



# Bare and Constructional Compositionality

David Gil<sup>1</sup>

Received: 6 July 2021 / Accepted: 24 November 2022  
© The Author(s) 2022

## Abstract

This paper proposes a typology of compositionality as manifest in human language and animal communication. At the heart of the typology is a distinction between *bare* compositionality, in which the meaning of a complex expression is determined solely by the meanings of its constituents, and *constructional* compositionality, in which the meaning of a complex expression is determined by the meanings of its constituents and also by various aspects of its structure. Bare and constructional compositionality may be observed in human language as well as in various animal communication systems, including primates and birds. Architecturally, bare compositionality provides the foundations for constructional compositionality, while phylogenetically, bare compositionality is a potential starting point for the evolution of constructional compositionality in animal communication and human language.

**Keywords** Compositionality · Bare compositionality · Constructional compositionality · Evolution of compositionality · Grammatical items · Configurational signs

## Introduction

All human languages make use of a set of simple signs, conventionalized pairings of forms and meanings. Nevertheless, however many such simple signs a language may have, its potential communicative needs remain far more numerous. In order to meet these needs, languages may combine two or more simple signs into a single more complex sign. However, for this to work, the meaning of the complex sign must be related to the meanings of its constituent parts, otherwise we wouldn't know what it means. This relationship is what is known as *compositionality*. A typical definition of compositionality is the following (Szabó, 2020):

---

Handling Editor: Joanna (Jo) M. Setchell

---

✉ David Gil  
gil@shh.mpg.de

<sup>1</sup> Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

(1) *Compositionality*

"The meaning of a complex expression is determined by its structure and the meanings of its constituents."

A crucial feature of the definition of compositionality is its *bipartite* nature. In accordance with the above definition, the meaning of a complex expression is determined by two factors: first, its structure, and second, the meaning of its constituents. Although the bipartite nature of compositionality is not explicit in Frege's original writings (e.g., Frege 1923: 55), it is a central feature of many or most subsequent definitions of compositionality (e.g., Partee, 1984: 281, Pagin, 2011: 52, Baggio et al., 2012: 656, Gleitman et al., 2012: 420, Hampton & Jönsson, 2012: 385, Löbner, 2012: 220, Werning 2012: 634, and Zimmermann, 2012: 82–83).

The bipartite nature of compositionality may be illustrated by the contrast between two simple sentences, one in Riau Indonesian (a colloquial variety of the national language of Indonesia spoken in east central parts of the island of Sumatra), the other in English. (Unless otherwise credited, all of the linguistic data cited in this paper is from the author's own fieldwork or general familiarity with the language in question.)

(2) *Riau Indonesian*

Ayam makan  
chicken eat

'Entity associated with chicken and eat' (e.g., 'The chicken is eating' / 'Some chickens have eaten' / 'The chicken is being eaten' / 'The chicken that is eating' etc. ...)

(3) *English*

The chicken is eating

Both sentences bring together a form that means 'chicken' and a form that means 'eat'; however, they do so in quite different ways.

Riau Indonesian permits the two words, *ayam* 'chicken' and *makan* 'eat' to be simply juxtaposed, without the addition of any further grammatical markers, either morphological or periphrastic. Accordingly, *Ayam makan* has an extremely wide range of possible interpretations, of which the ones exemplified in (2) are but a small sample. For example *ayam* 'chicken' may be understood as either singular or plural, and either definite or indefinite, while *makan* 'eat' may be associated with any combination of tense and aspect. Moreover, the thematic role that *makan* assigns to *ayam* is not specified; it could be agent, patient, or any other role that might make sense in a given context. Finally, the ontological category of the sentence as a whole is indeterminate; it may denote an activity (resulting in a clausal translation into English), a chicken (resulting in a phrasal relative-clause translation into English), or any other kind of entity that would make sense in the given context.

Riau Indonesian sentences such as *Ayam makan* are not multiply ambiguous, as one might perhaps be led to believe by their multiple possible translations into English; rather, they are endowed with a single vague or underspecified interpretation (Gil, 2001, 2005b, 2012, 2017 and elsewhere). Essentially, what *Ayam makan* is saying is

that there's a chicken and an eating, and the two are associated in some way; to the extent that greater specificity is called for, this can be filled in from the context of the utterance. The interpretation of *Ayam makan* may be represented in terms of the *Association Operator, A* (Gil, 2005b, 2012, 2017). In its monadic form, the Association Operator corresponds in its interpretation to familiar genitive or possessive constructions. For example,  $A(\text{JOHN})$  means 'entity associated with John', or simply *John's*, where the relationship between the associated entity and John is underspecified; thus, when modifying *book*, *John's book* could refer, depending on context, to the book that John owns, the book that John wrote, the book that is about John, and so forth. In its polyadic guise, in a formula such as  $A(X Y)$ , the resulting interpretation may be paraphrased as 'entity associated with X and Y'. In the formula  $A(X Y)$ , the order in which the two terms are listed is irrelevant; the relationship between them is purely symmetric. The interpretation of *Ayam makan* in (2) may thus be represented as  $A(\text{CHICKEN EAT})$ , or 'entity associated with chicken and eat'. The Association Operator is the sine qua non of semantic compositionality, reflecting the fact that whenever two or more signs are brought together, the result is a new sign whose meaning is built up from the meanings of the constituent signs.

(Notwithstanding a certain apparent affinity, the Association Operator differs in several respects from the Merge operator posited within the Minimalist Program, one of the current approaches to syntactic theory. First and foremost, while Merge is a syntactic operator that applies to forms (words and larger expressions), the Association Operator is a semantic operator that applies to meanings. Second, while Merge forms the basis for syntactic recursion, with the output of a Merge operation available as the input to a subsequent Merge operation, the Association Operator does not entail the presence of recursion, though it is consistent with it. Third, whereas in most versions of the theory, Merge is necessarily binary, the Association Operator is polyadic and may in principle apply to any number of arguments.)

Turning now to the English sentence in (3), quite obviously its meaning also has to do with chicken and eat: in this sense, the Association Operator and the formula  $A(\text{CHICKEN, EAT})$  also lie at the heart of its semantics, just as they do for Riau Indonesian (2). However, in comparison to (2), the semantics of (3) is more highly constrained; it cannot just mean anything related to chicken and eat. The greater semantic specificity of (3) correlates with its greater formal elaboration. Unlike Riau Indonesian, English does not generally permit the simple juxtaposition of forms such as *chicken* and *eat*. For a clause to be well-formed, additional grammatical markers must be present, such as, in (3), the definite article *the*, the combination of the auxiliary *be* and the suffix *-ing* marking progressivity, and the present tense form of the auxiliary, namely, *is*. The obligatory presence of grammatical markers such as these entails that the semantics of sentence (3) is narrower than just  $A(\text{CHICKEN, EAT})$ , in that it is specified also for semantic features such as number, definiteness, tense, aspect, thematic role, ontological category and others.

Thus, with reference to the definition of compositionality in (1), whereas the meaning of Riau Indonesian (2) is determined wholly by the meanings of its constituents, the meaning of English (3) is determined by the meaning of its constituents *in conjunction with* the grammatical construction within which these constituents occur. Unlike its Riau

Indonesian counterpart in (2), the meaning of the English sentence (3) thus reflects the bipartite nature of compositionality.

The distinction between (2) and (3) points towards a fundamental distinction between two types of compositionality, *bare* and *constructional*, which may be defined as follows:

(4) *Bare and constructional compositionality*

(a) *Bare compositionality*

The meaning of a complex expression is determined solely by the meanings of its constituents

(b) *Constructional compositionality*

The meaning of a complex expression is determined by the meanings of its constituents and also by its structure

Thus, while Riau Indonesian (2) instantiates bare compositionality, English (3) illustrates constructional compositionality.

In Szabó's definition of compositionality in (1), "structure" and "the meaning of [...] constituents" are listed as coordinate elements, seeming to imply that they are on a par with one another; however, this is actually not the case. Architectonically, the lexicon provides the foundations, while the grammar is the edifice that rests on top of those foundations. This is most clearly seen in the asymmetric distribution of the two: whereas grammar without lexicon is an impossibility, lexicon without an inventory of specific grammatical constructions offers an easily imaginable mode of expression attested in many different domains, namely bare compositionality.

This paper proposes a typology of compositionality distinguishing between bare and several kinds of constructional compositionality, and shows how the various types of compositionality are manifest in human languages and animal communication systems. The next section presents the typology in terms of an architecture of increasing complexity; the section after that illustrates the various types of compositionality with examples from human language; the following section provides an exploration of compositionality as has been suggested to exist in the realm of animal communication; and the final section offers some reflections regarding potential evolutionary trajectories involving different types of compositionality.

## A Typology of Compositionality

A typology of compositionality is presented in Table 1 below. The typology is intended to apply equally to human languages and the variegated communication systems of other animals. Its goal is to provide a conceptual framework for comparing human languages and animal communication systems, in order to identify commonalities and differences between the two domains, and in so doing, examine the extent to which it might be possible to bridge the gap between them.

**Table 1** A typology of compositionality

| <i>compositionality</i> |                   | <i>bare</i>            | <i>constructional</i>                      |  |
|-------------------------|-------------------|------------------------|--|--|
| <i>simple</i>           |                   | T1 A ( X Y )           | T2 A ( G <sup>0</sup> X )                  | T3 A ( C <sup>0</sup> X )                  |
| <i>relational</i>       | <i>monovalent</i> | *                      | T4 A <sub>rel</sub> ( G <sup>1</sup> X )   | T5 A <sub>rel</sub> ( C <sup>1</sup> X )   |
|                         | <i>bivalent</i>   | *                      | T6 A <sub>rel</sub> ( G <sup>2</sup> X Y ) | T7 A <sub>rel</sub> ( C <sup>2</sup> X Y ) |
|                         |                   | <i>non-grammatical</i> | <i>grammatical</i>                         | <i>configurational</i>                     |
|                         |                   | <i>material</i>        |  |  |

In Table 1, seven cells, numbered T1 to T7, represent seven distinct templates, or structures instantiating different kinds of compositionality. The primary distinction is between bare compositionality, in T1, and six distinct varieties of constructional compositionality, in T2–T7.

