 reported so far (21)] is considerably more than the 12 single call types. Given the diversity of structured vocal sequences found in chimpanzees and their potential to markedly increase the range of meanings, as shown in our study, we argue that chimpanzees fulfill all the criteria defined by Nowak *et al.* (17) for a general combinatorial system to evolve.

We did not use playback experiments to infer the potential meanings of the bigrams and are therefore agnostic as to how the information is perceived by the receiver. However, previous studies have shown a very good match between the event during which a call is emitted and the meaning perceived by the receiver in playbacks (24, 27, 30–34), even though not all calls emitted pertain to current events. For instance, in bonobos, the addition of a whistle before a high hoot is associated with a high likelihood for the caller to join another party in the future (57). Such call or call combination functions will not be captured by our methods because of strict assignment of utterances to current events, likely leading to an underrepresentation of influential events. The advantage of our probabilistic approach using Euclidean distances is that we can draw conclusions about potential combinatorial mechanisms across a large sample of call combinations, given the practical impossibility of testing all possible bigrams with playbacks.

## Conclusions

Our study suggests a highly versatile combinatorial communication system, with signal combinations consistent with the criteria of both compositional and noncompositional mechanisms, such that meanings may be generated, modified, or added together. Critically, the latter opens the possibility that some utterances may contain several meanings. Such a system offers considerable potential for meaning generation via the use of call combinations in a nonhuman animal, the chimpanzee. So far, at most, one meaning

modification mechanism has been reported per nonhuman species and, often in a single context, alarm. In contrast, chimpanzees not only appear to use diverse combinatorial mechanisms in their call combinations, but they use such mechanisms in a wide range of daily life events, not only alarm events. In a previous study, 46 of 58 chimpanzee bigrams were also recombined into trigrams (20) and likely longer utterances. The next step will be to determine which mechanisms are used when bigrams are recombined into longer utterances. To our knowledge, such a versatile use of combinatorial mechanisms as the one found in chimpanzees has not yet been described in a nonhuman animal species. Similar studies on other species will determine whether this reflects the extreme rarity of such signaling systems across the animal kingdom or rather reflects limited research effort in this area. We encourage future studies to assess in other species the usage of several combinatorial mechanisms in vocal and gestural signaling (58, 59), a prerequisite for which will be species that use a range of signal combinations across a range of daily life events. Combined with previous findings, the results of the current study show that chimpanzees fulfill the two criteria put forward in the evolutionary model of generalized combinatoriality by Nowak *et al.* (17). First, call combinations are used across a broad range of events and, second, single units are used across a large diversity of combinations. Nowak *et al.* (17) highlight that, upon fulfilling these criteria, call combinations should expand the range of messages that can be communicated. Here, we show that such an expansion can be aided particularly when several different combinatorial mechanisms operate within the same communication system, a criterion to add to those formulated by Nowak *et al.* (17). Such a communication system as we find here could be viewed as transitional between the more restricted combinatorial systems described so far in nonhuman animals and the extremely versatile and open-ended system that is human language.

## MATERIALS AND METHODS

### Ethics statement

Our study was purely observational and noninvasive. Observers followed the strict hygiene protocol of Tai Chimpanzee Project, which was adopted by International Union for Conservation of Nature as the best practice guideline for wild ape studies (60). Observers quarantined for 5 days before following the chimpanzees. During follows, observers disinfected their hands and boots and changed clothes before leaving and entering camps. In the forest, observers wore face masks and kept a minimum distance of 7 m between themselves and the chimpanzees to avoid disease transmission from humans to chimpanzees and to avoid disturbing the natural behavior of the observed individuals. The research presented here was approved by the “Ethikrat” of the Max Planck Society on 8 April 2018.

### Study subjects

T.B. collected data on 53 chimpanzees from three fully habituated communities (East, North, and South), Tai National Park, Ivory Coast, between two study periods: January to May 2019 and December 2019 to March 2020. We restricted our study to individuals 10 years and older because 10 is the age at which chimpanzees use the full adult repertoire of vocal sequences (45). T.B. collected data on 10 males and 12 females in the East group, 4 males and 8 females in the North group, and 5 males and 14 females in the South group.