While all seven templates are based on the Association Operator, they vary with regard to several additional structural features. Much of the variation pertains to the nature of the items that are related via the Association Operator. In the case of bare compositionality, in T1, the Association Operator applies solely to meanings of a general nature, represented here with X and Y; such meanings are commonly characterized by linguists as "contentful" or "lexical". In contrast, in the case of constructional compositionality, in T2–T7, the Association Operator applies to combinations of such general meanings and two more particular kinds of meanings. The first consists of meanings associated with *grammatical items*, which, like "contentful"/"lexical" meanings previously, are also of a *material* nature, expressed though basic forms typically consisting of morphemes formed from combinations of phonemes; these are represented in T2, T4, and T6 with G. The second comprises meanings associated with various more abstract *configurational signs*, involving linear order, repetition, voice modulation, and so forth; these are represented in T3, T5, and T7 with C. The distinction between general meanings, grammatical items and configurational signs is fleshed out in detail in the next section.

A further distinction pertains to the valences of the grammatical and configurational meanings, indicated in Table 1 with superscripts. For simple compositionality, in T1–T3, the two associated elements stand in a symmetric relationship to one another, upholding the fundamentally symmetric nature of the Association Operator. In contrast, for relational compositionality, in T4–T7, the associated elements stand in an asymmetric relationship whose effect is to constrain the interpretation of the utterance beyond the associational relationship; this further constraint is represented with the subscript A<sub>rel</sub>. In the monovalent case, in T4 and T5, the items G<sup>1</sup> and C<sup>1</sup> are associated with a meaning that relates in some way to a single element X, while in the bivalent case, in T6 and T7, the items G<sup>2</sup> and C<sup>2</sup> are associated with a meaning that pertains in some fashion to the two elements X and Y. The distinction between simple, monovalent, and bivalent meanings is also discussed and illustrated in greater detail in the next section.

As suggested by the typology in Table 1, the architecture of compositionality involves a series of successive increments in complexity. This is represented diagrammatically in Fig. 1 below:

In Fig. 1, types of compositionality are indicated in small capitals, with reference to the seven templates of Table 1. Arrows represent increments in complexity, the defining feature of each increment indicated in italics beside the respective arrow.

Starting at the top of the diagram, in order for compositionality to get off the ground, constructions must contain multiple signs. The simplest form of compositionality is thus bare compositionality. The first enrichment of bare compositionality comes with the introduction of grammatical items, in the left-hand column of the diagram, or, alternatively, configurational signs in the right-hand column. Further complexification is then associated with the increase in valency of the grammatical items or configurational signs, first to monovalent, and subsequently to bivalent.

In the next two sections, we embark on a comparative exploration of compositionality in human language and animal communication, following the typology laid out in Table 1.

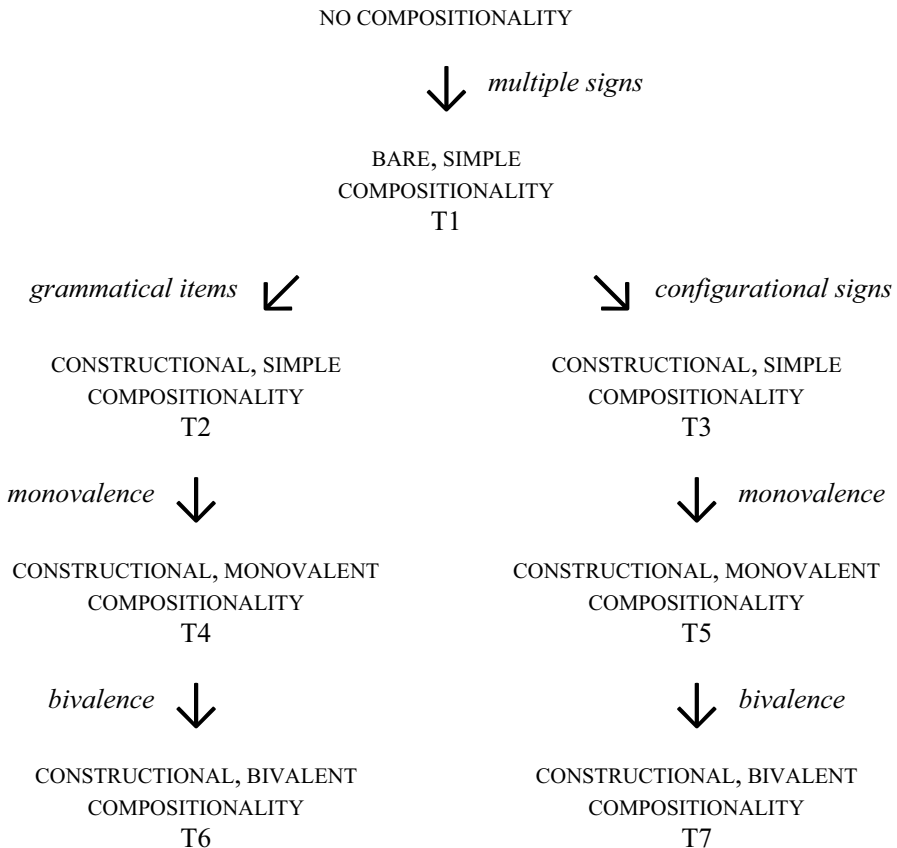


Fig. 1 The architecture of compositionality

## Compositionality in Human Language

The seven compositionality templates represented in Table 1 are illustrated in (5)–(28) below with examples from human languages. While many of these examples represent the kinds of constructions that grammarians typically concern themselves with, others involve a variety of phenomena that tend to fly under the radar of many linguists. However, in some cases at least, these latter constructions, involving features such as repetition and voice modulation, are the ones that appear, *prima facie*, to bear a closer resemblance to constructions in animal communication. Scholars of animal communication working their way through the details of the human-language examples are invited to try and consider which kinds of compositionality illustrated here are really specific to humans, and which others might also be present in various other animal communication systems; this is of course the question that is addressed in the following section below.

In several cases, the characterization of a construction as instantiating a certain kind of compositionality is dependent on a particular grammatical analysis; alternative and equally plausible analyses might yield different classifications. In general, the analyses assumed here tend to be concrete and rather superficial, eschewing alternative more abstract analyses that might reasonably be proposed, often within particular theoretical frameworks. One of the reasons for this approach is to present the human-language constructions in a way that would hopefully render them more amenable to comparisons with potentially similar constructions in animal communication systems. Still, the discussion in this section may, if so desired, be viewed not only as a typology of grammatical constructions but perhaps also as a typology of grammatical analyses.

One important respect in which the classification of various constructions in this section is analysis-dependent is in the assumption, implied in Table 1, that bare compositionality is mostly dyadic, applying to structures that are largely isomorphic to syntactic structures which are sometimes argued to be exclusively binary branching (e.g., Kayne, 1984). Alternative analyses of particular linguistic constructions might allow for compositionality in which the Association Operator applies to sets of more than two "contentful" or "lexical" meanings, for example A ( X, Y, Z ); such semantic structures might perhaps be suggested by syntactic analyses postulating "flat" or "non-configurational" structures (e.g., Hale, 1983 for Warlpiri et al. and Jackendoff and Wittenberg, 2014, 2017 more generally). The focus, in this paper, on dyadic structures is for ease of exposition only; no substantive matters hinge on this restriction.

Beginning with bare compositionality, as in T1, an example of this was already provided in Riau Indonesian (2); two additional examples are presented in (5) and (6) below:

### (5) *Riau Indonesian T1*

Aku pukul  
1SG hit  
A ( 1SG HIT )

'Entity associated with me and hit' (e.g., 'I hit (someone)' / '(Someone) hit me', etc. ...)

(6) *Riau Indonesian T1*

Makan tadi

eat PST.PROX

A ( EAT PST.PROX )

'Entity associated with eat and proximal past' (e.g., '(Someone) ate' / 'The person who was here just before is eating' / (Someone) is eating the thing referred to just before', etc. ...)

Examples (5) and (6) present the same structure as (2); their inclusion here gives a feel for the pervasiveness of bare compositionality in Riau Indonesian, in contrast to, say, English. In (5), the first person singular marker *aku* is not assigned any particular thematic role, and therefore, unlike superficially similar constructions in other languages, it may be understood as either the agent or the patient of *pukul* 'hit'. Its meaning is simply A ( 1SG HIT ), or 'entity associated with me and hit'. And in (6), the proximal past expression *tadi* may be understood as describing the activity *makan* 'eat', like a past-tense marker in other languages; however, it may also be understood as denoting some contextually-given entity connected to the past in some way, for example, the person that was just present, or the thing that was just referred to — in which latter cases, the recent-past entity is then associated in some way with *makan* 'eat', without specification of thematic role. Again, these and numerous other available interpretations may all be subsumed under the single bare associational interpretation A ( EAT PST.PROX ), or 'entity associated with eat and proximal past'.

While forms such as *aku* and *tadi* in (5) and (6) behave like most other forms in Riau Indonesian, including *ayam* 'chicken' and *makan* 'eat', in many other languages their counterparts do not; instead, they exhibit a variety of properties that justify their being set apart as belonging to categories such as pronouns and tense markers. Members of such categories are commonly referred to as grammatical items, and it is the meanings of such items that are represented in Table 1 with the letter G.

Although linguists talk of grammatical items all the time, it is actually quite difficult to come up with a straightforward definition that would clearly and unequivocally distinguish grammatical items from their complement set of non-grammatical items (see Boye & Kasper 2012 for discussion). A more practical approach, therefore, is to compile a list of properties that are commonly associated with grammatical items; the more such properties an individual item displays, the more prototypical it may be considered as an instantiation of the class of grammatical items (Heine & Reh, 1984; Lehmann, 2002; Croft, 2003: 224–225). A list of eight such properties, drawing from work by numerous scholars (including, among others, Heine et al., 1991, Heine & Kuteva, 2002, Lehmann, 2002), is presented in Table 2 below:

The above properties may be illustrated with reference to a typical example such as the English grammatical item *-s*, marking the plural form in a word such as *apples*. First, and most obviously, grammatical items tend to be shorter than other non-grammatical items, and also to make use of a more limited inventory of sounds;



**Table 2** Grammatical Items

|   |            | <i>Grammatical</i>      | <i>Non-grammatical</i>  |
|---|------------|-------------------------|-------------------------|
| a | Phonology  | more reduced            | less reduced            |
| b |            | more welded             | less welded             |
| c |            | smaller class           | larger class            |
| d |            | more strongly bound     | less strongly bound     |
| e | Semantics  | possibly paradigmatic   | non-paradigmatic        |
| f |            | more abstract           | more concrete           |
| g |            | poorer                  | richer                  |
| h | Pragmatics | less discursive primacy | more discursive primacy |

for example, *-s* consists of but a single segment. Second, grammatical items tend to exhibit stronger phonological welding to their hosts (in the sense of Haspelmath, 2021) than do other non-grammatical items; in the case at hand, the weldedness of *-s* is manifest in the morphophonemic alternations that it undergoes in accordance with the final segment of its host, for example, [-z] in [æplz] *apples*, [-ɪz] in [ɔrɪndʒɪz] *oranges*, and [-s] in [kɔ:kənʌts] *coconuts*. Moving on to the morphosyntax, grammatical items belong to smaller and closed grammatical classes, while non-grammatical items constitute larger and often open grammatical classes; thus, in *apples*, whereas *apple* can be replaced by a large and open ended set of nouns, including even nonsense ones such as the *tove* in Lewis Carroll's *T'was brillig and the slithy toves*, *-s* is *sui generis* — there is no other form in the language that exhibits similar behavior, and even Lewis Carroll wouldn't have been able to invent one. Next, grammatical items tend to be more strongly bound to their neighbors than do their non-grammatical counterparts, the most salient manifestation of this being their inability to stand alone; thus, while *apple* can occur on its own as a complete utterance, e.g., in response to a question such as *What flavor is this ice cream?* *-s* clearly cannot. Another property characteristic of grammatical items is that apparent concatenations of grammatical and non-grammatical items often lend themselves more appropriately to "Word and Paradigm" approaches (e.g., Blevins, 2016), which make reference to paradigms representing processes that apply to words; for example, a form such as *apples* might be analyzed not as *apple* plus *-s*, but rather as *apple*[plural], where the actual form of the plural marking may vary in accordance with the choice of word, e.g., *apple/apples*, *goose/geese*, *ox/oxen*, and so forth. In terms of their semantics, grammatical items are typically more abstract than non-grammatical items; thus, while *apple* refers to a concrete object, the meaning of *-s* makes reference to a more abstract notion of plurality. Related to this, the semantics of grammatical items is generally more impoverished than that of non-grammatical items; for example, it is often said that grammatical items such as *-s* are "semantically bleached" (Trask, 1993: 123, Heine & Kuteva, 2002) in comparison to their non-grammatical counterparts such as *apple*, and similarly, grammatical items may be characterized as "encyclopically poorer" than non-grammatical items (Gil, 2015: 317–318). Finally, in terms of their discourse function, grammatical items tend to

be backgrounded and less prominent than their non-grammatical counterparts, as reflected, *inter alia*, in their inability to bear focus; for example, while it is easy to assign contrastive stress to *apple*, as in, say, *John likes APPLES (not oranges)*, it is not possible to apply contrastive stress to *-s*, as in *\*John ate appleS (not just one apple)* (Boye & Kasper 2012). Of course, the eight criteria in Table 2 do not always coincide as neatly as they do in the contrast between grammatical *-s* and non-grammatical *apple*; many intermediate cases exist. Some examples of items straddling the boundary between grammatical and non-grammatical in English might include forms such as *have*, *while* and *top*. Still, the eight criteria in Table 2 correlate sufficiently well to justify treating them as instantiations of a single more general distinction between grammatical and non-grammatical items.

With all of this in mind, some examples of grammatical items taking part in simple constructional compositionality are provided in (7)–(9) below:

- (7) *Hebrew* T2  
 teyale  
 tea:DIM  
 A ( tea ENDEARMENT )  
 'tea' [expressing endearment]
- (8) *Ahousaht Nootka* (Kess & Copeland, 1984:17) T2  
 λułχaχ  
 good:DIM  
 A ( good BABY.TALK )  
 'good' [talking to baby]
- (9) *Tagalog* T2  
 Beynte po  
 twenty POL  
 A ( twenty POLITE )  
 'twenty' [with politeness]

In (7) and (8), the grammatical item is a diminutive suffix. In (7), uttered by a caregiver offering tea to a patient, the Hebrew suffix *-ale* expresses endearment, while in (8), the Ahousaht Nootka suffix *-χaχ* is a stylistic marker associated with baby talk — the register used by adults when addressing young infants. In (9), uttered by a marketplace vendor to a customer asking about the price of some merchandise, the grammatical item in question is the politeness clitic *po*. In all three examples, the meaning of the whole is a loose association of the meanings of the constituent parts, 'tea' plus endearment in (7), 'good' plus baby addressee in (8), 'twenty' plus politeness in (9) — thereby instantiating simple constructional compositionality. Such examples of grammatical items being involved in simple compositionality are actually relatively uncommon; more often, the meaning of the grammatical item also contains reference to the meaning of the sister constituent, resulting in relational compositionality — see examples (13)–(16) and (21)–(24) below.

As suggested in Table 1, non-grammatical meanings, represented with X and Y, and grammatical meanings, represented with G, share an important property, namely their material nature, associated, as they are, with combinations of individual segments. However, in many other cases, signs are built out of more abstract configurations; these are the meanings represented in Table 1 with the letter C. Configurational signs are a mixed bag, coming in a variety of flavors among which are repetition, linear order, and various suprasegmental features such as volume, duration and pitch.

Some examples of configurational signs taking part in simple constructional compositionality are provided in (10)–(12) below:

- (10) *Riau Indonesian* (Gil, 2005a:40) T3  
 Balai Balai Balai Balai Balai  
 Balai Balai Balai Balai Balai  
 A ( URGENCY BALAI )  
 'Balai' [with urgency]
- (11) *English* T3  
 Foul! [loud]  
 A ( AROUSAL FOUL )
- (12) *English* T3  
 Monday [high-rising-terminal intonation]  
 A ( TENTATIVENESS MONDAY )

Example (10) illustrates the widespread usage of repetition by vendors crying out their wares, in this case a man on a pier trying to attract passengers to his boat that is about to depart; repetition of the destination, Balai (actually an abbreviation of the fuller name Tanjung Balai), gives expression to the urgency of his call. Example (11) represents a typical shout by a football fan watching a match in progress, where the increased volume and pitch reflects the speaker's excitement. And example (12) is an instance of *upspeak*, a right-rising-terminal intonation contour; while the expressive functions of *upspeak* are variegated (Fletcher et al., 2002; Lowry, 2011; Warren, 2016 and others), one of its common usages, represented here, is to convey tentativeness, commonly associated with subordinate social status (Lakoff, 1973). In all three cases, the configurational sign expresses the speaker's mental state, whose connection to the meaning of the other item is underspecified, as represented by the formula making reference to the Association Operator.

Common to the examples of simple compositionality illustrated in (2) and (5)–(12) above is the loose semantic relationship between the two meanings, represented by the unconstrained application of the Association Operator: anything that has to do with the meanings of the two constituent signs is a potentially available interpretation of their combination. However, in the case of grammatical items and configurational signs, simple compositionality is probably less common than relational compositionality, in which the core meaning of the grammatical item or configurational sign is supplemented by an additional meaning specifying the

semantic relationship between the signs that, together, constitute the construction. Such additional semantic specification imposes constraints on the interpretation of the construction, narrowing it down from everything-goes simple compositionality. The more highly constrained application of the Association Operator is represented with the subscript  $A_{rel}$ . Examples of various kinds of relational compositionality are discussed in (13)–(28) below.