## Data collection

T.B. followed chimpanzees from dawn to dusk during ~12 hours per day. T.B. recorded vocalizations during 6-hour-long half-day focal animal samples (44) [see details in (20)]. Using a 2-s prerecord option, T.B. audio recorded each vocalization from the focal chimpanzee as well as any vocalization produced by individuals visible around the focal for whom the identity of the caller could be identified with certainty ad libitum (44). T.B. recorded the vocalizations using a Sennheiser ME67 directional microphone (digitized at a 48-kHz sampling rate and 24-bit sampling depth) connected to a Tascam DR-40X digital recorder. For each vocalization recorded, T.B. noted the events related to the vocalization directly in the field following each event according to a preestablished ethogram, either on a smartphone using Cybertracker software (61) or via a vocal comment at the end of each recording. T.B. recorded what the caller was doing immediately before and at the time that each vocalization was emitted, what the caller was looking at while vocalizing, and the acoustic stimuli the caller were exposed to immediately before vocalizing. This led to the categorization of three kinds of events: the activity of the focal individual, social interaction directly involving the focal individual, and external events such as fusions or animal encounters occurring while vocalizing. In total, we defined 22 different events known to elicit vocalizations in chimpanzees and belonging to three categories (see table S1 for a complete list of events and their definitions): activities (feed, rest, and travel), social interactions (affiliate, approach, play, receive aggression, groom, beg, give aggression, food share, copulate, and solicit copulation), and changes in the environment (animal encounter, bystander to aggression, distress, hunt, intergroup encounter, fusion, interparty communication, nest time, and outside party noise).

We only considered the activity in which the focal animal was engaged in at the time of vocalizing as an event potentially triggering vocalization when it did not co-occur with any social interaction. T.B. focused on 53 chimpanzees 10 years and older for 646 hours and collected ad libitum data for an additional 387.8 hours. Overall, T.B. recorded 5399 vocal utterances with a mean  $\pm$  SE of  $101.9 \pm 7.41$  vocal utterances per chimpanzee. We did not include vocal sequences consisting of more than two different call types in this analysis.

## Vocal repertoire

For this study, we divided the chimpanzee vocal repertoire into 12 single call types following (20). A chimpanzee call type is defined as either single or repeated strings of a particular exhaled vocal element (e.g., “hoo” and “bark”) or “panted” forms of the same unit whereby a voiced inhalation is inserted between each exhaled element. Accordingly, our repertoire comprised seven single vocal units—bark (BK), grunt (GR), hoo (HO), pant (PN), scream (SC), whimper (WH), and nonvocal calls (NV)—and five panted forms of calls—panted bark (PB), panted grunt (PG), panted hoo (PH), panted roar (PR), and panted scream (PS) [see details in (20) for the call definition and classification]. In our sample, all single call types have been uttered singly with the exception of PR that only appeared in sequences. Accordingly, we could not compare event distribution between the study bigram and single unit PR and we excluded PR from the analysis, resulting in analysis of the event distribution of only 11 single units. As in (20), we defined a single utterance as a unit emitted alone or repeated within 2-s intervals. We defined a sequence as different types of vocal units emitted within less than

a 1-s interval. Different variants of the same call type [e.g., “rest” or “alert” hoo in (25)] are not differentiated here but are considered as the same unit type.

## Assigning units to recorded vocal sequences

We examined each recording using PRAAT spectrograms, which show the frequency distribution across the call (62), and determined each single call type using a concurrent visual inspection of the spectrogram and listening of the audio file [see details in (20)]. Call types can be differentiated because of their distinctive acoustic features (20). For the analysis, we considered only calls of high quality, with the lowest frequency band visible, recorded from the beginning to the end, and with the signaler ID defined. We did not include in the analysis utterances containing hard-to-identify calls. T.B. coded all the data. A total of 6% of the data (301 calls across all call types) was subjected to interrater reliability testing with a blind coder. At the end of the training, T.B. and the blind coder reached a 94.6% agreement on the call classification [see details in (21)].