Examples (13)–(16) below illustrate the case of compositionality involving two elements, where one of them is a grammatical item, which, as suggested by the superscripted  $G^1$  in Table 1, is monovalent, specifying a semantic relationship between it and its other sister meaning:

- |      |   |    |
|------|---|----|
| (13) | <i>Roon</i>   | T4 |
|      | imun<br>1SG:hit<br>$A_{rel} ( 1SG HIT )$<br>'I hit (someone)'         |    |
| (14) | <i>English</i>  | T4 |
|      | ate<br>$A_{rel} ( EAT PST )$  |    |
| (15) | <i>Hebrew</i>   | T4 |
|      | znavnav<br>tail ~ DIM<br>$A_{rel} ( TAIL SMALL )$<br>'little tail'    |    |
| (16) | Riau Indonesian   | T4 |
|      | Ke kedai<br>DIR shop<br>$A_{rel} ( DIRECTION SHOP )$<br>'to the shop' |    |

Example (13) in Roon (an Austronesian language spoken on the eponymous island in the Cenderawasih Bay off the north coast of New Guinea) presents a garden-variety case of verbal argument indexation. With respect to the core meanings of its constituent parts it completely parallels Riau Indonesian (5); however, unlike *aku* in (5), the prefix *i-* in (13) does not merely express the 1st person singular, it also marks the relationship to its host *-mun*, which assigns it the thematic role of agent. Thus, whereas the interpretation of (5) is wholly determined by the Association Operator, that of (13) is further constrained by the relational nature of the grammatical prefix *i-*. Example (14) in English provides a commonplace illustration of verbal tense marking. Again, with regard to the basic meanings of its constituent parts it closely resembles Riau Indonesian (6); however, unlike *tadi* in (6), the inflection on the verb in (14) does not just denote a past time, but, crucially, entails that the time in question is an

attribute of the activity denoted by the verb. Accordingly, while the interpretation of (6) is determined entirely by the Association Operator, that of (14) is further narrowed down by the relational nature of the past-tense inflection on the verb. Example (15) in Hebrew illustrates a diminutive construction, formed by reduplication of the final two consonants of the root *z-n-v* 'tail' plus vocalic intercalation; the core meaning of the diminutive here is 'small'. Although formally resembling the diminutive constructions in (7) and (8) earlier, it is more highly constrained in its semantics: whereas the meanings of (7) and (8) are wholly determined by the Association Operator, that of (15) is more specific, in that 'small' necessarily stands in an attributive relationship to 'tail' — *znavnav* cannot just refer to anything that has to do with 'tail' and 'small', such as for example 'tail of a small animal'. Finally, example (16) in Riau Indonesian is a run-of-the-mill instance of argument flagging, in this case by the directional marker *ke*. If the relationship between the meanings of *ke* and *kedai* were unconstrained, a variety of alternative interpretations would be available, such as for example 'shop selling directions'; in fact, however, the only available interpretation is that in which the shop is the location constituting the goal of the direction. In summary, then, the above four examples provide a feel for the widespread occurrence of relational compositionality involving monovalent grammatical items; indeed, this is possibly the most common kind of compositionality into which grammatical items may enter.

Examples (17)–(20) illustrate cases of relational compositionality in which the monovalent element is not a grammatical item but rather a configurational sign:

- (17) *Riau Indonesian* (Gil, 2005a:45) T5  
 Nyilam nyilam nyilam nyilam nyilam nyilam  
 dive dive dive dive dive dive  
 A<sub>rel</sub> ( ITERATION DIVE )  
 'dive repeatedly'
- (18) *Riau Indonesian* (Gil, 2022:486) T5  
 panjang-panjang  
 INT ~ long  
 A<sub>rel</sub> ( INTENSIFICATION LONG )  
 'very long'
- (19) *English* T5  
 tea? [rising intonation]  
 A<sub>rel</sub> ( YNQ TEA )
- (20) *English* T5  
 Messi! [with sudden increase in pitch and volume midway through word]  
 A<sub>rel</sub> ( AROUSAL MESSI )

Examples (17) and (18) in Riau Indonesian illustrate the cross-linguistically widespread iconically-motivated usages of repetition and reduplication to express notions such as iteration and intensification. In (17), repetition expresses iteration. It may seem too obvi-

ous to be worthy of mention, but the combination of 'dive' and iteration is relational: it is the diving that is repeated, not some other activity that is only more loosely connected to the diving — the sentence cannot mean, say, 'dived (once), and (during the dive) shot repeatedly at fish'. And in (18), reduplication expresses intensification; here too the combination of 'long' and intensification is relational, with the intensification applying to 'long' and not to some other contextually-determined entity. Thus, whereas in (10), repetition is a simple configurational sign, in (17) and (18) repetition and reduplication are relational, imposing further restrictions on the interpretations of their respective constructions beyond mere associationality.

Examples (19) and (20) in English illustrate variegated usages of voice modulation as configurational signs. In (19), rising intonation marks the utterance as constituting a yes/no question; depending on the context, it might mean 'Would you like some tea?', 'Are those bushes tea bushes?', or any number of other possible yes/no questions. However, the yes/no question must involve 'tea' as part of the propositional content that is being questioned; in this sense it constitutes a relational configurational sign. In (20), a television commentator is reporting live on a football match. As superstar Messi dribbles the ball towards the opponents' goal, passing one defender after another, the commentator reports on the player's progression by repeating the player's name, *Messi ... Messi ...*. Suddenly, Messi shoots, and scores a goal. Rather than saying something like *He scores*, the commentator reports on the event by means of intonation alone, with a sudden rise in pitch and volume, subsequently drawn out over several seconds — all of this somewhere in the middle of the word *Messi*. In this particular style of live sports commentary, it is a convention that the dramatic intonation contour means that Messi scored a goal, and not merely that some other exciting event took place, regardless of whether it actually involved Messi. Thus, this particular intonation contour, which might be dubbed the *scores* contour, is also a relational configurational sign: not only does it denote the particular activity of scoring a goal, but in addition, it predicates the activity of the player denoted by the name that plays host to the intonation contour. With respect to their semantic specificity, examples (19) and (20) thus stand in contrast to the simple compositionality of examples (11) and (12) earlier, in which the meaning associated with intonation did not enter into a semantic relationship with the expression hosting the intonation contour.

Examples (13)–(20) above illustrated various cases of relational compositionality in which the grammatical item or configurational sign are monovalent, applying to a single sister element. In other cases, however, the grammatical item or configurational sign is bivalent, applying simultaneously but differentially to two sister terms. Some examples of bivalent relational compositionality are provided in (21)–(28) below.

Examples (21)–(24) illustrate cases of relational compositionality in which the bivalent element is a grammatical item, represented in Table 1 with  $G^2$ :

- |      |                                      |    |
|------|--------------------------------------|----|
| (21) | <i>English</i>                       | T6 |
|      | listen to music                      |    |
|      | $A_{rel}$ ( DIRECTION LISTEN MUSIC ) |    |
| (22) | Vietnamese                           | T6 |
|      | hai con chó                          |    |

- two CLF dog  
 A<sub>rel</sub> ( LIVE.OBJECT TWO DOG )  
 'two dogs'
- (23) *Tagalog* T6
- pangit na aso  
 ugly LNK dog  
 A<sub>rel</sub> ( ATTRIBUTION UGLY DOG )  
 'ugly dog'
- (24) *Amarasi Meto* (Edwards, 2020:4,256) T6
- faut ko?u  
 stone\CONSTR big  
 A<sub>rel</sub> ( ATTRIBUTION STONE BIG )  
 'big stone'

Example (21) in English illustrates the grammatical item *to*, associated, very broadly, with a directional meaning, in construction with *listen* and *music*. On the face of it, *to music* in (21) presents a parallel to *ke kedai* (DIR shop) in Riau Indonesian (16); however, there is a crucial difference between the two constructions. To begin with, while the directional meaning of *ke* is usually literal, that of *to*, in (21) and many other cases, is metaphorical. Its metaphorical nature is related to an arbitrary grammatical fact, namely that its presence is lexically licensed by the verb *listen*: if *listen* were replaced with the semantically similar *hear*, the presence of *to* would no longer be possible. Thus, in (21), the grammatical item *to* is bivalent, its presence relating, albeit in different ways, to both *music* and *listen*. In contrast, in Riau Indonesian, *ke kedai* can be preceded by any activity expression whatsoever, for example *makan ke kedai* (eat DIR shop) 'go to the shop and eat', *nonton ke kedai* (watch DIR shop) 'go to the shop to watch a show', and so forth — suggesting that, unlike *to*, *ke* is monovalent. The distinction between monovalent *ke* and bivalent *to* corresponds, albeit imperfectly, to a distinction drawn by some grammarians between two different kinds of case assignment, sometimes referred to as "inherent" and "structural" (Chomsky, 1995).

Examples (22)–(24) illustrate a variety of strategies in which an attributive noun-modifier construction is marked as such by an additional grammatical item occurring in between the two terms. In (22) in Vietnamese, for a numeral to quantify a noun, a numeral classifier must also be present, providing additional information about the semantic category to which the noun belongs. In (23) in Tagalog, attributive constructions are marked with the "linker" or "ligature" *na* (or its allomorph *-ng*); if the linker is replaced by the article/topic marker *ang*, the interpretation would change from attributive to predicative — *Pangit ang aso* 'The dog is ugly'. Similarly, in (24) in Amarasi Meto, attribution is marked by the construct form of the head noun, formed by metathesis; with the unmarked absolute form of the noun, the interpretation would, again, be predicative — *Fatu ko?u* 'The stone is big'. Thus, in all three cases, the grammatical item in question functions bivalently, by qualifying the semantic relationship that obtains between the two other terms, and thereby restricting the range of possible interpretations of the construction.

Finally, examples (25)–(28) illustrate cases of relational compositionality in which the bivalent element is a configurational sign, represented in Table 1 with  $C^2$ :

|      |  |    |
|------|--|----|
| (25) | <i>Nage</i>  | T7 |
|      | Manu ka<br>bird eat<br>$A_{rel}$ ( AGENT BIRD EAT )<br>'The bird is eating'                                  |    |
| (26) | <i>Sri Lankan Malay</i> (Peter Slomanson p.c.)   | T7 |
|      | Kotor ayer<br>dirty water<br>$A_{rel}$ ( ATTRIBUTION DIRTY WATER )<br>'dirty water'                          |    |
| (27) | <i>Riau Indonesian</i>   | T7 |
|      | tak enak [single intonational phrase]<br>NEG TASTY<br>$A_{rel}$ ( APPLICATION NEGATION TASTY )<br>'not nice' |    |
| (28) | <i>Hebrew</i>  | T7 |
|      | hazrait<br>pig/truck<br>$A_{rel}$ ( HYBRID PIG TRUCK )<br>'pig/truck hybrid'                                 |    |

In (25) and (26), the configurational sign is linear order. Example (25) in Nage (an Austronesian language spoken in central Flores island in Indonesia) illustrates the cross-linguistically widespread use of word order to distinguish thematic roles, or, under alternative analyses, grammatical relations. Specifically, with *manu* preceding *ka*, it is interpreted as its agent; however, if *manu* were to follow *ka*, it would be interpreted as its patient — *Ka manu* 'eating the bird'. Example (26) in Sri Lankan Malay shows how linear order may also be used to distinguish attributive from predicative interpretations, in similar fashion to the grammatical items in examples (23) and (24) earlier. With *kotor* occurring before *ayer*, it is understood attributively; however, if *kotor* were to follow *ayer*, it would be understood predicatively — *Ayer kotor* 'The water is dirty'. In both (25) and (26), then, linear order serves to narrow down the semantic relationships that may obtain between the two terms.

In (27) in Riau Indonesian, the configurational sign is suprasegmental, illustrating a commonplace instance of intonation being used to mark syntactic nexus and constituency. Bearing a single intonational phrase, *Tak enak* means 'not nice', with the negative marker applying directly to its host word; however, with an intonational break between the two terms, *Tak, enak* would mean 'no, it's nice'.



In (28) in Hebrew, the configurational sign is blending, the fusion of two words into one, typically involving the shortening of each of the original constituent words. Familiar examples of blends in English are *motel* (from *motor* and *hotel*) and *smog* (from *smoke* and *fog*). However, while the preceding examples are conventionalized, and may be considered part of the English lexicon, *ħazrait* in (28) is not a real word in Hebrew; instead, it was produced creatively, on the fly, in an experimental setting in which subjects were asked to offer descriptions of images portraying hybrid entities (Shen & Gil, 2017:1183–5). In the case at hand, the image was part pig and part truck, and the subject chose to describe the image with a playful and innovative combination of the Hebrew words for 'pig' *ħazir* and 'truck' *masait*. In examples such as (28), then, blending is a configurational sign whose meaning is an iconic reflection of the hybrid nature of the blend's referent.

In summary, then, examples (5)–(28) in this section show how the typology of compositionality is instantiated in human languages. The primary distinction is between bare compositionality in (5)–(6) and constructional compositionality in (7)–(28); within constructional compositionality, further distinctions obtain between simple compositionality in (7)–(12), monovalent compositionality in (13)–(20), and bivalent compositionality in (21)–(28), and cross-cutting these, between compositionality involving grammatical items in (7)–(9), (13)–(16) and (21)–(24), and compositionality involving configurational signs in (10)–(12), (17)–(20) and (25)–(28).

The examples discussed in this section were all selected for their simplicity; specifically, for the fact that they consist of only two, or, in the case of bivalent relational compositionality, three signs. In actuality, much human linguistic behavior involves complex utterances consisting of more than two or even three signs; for example, as pointed out near the beginning, even a simple English sentence such as (3) contains a multiplicity of signs: *the*, *chicken*, *is + ing*, present tense, *eat*, and additional configurational signs pertaining to linear order and intonation. Examples such as (3) raise the question whether relational compositionality is limited to monovalent and bivalent types, or whether there might exist constructions in which a grammatical item or configurational sign is associated with a higher valency, applying to three or more sister signs.

It may be conjectured that human languages do not permit relational configurationality involving degrees of valency that are higher than two. The plausibility of this conjecture is of course crucially dependent on the grammatical analysis to which any putative counterexample to this conjecture is subjected. What it says, essentially, is that any apparent instance of higher valency is more appropriately reduced to a hierarchical structure in which, at every level, the valency is at most two. For example, a hypothetical structure of the form  $A_{rel} ( X Y Z G^3 )$  containing a would-be trivalent grammatical item might be more appropriately analyzed as, say, an instance of bivalent relational compositionality embedded within a higher structure, for example  $A ( X A_{rel} ( Y Z G^2 ) )$ . If true, this conjecture would constitute an important constraint on compositionality in human language.

## Compositionality in Animal Communication

Having presented the typology of compositionality and illustrated it with examples from human languages, we now turn to an exploration of reported instances of compositionality in animal communication. Following the typology laid out in

Table 1, this section examines the extent to which compositionality in animal communication resembles that observed in human languages. As we shall see, most of the types of compositionality proposed in Table 1 have potential instantiations in various reported cases of animal communication; however, alongside such parallels there are also significant differences between the ways in which compositionality is manifest in the animal and human domains. The present endeavour is conducted in the spirit of "formal monkey linguistics" (Schlenker et al., 2016a, 2016b; see also Sabaté et al., 2022), attempting to apply methods and concepts from the study of human languages to the communicative systems of other animals.

The exploration of compositionality in animal communication conducted in this section is subject to a number of substantial methodological limitations. First, since animal communication is not my area of expertise, the discussion and analysis is based on an outsider's reading of published work by experts in the field, so there is ample potential for omissions and misrepresentations on my part, for which I can only beg forbearance. But secondly, my impression is that when compared to linguistics, the field of animal communication is still very much in its infancy, and the body of shared knowledge is increasing in leaps and bounds from year to year, as primatologists and other animal communication specialists make more and more exciting discoveries about how various species communicate. This means not only that I may have missed out on significant recent reports, but also that whatever we know today might easily be eclipsed by significant future discoveries and increases in our understanding lying just around the corner. Thirdly, animal communication systems are, by their very nature, much less accessible to human researchers than human languages: we cannot simply draw on our own familiarity with our native languages, or ask a friendly orang-utan or whale what the meaning was of something that they just said. Finally, as in linguistics but perhaps even more acutely, alternative analyses of the same data may result in conflicting characterizations of the same construction as instantiating different types of compositionality — some examples of this are considered below.

When attempting to apply the typology of compositionality in Table 1 to animal communication, an immediate problem arises: it is not obvious to what extent animal communication systems present analogues to the three-way distinction, represented in the columns of Table 1, between non-grammatical items, grammatical items, and configurational signs. With regard to the distinction between non-grammatical and grammatical items, a few scholars have proposed certain analogies, such as the characterization, discussed below, of calls such as the Diana monkeys' -A and the Campbell monkeys' -oo as suffixes. However, in many cases, it is not clear how the characteristic features of grammatical items presented in Table 2 might apply to animal communication systems. As for the distinction between material and configurational signs, this is based on the presumed primacy of the segmental tier, containing the phonemes, morphemes, words, and phrases of human languages, as opposed to everything else. However, when dealing with radically different systems of communication, such as those of various animals, it is not always obvious to what extent such a distinction can be maintained. For this reason, the classification of various instances of compositionality in animal communication in accordance with the typology in Table 1 must necessarily remain tentative.

Before beginning the exploration of compositionality in animal communication, it is first necessary to clear the deck by eliminating all sorts of phenomena that do not qualify as exhibiting compositionality. First, compositionality cannot exist without meaning; this excludes various dance- or music-like behaviours lacking in meaning, such as, for example, gorilla gesture sequences (as described by Tanner, 2004 and Tanner & Perlman, 2016). Secondly, for compositionality to be present, there must be two or more elements in combination; this rules out forms such as a single cat's meow or dog's bark. However, for compositionality to obtain, these two conditions do not suffice, and a third condition must be met, namely that meaning be associated with forms at least two distinct hierarchical levels, that of the whole and that of its constituent parts.

This third condition may fail to be met in two different ways. The first is when meanings are present at the level of the constituent parts but not that of the whole. A possible example of this is provided by the calls of black-fronted titi monkeys (*Callicebus nigrifrons*) (as analyzed by Schlenker et al., 2016a, 2016b; Schlenker et al., 2017). Titi monkey calls consist of sequences of *A* and *B* calls, where *A* calls refer to serious non-ground threats, while *B* calls denote noteworthy events. Each individual call is analyzed as an independent utterance reflecting the state of the environment at the precise time of production; meanings are thus associated with individual calls, but not with the sequence as a whole. This state of affairs is analogous to that which obtains, in human languages, between a string of utterances that constitute a discourse: while each of the individual utterances is associated with a meaning, the discourse as a whole is not.

A second diametrically opposed way in which the third condition may fail to be met is when meaning is associated with a sequence of forms but not with any of its constituent parts. Examples of this include the songs of European starlings (*Sturnus vulgaris*) (as analyzed by Gentner et al., 2006), humpback whales (*Megaptera novaeangliae*) (as described by Suzuki, Buck, & Tyack 2006), and white-handed gibbons (*Hylobates lar*) (as per Clarke et al., 2006). In such cases, although meanings of various kinds may be attributed to entire sequences, there is no evidence that such meanings are built up out of smaller meanings associated with the sequences' constituent parts. This case is analogous to the way in which, in human languages, a morpheme is built up out of a sequence of segments: while the morpheme as a whole bears a meaning, each individual segment does not.

With all of the preceding considerations in mind, we now turn to examine possible instances of compositionality in animal communication. Examples (29)–(40) below provide putative examples of compositionality, in accordance with the typology of compositionality proposed in Table 1. The discussion follows the format and order of presentation adopted in the previous section for human language compositionality, in order to facilitate a direct comparison between animal communication and human language, and highlight the similarities and differences between the two.