## Bigrams studied

For this study, we restricted our analyses to bigrams (i.e., vocal sequences with two vocal units) commonly used by chimpanzees and utterances consisting of single vocal units. Specifically, we used all bigrams emitted by at least 10 different chimpanzees, resulting in the following 16 bigrams: GR\_BK, GR\_HO, GR\_PG, GR\_PN, HO\_GR, HO\_PG, HO\_PH, HO\_PN, PG\_GR, PG\_PN, PH\_HO, PH\_PB, PH\_PG, PH\_PS, PH\_SC, and PN\_PG (Fig. 2A). Of the 16 bigrams, 8 are composed of four pairs of calls in which the order of emission of the composing calls showed reversal (i.e., GR\_HO and HO\_GR, GR\_PG and PG\_GR, HO\_PH and PH\_HO, and PG\_PN and PN\_PG). These pairs allow us to assess the effect of call order on the events related to each bigram and, therefore, on the information conveyed by vocal sequences.

## Event analysis

Our analysis aimed to compare the distribution of the events that were associated with each single vocal unit with the events associated with each of the 16 bigrams. Thus, we used 4323 vocal utterances comprising 3589 single units and 734 bigrams. Each bigram was recorded in the dataset with a mean  $\pm$  SE of  $45.9 \pm 9.6$  (range: 11 to 136) occurrences.

Chimpanzee studies demonstrate that single calls can be event specific, such as nonvocal sounds being emitted almost exclusively during grooming (63), grunts emitted during feeding (48), and panted grunts emitted during greetings (28). Also, several acoustically different variants have been identified per call type, which each demonstrate event specificity, such as hoo variants during hidden threats, resting, and traveling (24, 26, 64). Here, we do not distinguish call variants, as this requires acoustic analysis. Hence, we expect that most call types will be associated with several events, for example, “hoos” will be associated with alarm, rest, and travel.

## Statistical analyses

We coded our multinomial model in Stan and fitted it using the “cmdstanr” interface in R (65, 66). We constructed a matrix with 22 columns representing events and 4323 rows representing observations. Each line was filled such that a value was 1 if the event was present when the utterance was produced and 0 if it was absent. In this model, we fitted the utterance type (i.e., the name of the single vocal unit or



the bigram emitted for a given utterance with 27 possibilities, 11 single vocal unit types, and 16 bigram types) and the identity of the caller as varying intercepts. We generated 4000 posterior draws (500 iterations in eight chains, with 3000 warmup samples in each chain).

To calculate the Euclidean distances, we extracted from the multinomial model the posterior distribution of the probability vector of each event to occur for each utterance type for an average individual (i.e., ignoring the individual-level intercept). For a given utterance, these estimated probabilities across the 22 events sum to 1. Crucially, however, 22% (933 of 4322) of utterances were observed to occur during concomitant events (e.g., approach + fusion + feed). To obtain event distributions that reflect this feature, we multiplied each probability vector with the expected number of events per recording of a given utterance type. For instance, if we obtained three recordings of utterance A, one during a feed event, one during concomitant feed + fusion events, and one during an approach event, we obtain a raw probability vector (which sums to 1) of  $[2 + 1 + 1]/4 = [0.5, 0.25, 0.25]$  (feed, fusion, approach). The expected number of events per recording would be  $(2 + 1 + 1)/3 = 1.33$ . We then multiplied the probability vector  $[0.5, 0.25, 0.25]$  by 1.33 to arrive at the desired probabilities for utterance A for each event:  $[0.66, 0.33, 0.33]$ . This reflects the fact that the feed event was recorded two times out of three recordings (prop. = 0.66) and fusion and approach one time out of three each (prop. = 0.33). It is these adjusted probability vectors for a given utterance that formed the basis for the calculation of the Euclidean distances.