As in the preceding section, we begin with instances of bare compositionality, in (29)–(32) below, corresponding to human language examples (2), (5), and (6) earlier. The clearest cases of bare compositionality in animal communication come from the behaviour of captive great apes making use of various communication systems taught to them by their human caregivers. Following are examples of bare

compositionality produced by the bonobo Kanzi (Greenfield & Savage-Rumbaugh, 1990) and the orang-utan Chantek (Miles, 1990):

(29) *Bonobo, using lexigrams* T1

LIZ HIDE  
Liz hide  
A ( LIZ HIDE )

(30) *Orang-utan, using American Sign Language* T1

BEARD PULL  
Beard pull  
A ( BEARD PULL )

As shown by numerous studies, the communicative repertoire of Kanzi and Chantek clearly makes use of compositionality: they are manifestly capable of taking signs that they have already mastered, and combining them in novel ways to build up new meanings from the meanings of the constituent parts. However, their compositionality is of the bare variety, lacking any of the constructional constraints that are characteristic of much human language. Thus, Kanzi also produces utterances such as WATER HIDE, showing that the first of the two signs is unspecified for thematic role, and may be understood as either the agent or the patient of HIDE, as well as utterances such as HIDE AUSTIN, showing that word order plays no role in thematic role assignment, and that the agent of HIDE may occur either before it or after it. Similarly, Chantek also produces utterances such as YOU PULL, showing that the first of the two signs is unspecified for thematic role, and may be understood as either the agent or the patient of PULL, as well as utterances such as PULL BEARD, showing that word order plays no role in thematic role assignment, and that the patient of PULL may occur either before it or after it. Thus, just as in Riau Indonesian examples (2), (5), and (6) in the preceding section, the interpretations of captive great ape utterances (29) and (30) are most appropriately represented in terms of the Association Operator, as A ( LIZ HIDE ) and A ( BEARD PULL ), without any further constraints of a constructional nature: that is to say, they exhibit bare compositionality.

Although reflective of Kanzi and Chantek's cognitive abilities, the kind of communicative behaviour represented in (29) and (30) is in a sense unnatural. One might perhaps have hoped for Kanzi's and Chantek's captive conspecifics to pick up their behaviour and even pass it on to subsequent generations, thereby creating communities of speaking apes; but for whatever reason, such learning does not usually happen; the only known case of this being the chimpanzee Louris, who was reported to have acquired some signs of American Sign Language from another chimpanzee, Washoe. (Gardner et al., 1989). Moreover, no similar communicative systems have yet been attested amongst great apes or other primates in the wild. The reasons for this are potentially many, but one obvious factor is purely structural. Whereas the vocabularies of lexigrams and American Sign Language mastered by Kanzi and Chantek respectively are relatively large, numbering in the dozens, the repertoires of individual cries that have been described for primates in the wild are mostly much

smaller, consisting at most of a handful of distinct cries, generally not more than the 12 that have been suggested for chimpanzees (Girard-Buttoz et al., 2022). Given such limited lexicons, there would seem to simply be less practical motivation to combine such cries compositionally.

Outside of primates, potential instances of bare compositionality may be observed amongst the birds and the bees. One apparent example is that of Japanese tits (*Parus minor*) (as described and analyzed by Suzuki et al., 2019; Suzuki, 2021); under their analysis, Japanese tits' calls would qualify as exhibiting bare compositionality. However, under a subsequent analysis, cf. discussion of (39) below, their calls are more appropriately viewed as exhibiting a more complex variety of constructional compositionality. More clear cut cases of bare compositionality are provided by two species of chickadee, the black-capped chickadee (*Poecile atricapillus*) (as described by Hailman et al., 1985; Hailman & Ficken, 1986), and the Mexican chickadee (*Poecile sclateri*) (as described by Ficken et al., 1994) and exemplified below:

- (31) *Mexican Chickadee* T1
- |  |                     |
|--|---------------------|
| $A^k$                                  | $C^m$               |
| RESTLESSNESS                           | DISTURBING.STIMULUS |
| A ( RESTLESSNESS DISTURBING.STIMULUS ) |                     |

Mexican chickadees are described as having four note types, represented as *A*, *B*, *C*, and *D*. While *B* occurs too infrequently to enable its meaning, if any, to be determined, the remaining three note types are associated with particular meanings: *A* denotes restlessness and flight, either incipient or in progress, *C* conveys a disturbing stimulus and a tendency to change direction, while *D* indexes a perching location. Calls consist of a variable number, zero, one, or many, of each of the four note types, but in strict sequence,  $A^k-B^l-C^m-D^n$ , where superscripts denote the number of occurrences of each note type. The semantics is compositional, with the meaning of the whole call derived straightforwardly from the meaning of its constituent parts, in accordance with the Association Operator; for example, in (31) above, the meaning of the call combines restlessness with disturbing stimulus. Thus, Mexican Chickadee calls resemble the communication systems of Kanzi and Chantek in the presence of bare compositionality; however, they differ from them in two other important respects. First, their lexicon is more limited in size; secondly, whereas the captive ape signs may occur in any order, the relative order of the chickadee note types is fixed. However, since the linear order does not contribute to the meaning in any way, the combinations of *A*, *B*, *C*, and *D* note types may be characterized as exhibiting bare rather than constructional compositionality. In contrast, a somewhat different type of compositionality is exhibited within each sequence of identical note types, discussed below, in (35).

A very different case of bare compositionality is provided by the well-known waggle dance of the honeybee (*Apis mellifera*), (as described by Von Frisch, 1967; Riley et al., 2005; Preece & Beekman, 2014):

- (32) *Honeybee* T1

## ORIENTATION\DURATION

Direction\distance

A ( DIRECTION DISTANCE )

The function of the waggle dance is to convey the location of a desirable resource such as a patch of flowers or a source of water. Simplifying somewhat, the waggle dance consists of two meaningful components: its orientation, indicating the direction to the resource, and its duration, representing the distance to the resource. As indicated in (32) above, these two components stand in a relationship of bare compositionality. In structural terms, the waggle dance differs in two important respects from the captive apes' artificial languages and the chickadee calls. First, whereas the great ape signs and the chickadee calls are produced sequentially, the two components of the bees' waggle dance are produced simultaneously. Secondly, whereas the lexicons of the great apes and the chickadees consist of discrete and arbitrary signs, those of the bees are continuous and iconic, constructed out of physical dimensions of orientation and duration. Still, notwithstanding these differences and others, the waggle dance shares with the communicative systems of both the captive apes and the chickadees its reliance on bare compositionality — the meaning of the whole, namely the specification of a location, being derived straightforwardly and without any further constraints from the meanings of its constituent parts, the orientation and the duration.

Examples such as those exemplified in (29)–(32) represent the only clear cases that I have been able to identify of bare compositionality being reported in animal communication. Of course there may be others, ones that may have escaped my attention, or alternatively that are still waiting to be discovered and appropriately analyzed. However, at this point at least, there would appear to be no clear-cut case of bare compositionality amongst primates in the wild. Instead, available descriptions of compositionality amongst wild primates all seem to instantiate one or another form of constructional compositionality. Examples of constructional compositionality in primates and birds are presented in (33)–(40) below, beginning with simple constructional compositionality in (33)–(35).

An apparent instance of constructional compositionality of the simple variety making use of a putative grammatical marker is provided by the calls of female Diana monkeys (*Cercopithecus diana*) (as analyzed by Stephan & Zuberbühler, 2008; Candiotti et al., 2016). Female Diana monkeys are described as having four distinct cries: *H* for socially positive or relaxed situations, *L* for neutral situations, *R* for socially negative or potentially dangerous situations, and *A* whose function is to identify the caller. The latter *A* cry typically occurs after one of the preceding three cries, as in the following example:

(33) *Female Diana monkey*

T2

H-A

SOCIALLY.POSITIVE.RELAXED.SITUATION-IDENTITY

A ( SOCIALLY.POSITIVE.RELAXED.SITUATION IDENTITY )

Under one analysis (Veselinović et al., 2014), when occurring after *H*, *L*, or *R*, *A* is a "suffix", which may accordingly be written -*A*: the argument appeals to the absence

of any discernible pause between the *H*, *L*, or *R* and the following *-A*. To the extent that the characterization of *-A* as a suffix holds water, the *-A* cry would seem to constitute a grammatical item, as distinct from the *H*, *L*, and *R* calls. It should, however, be acknowledged that the *-A* cry lacks a characteristic property of grammatical items in human languages, namely that they constitute a small and typically closed class of items (property *c* in Table 2 earlier): in the case at hand, each monkey would seem to have his or her own individually identifying *-A* cry. Turning now to the semantics, the compositionality involved is simple; as represented by the formula  $A ( \text{SOCIALY. POSITIVE.RELAXED.SITUATION IDENTITY} )$ , the meaning of *H-A* is a loose association of the meanings of *H* and *-A*, that is to say 'entity associated with socially positive and relaxed situation and me talking'. Thus, the female Diana monkey cry in (33) presents a close parallel to instances of simple grammatical compositionality in human language exemplified in (7)–(9) earlier.