Because the number of events might vary according to the utterance type, we also modeled the “event length” (the number of events in a recording). We used a zero-truncated Poisson likelihood, reflecting the fact that each recording comprised at least one event. As above, we fitted varying intercepts for utterance type and caller ID. Using the estimated intercept from this model for a given utterance type, we simulated 1000 recordings (where the event length was a random number drawn from a truncated Poisson distribution with the estimated mean for the current utterance type) and extracted which event(s) occurred. In each simulated recording, we calculated the adjusted probability vector for a given utterance type, as described above. The final result of these 1000 simulated recordings was then returned as a vector of means per event type. We did this for each utterance type, which then resulted in a set of 27 adjusted probability vectors. This set of adjusted vectors then served as a basis for calculating Euclidean distances between pairs of utterance types.

## Supplementary Materials

### This PDF file includes:

Supplementary Methods  
Supplementary Results  
Figs. S1 to S9  
Table S1

## REFERENCES AND NOTES

- G. K. Pullum, A. Zwicky, The syntax-phonology interface. *Linguist. Camb. Surv.* **1**, 255–280 (1988).
- P. Marler, “Animal communication and human language” in *Origin and Diversification of Language* (California Academy of Sciences, 1998), 19 pp.
- J. R. Hurford, *The Origins of Grammar: Language in the Light of Evolution II* (OUP Oxford, 2012).
- K. Collier, B. Bickel, C. P. van Schaik, M. B. Manser, S. W. Townsend, Language evolution: Syntax before phonology? *Proc. Biol. Sci.* **281**, 20140263 (2014).
- A. F. Russell, S. W. Townsend, Communication: Animal steps on the road to syntax? *Curr. Biol.* **27**, R753–R755 (2017).
- K. Zuberbühler, Evolutionary roads to syntax. *Anim. Behav.* **151**, 259–265 (2019).
- W. T. Fitch, M. D. Hauser, N. Chomsky, The evolution of the language faculty: Clarifications and implications. *Cognition* **97**, 179–210 (2005).
- J. J. Bolhuis, I. Tattersall, N. Chomsky, R. C. Berwick, How could language have evolved? *PLoS Biol.* **12**, e1001934 (2014).
- S. W. Townsend, S. Engesser, S. Stoll, K. Zuberbühler, B. Bickel, Compositionality in animals and humans. *PLOS Biol.* **16**, e2006425 (2018).
- J. Amphaeris, D. T. Blumstein, G. Shannon, T. Tenbrink, A. Kershenbaum, A multifaceted framework to establish the presence of meaning in non-human communication. *Biol. Rev. Camb. Philos. Soc.* **98**, 1887–1909 (2023).
- M. Leroux, A. M. Schel, C. Wilke, B. Chandia, K. Zuberbühler, K. E. Slocombe, S. W. Townsend, Call combinations and compositional processing in wild chimpanzees. *Nat. Commun.* **14**, 2225 (2023).
- T. N. Suzuki, D. Wheatcroft, M. Griesser, Experimental evidence for compositional syntax in bird calls. *Nat. Commun.* **7**, 10986 (2016).
- S. Engesser, A. R. Ridley, S. W. Townsend, Meaningful call combinations and compositional processing in the southern pied babbler. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 5976–5981 (2016).
- K. Arnold, K. Zuberbühler, Meaningful call combinations in a non-human primate. *Curr. Biol.* **18**, R202–R203 (2008).
- C. Coye, K. Ouattara, K. Zuberbühler, A. Lemasson, Suffixation influences receivers’ behaviour in non-human primates. *Proc. R. Soc. B* **282**, 20150265 (2015).
- E. Howard-Spink, M. Hayashi, T. Matsuzawa, D. Schofield, T. Gruber, D. Biro, Nonadjacent dependencies and sequential structure of chimpanzee action during a natural tool-use task. *PeerJ* **12**, e18484 (2024).
- M. A. Nowak, J. B. Plotkin, V. A. A. Jansen, The evolution of syntactic communication. *Nature* **404**, 495–498 (2000).
- M. A. Nowak, N. L. Komarova, Towards an evolutionary theory of language. *Trends Cogn. Sci.* **5**, 288–295 (2001).
- C. Crockford, “Why does the chimpanzee vocal repertoire remain poorly understood and what can be done about it?” in *The Chimpanzees of the Tai Forest: 40 Years of Research*, C. Boesch, R. M. Wittig, C. Crockford, L. Vigilant, T. Deschner, F. H. Leendertz, Eds. (Cambridge University Press, 2019), pp. 394–409.
- C. Girard-Buttoz, E. Zaccarella, T. Bortolato, A. D. Friederici, R. M. Wittig, C. Crockford, Chimpanzees produce diverse vocal sequences with ordered and recombinatorial properties. *Commun. Biol.* **5**, 410 (2022).
- T. Bortolato, A. D. Friederici, C. Girard-Buttoz, R. M. Wittig, C. Crockford, Chimpanzees show the capacity to communicate about concomitant daily life events. *iScience* **26**, 108090 (2023).
- M. Leroux, A. B. Bosshard, B. Chandia, A. Manser, K. Zuberbühler, S. W. Townsend, Chimpanzees combine pant hoots with food calls into larger structures. *Anim. Behav.* **179**, 41–50 (2021).
- C. Girard-Buttoz, T. Bortolato, M. Laporte, M. Grampp, K. Zuberbühler, R. M. Wittig, C. Crockford, Population-specific call order in chimpanzee greeting vocal sequences. *iScience* **25**, 104851 (2022).
- C. Crockford, T. Gruber, K. Zuberbühler, Chimpanzee quiet hoo variants differ according to context. *R. Soc. Open Sci.* **5**, 172066 (2018).
- C. Crockford, R. M. Wittig, K. Zuberbühler, Vocalizing in chimpanzees is influenced by social-cognitive processes. *Sci. Adv.* **3**, e1701742 (2017).
- C. Girard-Buttoz, M. Surbeck, L. Samuni, P. Tkaczynski, C. Boesch, B. Fruth, R. M. Wittig, G. Hohmann, C. Crockford, Information transfer efficiency differs in wild chimpanzees and bonobos, but not social cognition. *Proc. Biol. Sci.* **287**, 20200523 (2020).
- A. Bouchard, K. Zuberbühler, An intentional cohesion call in male chimpanzees of Budongo Forest. *Anim. Cogn.* **25**, 853–866 (2022).
- M. N. C. Laporte, K. Zuberbühler, Vocal greeting behaviour in wild chimpanzee females. *Anim. Behav.* **80**, 467–473 (2010).
- P. Fedurek, P. J. Tkaczynski, C. Hobaiter, K. Zuberbühler, R. M. Wittig, C. Crockford, The function of chimpanzee greeting calls is modulated by their acoustic variation. *Anim. Behav.* **174**, 279–289 (2021).
- R. M. Seyfarth, D. L. Cheney, P. Marler, Monkey responses to three different alarm calls: Evidence of predator classification and semantic communication. *Science* **210**, 801–803 (1980).
- S. A. Gill, S. G. Sealy, Functional reference in an alarm signal given during nest defence: Seet calls of yellow warblers denote brood-parasitic brown-headed cowbirds. *Behav. Ecol. Sociobiol.* **56**, 71–80 (2004).
- J. Kirchhof, K. Hammerschmidt, Functionally referential alarm calls in tamarins (*Saguinus fuscicollis* and *Saguinus mystax*) – Evidence from playback experiments. *Ethology* **112**, 346–354 (2006).
- T. J. Bergman, J. C. Beehner, D. L. Cheney, R. M. Seyfarth, Hierarchical classification by rank and kinship in baboons. *Science* **302**, 1234–1236 (2003).