Instances of constructional compositionality involving configurational signs are perhaps more commonplace. When a form is called out more loudly, or repeated, it is reasonable to suppose that it is imbued with greater assertiveness or urgency. As a pet owner, I would assume that whatever my cat's meow or dog's bark means, this meaning is amplified when the meow or bark is louder or repeated — though of course it remains to be rigorously demonstrated that such semantic amplification is intended by the animal rather than merely attributed to it by its personifying owner. However, for at least some primates and birds, this has indeed been argued to be the case. For primates, an example is provided by the cries of male putty-nosed monkeys (*Cercopithecus nictitans martini*) (as described and analyzed by Arnold & Zuberbühler, 2013; Schlenker et al., 2016a). Male putty-nosed monkeys use combinations of two basic cries, *pyow* denoting a general alert and *hack* denoting a serious non-ground-movement-related alert. Further down, in the discussion of (38), we examine the use of these two cries in combination, but for now let us consider a string of, say, five *hack* cries:

(34) *Male putty-nosed monkey*

T3

hack hack hack hack hack

SERIOUS.NON.GROUND.MOVEMENT.RELATED.ALERT

$A ( \text{ALARM.LEVEL.5 SERIOUS.NON.GROUND.MOVEMENT.RELATED.ALERT} )$

In accordance with a proposed Alarm Level Rule, a sequence of *n* calls is associated with alarm level *n*; that is to say, the more calls in the string, the higher the level of alarm. Thus, example (34) exhibits simple compositionality involving a configurational sign, in the case at hand repetition; as suggested by the formula  $A ( \text{ALARM.LEVEL.5 SERIOUS.NON.GROUND.MOVEMENT.RELATED.ALERT} )$ , the meaning of the sequence of *hack* cries is a loose association of the meanings of *hack*, namely serious non-ground-movement-related alert and alarm level 5.

Similar effects have been observed also amongst birds. Recall the cries of the Mexican chickadee discussed earlier. While combinations of different note types, *A*, *B*, *C*, and *D*, as in (31), exhibit bare compositionality, sequences of the same note type are argued to express intensity (Ficken et al., 1994). For example, when the *A* note



type, conveying restlessness and flight, is repeated, as in (35) below, this is suggested to expressed increased restlessness and speed of movement:

(35) *Mexican chickadee*

T3

A A A A A  
 RESTLESSNESS  
 A ( INTENSITY RESTLESSNESS )

As spelled out in the formula A ( INTENSITY RESTLESSNESS ), the meaning of the sequence of A tones is a loose association of the meanings of the A tone, namely restlessness, and of the repetition, namely intensity. Accordingly, as in the preceding example, here too, simple compositionality involves repetition as a configurational sign, resulting in simple constructional configurationality. Similar cases of repetition and repetition rate as a constructional sign have been reported for black-capped chickadees (Hailman et al., 1985; Hailman & Ficken, 1986; Templeton et al., 2005), as well as for *Ficedula* flycatchers (Wheatcroft, 2015). Thus, the repeated calls by monkeys and birds in (34) and (35) bear a close parallel to the cases of simple configurational compositionality in human language discussed in (10)–(12), and in particular to the Riau Indonesian example (10), in which a similar instance of repetition, here the boatman’s calling out of his boat’s destination, has a similar function, namely the expression of urgency.

To this point, all of the instances of compositionality in animal communication that we have examined involved simple compositionality. We now turn to consider some potential examples representing the more complex case of relational compositionality.

A possible example of relational compositionality involving a grammatical item is provided by the cries of Campbell’s monkeys (*Cercopithecus campbelli campbelli*) (as described and analyzed by Ouattara et al., 2009a, 2009b; Kuhn et al., 2014; Schlenker et al., 2014, 2016a, 2016b). Campbell’s monkeys make use of four cries: *boom* denoting a non-predator disturbance, *hok* denoting a non-terrestrial disturbance, *krak* denoting a general disturbance, and *-oo* with an attenuating effect. The latter form, *-oo*, is argued to be a suffix, which may attach to either *hok* or *krak* (but not *boom*) — as in the following example:

(36) *Campbell’s monkey*

T4

hok-oo  
 NON.TERRESTRIAL.DISTURBANCE-ATTENUATION  
 A<sub>rel</sub> ( NON.TERRESTRIAL.DISTURBANCE ATTENUATION )

The proposed analysis of *-oo* as a suffix, appealing to phonetic properties and to the fact that it attaches to a host cry, suggests that it is may perhaps be considered as a grammatical item. In this respect, it would resemble the female Diana monkeys’ -A cry illustrated in (33) above. However, the Campbell’s monkeys’ *-oo* cry would seem to be associated with a more complex semantic structure. Whereas the meaning of the Diana monkeys’ -A cry is independent of its host cry, expressing the identity of the caller, the



meaning of the Campbell's monkeys' *-oo* cry makes crucial reference to the meaning of its host cry — by entailing a broadening, weakening, or attenuation of the host cry's meaning. For this reason, it may be considered as a relational grammatical item. Thus, the Campbell's monkey cry in (36) closely parallels instances of relational grammatical compositionality in human language examples (13)–(16). Fleshing out the analogy, the relational Campbell's monkey suffix *-oo* is to the simple Diana monkey *-A* cry as the likes of relational Roon prefix *i-* are to the simple Hebrew suffix *-ale* and its ilk.

A potential instance of relational compositionality making use of a configurational sign is offered by the roars of Guereza colobus monkeys (*Colobus guereza*). Guereza colobus monkeys make use of two calls, *snort* and *roar*. In many cases these two calls combine; however, such combinations present significant analytical difficulties (Schlenker et al., 2016a, 2016b) and are hence not considered further in this paper. Instead, we consider sequences formed entirely by *roar* cries (Schel et al., 2010). While leopards trigger alarms consisting of many roaring sequences of only a few calls each, eagles trigger alarms comprising few sequences each with many calls, as in the following:

(37) *Guereza colobus monkey*

T5

roar roar roar roar roar roar, roar roar roar roar roar roar roar roar  
 predator\eagle  
 $A_{rel}$  ( EAGLE PREDATOR )

Two alternative analyses of Guereza colobus monkey roars present themselves. In accordance with the first, the roar itself is meaningless, while the meaning, leopard or eagle, is derived entirely from the configuration. Under this analysis, there would be no compositionality: meaning would be present at the level of the whole but not at that of the constituent parts — Guereza colobus monkey roars would thus resemble the songs of European starlings, humpback whales, and gibbons, discussed earlier. However, in accordance with a second analysis, the roar would be associated with a general meaning of predator, while the particular configuration would narrow down the meaning to that of leopard or eagle, as appropriate. Under such an analysis, Guereza colobus monkey roars would provide an instance of configurational compositionality, but one that differs in a subtle but important way from that of the repeated *hack* cries of the putty-nosed monkeys in (34) earlier. In (34), the two constituent meanings, alarm level 5 and serious non-ground-movement-related alert, although contextually closely related, are logically independent of each other — each bears its own well-defined meaning regardless of the other. In contrast, in (37), under the second analysis, the meaning conveyed by the configuration, namely leopard or eagle, is a semantic modification, or narrowing-down, of the meaning expressed by the roar, which is that of an underspecified predator. Under the second analysis, the compositionality of Guereza colobus monkey roars would accordingly be relational, presenting a parallel to the human language examples of relational configurational compositionality in (17)–(20) earlier.

The two potential cases of relational compositionality in Campbell's and Guereza colobus monkeys discussed in (36) and (37) both involve monovalent relations. As

for potential cases of bivalent relational compositionality, I have encountered no reports of potential bivalent relational compositionality involving grammatical items — the only cell in the typology of compositionality in Table 1 that remains unpopulated by an example of animal communication. However, a handful of potential cases of bivalent relational compositionality may be observed involving a configurational sign.

One such example is associated with the *pyow* and *hack* cries of male putty-nosed monkeys considered earlier. Whereas example (34) consisted of just a string of *hack* cries, example (38) illustrates the two cries, *pyow* and *hack*, in combination. (Most commonly, *pyow-hack* sequences consist of a few *pyow* cries followed by a few *hack* cries; the choice of a single *hack* followed by a single *pyow* in (38) is for ease of exposition only.)

(38) *Male putty-nosed monkey*

T7

*pyow hack*

ALERT SERIOUS.NON.GROUND.MOVEMENT.RELATED.ALERT

$A_{rel}$  ( RESTRICTION ALERT SERIOUS.NON.GROUND.MOVEMENT.RELATED.ALERT )

While earlier analyses (Arnold & Zuberbühler, 2006, 2012) consider such combinations to be non-compositional, later analyses (Arnold & Zuberbühler, 2013 and Schlenker et al., 2016a, 2016b) posit a compositional semantic relationship between the two calls, in which *pyow*, denoting a general alert, is semantically restricted by *hack*, denoting a serious non-ground-movement-related alert. Such semantic restriction is a bivalent semantic relation that holds between the two cries. Moreover, it is licensed by a configurational sign, namely the linear order of the two calls in which *hack* follows *pyow*. Thus, male putty-nosed monkey *pyow-hack* cries exhibiting bivalent relational compositionality involving a configurational sign are analogous to human language examples such as (25)–(28), and in particular (25) and (26), in which the configurational sign involved is also linear order.