34. R. M. Wittig, C. Crockford, K. E. Langergraber, K. Zuberbühler, Triadic social interactions operate across time: A field experiment with wild chimpanzees. *Proc. Biol. Sci.* **281**, 20133155 (2014).
35. C. Crockford, R. M. Wittig, R. M. Seyfarth, D. L. Cheney, Baboons eavesdrop to deduce mating opportunities. *Anim. Behav.* **73**, 885–890 (2007).
36. B. C. Wheeler, J. Fischer, Functionally referential signals: A promising paradigm whose time has passed. *Evol. Anthropol.* **21**, 195–205 (2012).
37. J. Fischer, T. Price, Meaning, intention, and inference in primate vocal communication. *Neurosci. Biobehav. Rev.* **82**, 22–31 (2017).
38. M. Berthet, C. Coye, G. Dezechache, J. Kuhn, Animal linguistics: A primer. *Biol. Rev. Camb. Philos. Soc.* **98**, 81–98 (2023).
39. M. Fröhlich, N. Bartolotta, C. Frys, C. Wagner, L. Momon, M. Jaffrezic, T. Mitra Setia, M. A. van Noordwijk, C. P. van Schaik, Multicomponent and multisensory communicative acts in orang-utans may serve different functions. *Commun. Biol.* **4**, 917 (2021).
40. M. Grampp, L. Samuni, C. Girard-Buttoz, J. León, K. Zuberbühler, P. Tkaczynski, R. M. Wittig, C. Crockford, Social uncertainty promotes signal complexity during approaches in wild chimpanzees (*Pan troglodytes verus*) and mangabeys (*Cercocebus atys atys*). *R. Soc. Open Sci.* **10**, 231073 (2023).
41. T. A. Farmer, M. H. Christiansen, P. Monaghan, Phonological typicality influences on-line sentence comprehension. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 12203–12208 (2006).
42. P. Monaghan, M. H. Christiansen, T. A. Farmer, S. A. Fitneva, Measures of phonological typicality: Robust coherence and psychological validity. *Ment. Lex.* **5**, 281–299 (2010).
43. R. M. Wittig, “Tai chimpanzees” in *Encyclopedia of Animal Cognition and Behavior*, J. Vonk, T. K. Shackelford, Eds. (Springer International Publishing, 2018).
44. J. Altmann, Observational study of behavior: Sampling methods. *Behaviour* **49**, 227–266 (1974).
45. T. Bortolato, R. Mundry, R. M. Wittig, C. Girard-Buttoz, C. Crockford, Slow development of vocal sequences through ontogeny in wild chimpanzees (*Pan troglodytes verus*). *Dev. Sci.* **26**, e13350 (2023).
46. H. Notman, D. Rendall, Contextual variation in chimpanzee pant hoots and its implications for referential communication. *Anim. Behav.* **70**, 177–190 (2005).
47. P. Fedurek, K. Zuberbühler, C. D. Dahl, Sequential information in a great ape utterance. *Sci. Rep.* **6**, 38226 (2016).
48. K. E. Slocombe, K. Zuberbühler, Functionally referential communication in a chimpanzee. *Curr. Biol.* **15**, 1779–1784 (2005).
49. A. M. Schel, Z. Machanda, S. W. Townsend, K. Zuberbühler, K. E. Slocombe, Chimpanzee food calls are directed at specific individuals. *Anim. Behav.* **86**, 955–965 (2013).
50. K. Ouattara, A. Lemasson, K. Zuberbühler, Campbell's monkeys use affixation to alter call meaning. *PLOS ONE* **4**, e7808 (2009).
51. T. Gruber, K. Zuberbühler, Vocal recruitment for joint travel in wild chimpanzees. *PLOS ONE* **8**, e76073 (2013).
52. J. Goodall, *The Chimpanzees of Gombe: Patterns of Behavior* (Harvard Univ. Press, 1986).
53. C. Boesch, The effects of leopard predation on grouping patterns in forest chimpanzees. *Behaviour* **117**, 220–241 (1991).
54. Y. Sato, F. Kano, N. Morimura, M. Tomonaga, S. Hirata, Chimpanzees (*Pan troglodytes*) exhibit gaze bias for snakes upon hearing alarm calls. *J. Comp. Psychol.* **136**, 44–53 (2022).
55. B. M. Waller, J. Whitehouse, J. Micheletta, Macaques can predict social outcomes from facial expressions. *Anim. Cogn.* **19**, 1031–1036 (2016).
56. P. Schlenker, C. Coye, M. Leroux, E. Chemla, The ABC-D of animal linguistics: Are syntax and compositionality for real? *Biol. Rev. Camb. Philos. Soc.* **98**, 1142–1159 (2023).
57. I. Schamberg, D. L. Cheney, Z. Clay, G. Hohmann, R. M. Seyfarth, Call combinations, vocal exchanges and interparty movement in wild bonobos. *Anim. Behav.* **122**, 109–116 (2016).
58. K. Liebal, J. Call, M. Tomasello, Use of gesture sequences in chimpanzees. *Am. J. Primatol.* **64**, 377–396 (2004).
59. E. Genty, R. W. Byrne, Why do gorillas make sequences of gestures. *Anim. Cogn.* **13**, 287–301 (2010).
60. K. V. Gilardi, T. R. Gillespie, F. H. Leendertz, E. J. Macfie, D. A. Travis, C. A. Whittier, E. A. Williamson, *Best Practice Guidelines for Health Monitoring and Disease Control in Great Ape Populations* (IUCN SSC Primate Specialist Group, 2015).
61. J. Steventon, Cybertracker v. 3.284 (2002).
62. P. Boersma, D. Weenink, Praat: Doing phonetics by computer (version 5.1.05) (2009) [retrieved 1 May 2009].
63. P. Fedurek, K. E. Slocombe, J. A. Hartel, K. Zuberbühler, Chimpanzee lip-smacking facilitates cooperative behaviour. *Sci. Rep.* **5**, 13460 (2015).
64. C. Crockford, R. M. Wittig, R. Mundry, K. Zuberbühler, Wild chimpanzees inform ignorant group members of danger. *Curr. Biol.* **22**, 142–146 (2012).
65. Stan Development Team, Stan Modeling Language Users Guide and Reference Manual, v. 2.33.1. (2023); <https://mc-stan.org>.
66. J. Gabry, R. Češnovar, A. Johnson, S. Bröder, cmdstanr: R Interface to “CmdStan” (2023); <https://mc-stan.org/cmdstanr/>.

**Acknowledgments:** We thank the Ministère de l'Enseignement Supérieur et de la Recherche Scientifique, the Ministère de Eaux et Forêts in Côte d'Ivoire, and the Office Ivoirien des Parcs et Réserves for permitting the study. We are grateful to the Centre Suisse de Recherches Scientifiques en Côte d'Ivoire for their logistical support and to K. Kolff and the staff members of the Tai Chimpanzee Project for support and assistance in collecting the data, and we are indebted to the late C. Boesch for establishing and nurturing the Tai Chimpanzee Project for 30 years. We thank D. Taylor for helpful comments on this manuscript. **Funding:** This study was funded by the Max Planck Society (M.F.EVAN8103 to C.C. and R.M.W. through the Evolution of Brain Connectivity Project) and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program awarded to C.C. (grant agreement no. 679787). C.N. was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) project ID 254142454/GRK 2070. **Author contributions:** Conceptualization: A.D.F., C.C., C.G.-B., E.Z., and R.M.W. Data curation: C.C., C.G.-B., C.N., R.M.W., and T.B. Formal analysis: C.G.-B. and C.N. Funding acquisition: A.D.F., C.C., and R.M.W. Investigation: T.B. Methodology: C.C., C.G.-B., C.N., E.Z., and R.M.W. Project administration: C.C., R.M.W., and C.G.-B. Resources: C.C., C.N., R.M.W., and T.B. Software: C.N. Supervision: C.C., R.M.W., and C.G.-B. Validation: C.C., C.G.-B., C.N., and T.B. Visualization: A.D.F., C.C., C.G.-B., and C.N. Writing—original draft: C.G.-B. Writing—review and editing: A.D.F., C.C., C.G.-B., C.N., E.Z., R.M.W., and T.B. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. The datasets generated during and/or analyzed during the current study are available at [https://github.com/tozbu/Chimpanzee\\_bigram\\_meaning](https://github.com/tozbu/Chimpanzee_bigram_meaning) and at <https://zenodo.org/records/13889876>. No custom computer code or algorithm was generated to analyze the data here. The R and STAN codes for the analysis are available at [https://github.com/tozbu/Chimpanzee\\_bigram\\_meaning](https://github.com/tozbu/Chimpanzee_bigram_meaning).

Submitted 8 May 2024

Accepted 8 April 2025

Published 9 May 2025

10.1126/sciadv.adq2879