A possibly similar example is provided by calls of Japanese tits, for which a number of alternative analyses present themselves. Two basic facts seem to be agreed upon: first that they possess two basic calls, an *ABC* call denoting a general predatory threat, and a *D* call used to recruit others to non-dangerous social contexts; and secondly that these two basic calls may occur in sequence, *ABC-D*, which (as described by Suzuki, 2014), is associated with the mobbing of stationary predators. The sequential use of these two calls may be represented as follows:

(39) *Japanese tits*

T7

*ABC*

*D*

PREDATORY.THREAT RECRUIT

$A_{rel}$  ( MOBBING PREDATORY.THREAT RECRUIT )

What is at issue is the appropriate analysis of the *ABC-D* call sequence. On the one hand, its association with mobbing behaviour might point towards its characterization as non-compositional (Greisser et al., 2018). However, such a characteriza-

tion ignores the obvious sense in which the mobbing interpretation of the sequence derives from the predatory threat and recruitment interpretations of its constituent parts. A series of experiments (Suzuki et al., 2019) suggests that the meaning of the *ABC-D* sequence can indeed be broken down into two components, predatory threat and recruitment. Under such an analysis, the *ABC-D* call sequence would thus constitute an instance of bare compositionality. However, as I understand it, this analysis does not seem to account for the fact that the interpretation of the *ABC-D* call sequence, namely mobbing, is more restrictive than a simple combination of predatory threat and recruitment; indeed, the notion of mobbing even seems to cancel out one of the characteristic features of the recruitment call *D*, namely that on its own, it is associated with non-dangerous social contexts. Instead, a more appropriate analysis might build on the characterization of the *ABC-D* call sequence as compositional, but with the additional proviso that the combination of the *ABC* and *D* calls into a unitary call sequence should itself be viewed as a configurational sign whose semantic effect is to restrict the meaning of the *ABC-D* call sequence to mobbing — as represented in (39). Under such a refinement of the original (Suzuki et al., 2019) analysis, the Japanese tits' *ABC-D* call sequence would thus resemble male putty-nosed monkey *pyow-hack* cries in (38), as well as human language examples (25) and (26), all of which instantiate bivalent relational compositionality with a configurational sign involving linear order. A potentially similar analysis might be available also for the cry sequences of the southern pied babbler (*Turdoides bicolor*), in which (as described by Engesser et al., 2016) alert and recruitment calls are brought together to produce a complex call whose meaning, once again, is associated with the more specific notion of mobbing.

The final example to be considered here is that of chimpanzee drumming behavior (as described by Boesch, 1991 and subsequently reanalyzed by Gabrić, 2021). Chimpanzee drumming consists of a sequence of drumming events, each denoted with the letter *D*; each such drumming event is associated with a particular tree whose identity is represented with a numeral subscript (Gabrić, 2021). (The notation for chimpanzee drumming adopted here differs from the original Gabrić, 2021 notation, which is more difficult to read.) Two main sequences are described:  $D_1 D_1$ , expressing resting period initiation, and  $D_1 D_2$ , denoting travel direction change. Of interest is what happens when these two sequences are combined. Rather than simply being juxtaposed, as in, say  $D_1 D_1 D_1 D_2$ , the resulting combination is shortened, as for example in (40) below.

(40) *Chimpanzee*

T7

$D_1 D_1 D_2$

RESTING.PERIOD.INITIATION\TRAVEL.DIRECTION.CHANGE

$A_{rel} ( \text{SEQUENTIALITY RESTING.PERIOD.INITIATION TRAVEL.DIRECTION.CHANGE} )$

(An alternative acceptable shortening is  $D_1 D_2 D_2$ .) Formally, the shortening of the combined drumming event sequence from four to three events may be considered to be a configurational sign. Moreover, the semantics of the construction is not the mere combination of the two component parts, resting period initiation and travel

direction change, which, as they stand, are mutually contradictory. Rather, the configurational sign, that is to say the shortening, introduces an additional meaning component into the mix, namely sequentiality: first we rest, then we go off in a different direction. Thus, chimpanzee drumming behavior may also be considered to instantiate bivalent relational compositionality, albeit with a configurational sign of a somewhat different nature, involving the shortening, or blending of the two constituent signs. In this respect, then, chimpanzee drumming presents a close parallel to human language blends such as in (28) (as indeed pointed out by Gabrić, 2021).

This section has explored some of the ways in which the typology of compositionality in Table 1 may potentially be instantiated by various instances of animal communication. It must be emphasized that this section does not come close to providing a comprehensive survey of compositionality in animal communication. As pointed out earlier, the exploration was constrained by a number of methodological limitations stemming from my own outsider perspective, the rapidly changing state of the art, the inherent difficulties in understanding animal communication, and the sometimes very different competing analyses that are proposed for one and the same construction. Nevertheless, the present survey will hopefully have provided a feel for some of the similarities between compositionality in animal communication and human language — but also some of the differences. In the final section we shall examine some possible implications of the preceding study with respect to the evolution of compositionality.

## The Evolution of Compositionality

The typology of compositionality proposed in this paper provided the conceptual tools for the detailed comparison of compositionality in human language and animal communication systems presented above. The main result to emerge from this comparison is that almost all types of compositionality present in human language — six out of seven in the typology of Table 1 — appear to possess potential analogues in the domain of animal communication. Thus, not only is compositionality per se not specific to human language, but even the typology of compositionality, with the various distinctions that it embodies, is just as applicable to animal communication systems as it is to human languages.

Nevertheless, in spite of the applicability of the typology of compositionality to both animal communication systems and human languages, the results of the comparison bear witness to an enormous gap between humans and other animals with respect to the distribution and prevalence of compositionality. While in human languages it is ubiquitous, in animal communication systems it is relatively uncommon, limited to a mere handful of clearly attested cases.

While the typology of compositionality proposed in this paper is expressly synchronic, it lends itself straightforwardly to a phylogenetic interpretation. Couched in terms of increasing formal complexity, the architecture of compositionality represented in Fig. 1 may also be understood as constituting a set of hypotheses about how compositionality evolved. The arrows in Fig. 1 may be interpreted as representing evolutionary paths, with each type of compositionality evolving out of the

type preceding it, and in turn forming the basis for the development of the type that follows it. As in many other domains, here too, evolution proceeds from simple to complex.

As suggested in Fig. 1, the evolution of compositionality may be broadly divided into two distinct stories. The first story, that represented at the top of the diagram, is that in which compositionality first emerges out of an earlier stage in which compositionality is absent. This is the focus of the lions' share of studies dealing with how compositionality might have evolved; in particular, it is the locus of the debate concerning whether compositionality evolved combinatorically, by the addition of syntax to an already existing lexicon, or holophrastically, by the breaking down of larger, unitary utterances (see the debate in Arbib & Bickerton, 2010). However, the focus of the present study is less with how and why compositionality first arose, and more with the story of its subsequent evolution, and the various stages that it might have traversed. Thus, Fig. 1 suggests that the evolution of compositionality should perhaps be viewed not in terms of a single giant leap from nothing to the full splendour of contemporary human-language compositionality, but rather as a more gradual process with a number of discrete waystations involving different types of compositionality of increasing complexity. In particular, it suggests that bare compositionality may be evolutionarily prior to the various types of constructional compositionality.

Empirical evidence for the evolutionary primacy of bare compositionality amongst humans derives from a number of recent and in-progress studies in linguistic typology examining the cross-linguistic distribution of bare and constructional compositionality. As shown in this paper, contemporary human languages make use of both; recall the contrast between bare compositionality in Riau Indonesian (2) *Ayam makan* and constructional compositionality in English (3) *The chicken is eating*. But on what basis can we infer that one of these constructions, specifically that in (2), is in any sense evolutionarily prior to the other, in this case that in (3)? A potential rationale for drawing such inferences is provided by observed typological correlations between linguistic and socio-political complexity (Gil, 2021:3): "[i]n domains where linguistic complexity correlates positively with cultural or socio-political complexity, simpler linguistic structures may be inferred to be evolutionarily prior to their more complex counterparts". The justification for this principle lies in the already well-established trajectory of cultural and socio-political evolution from simpler to more complex. Until yesterday we were all hunter-gatherers (Diamond, 2012). Accordingly, if we observe certain linguistic structures of lower complexity to be associated with contemporary societies of lower socio-political complexity, we may reconstruct the linguistic structures in question to an earlier stage in the evolution of language.

Indeed, a number of empirical studies suggest that simple bare compositionality correlates with lower socio-political complexity, while more complex types of constructional compositionality are found more commonly in languages associated with higher socio-political complexity. The most extensive study to show this is the in-progress Association Experiment (Gil, 2007, 2008), conducted on a world-wide sample of 69 languages. Preliminary results show that while bare compositionality, as in (2), is more common in languages of lower socio-political complexity, such as Ju'h'oan, Yali, and

Tikuna, constructional complexity, as in (3), is more prevalent in languages of higher socio-political complexity, such as French, Hebrew, and Japanese (Gil & Shen, 2019: 3–8). Additional support for the correlation is provided in other cross-linguistic studies of particular grammatical features that are connected in one way or another to the distinction between bare and constructional compositionality. In a typological study of case and tense–aspect–mood markers, two of the major grammatical features contributing to constructional complexity, such markers are argued to be either absent or less frequently occurring in Creole and sign languages — two classes of languages that are relatively newly evolved and generally associated with lower socio-political complexity (Gil, 2014). In yet another study, tense–aspect–mood marking is shown to occur less frequently in languages belonging to smaller language families, whose histories reflect a lower degree of socio-political complexity (Gil, 2021). Finally, in an ongoing study, a range of linguistic features associated with greater morphosyntactic complexity and hence more constructional compositionality are shown to be correlated with an array of features reflective of greater socio-political complexity (Benítez-Burraco et al., 2022). In conjunction, these studies provide good reason to believe that within the relatively recent lead-up to contemporary human languages, compositionality developed along an evolutionary trajectory from bare to compositional, as reflected in Fig. 1.

How, then, does this conclusion fare in face of the variegated instances of constructional compositionality in animal communication surveyed in this paper? The bare compositionality of captive apes such as Kanzi and Chantek may perhaps constitute an initial point of departure for the above-mentioned evolutionary trajectory from bare to constructional compositionality (Gil, 2017). However, when that paper was written, much less was known about compositionality in animal communication; moreover, I was not yet familiar with some of the material that had already been published, supporting the presence of bare and constructional compositionality in primates and in birds. The results of the present survey suggest that compositionality, of both bare and constructional varieties, may have evolved independently in several diverse branches of the animal kingdom, of which the recent branch leading towards modern humans is just one.

However, as they now stand, the results of the present survey provide no direct evidence in favour of the phylogenetic interpretation of Fig. 1 and the evolutionary trajectory from bare to constructional compositionality. They are not inconsistent with it; the problem is just that the evidence of compositionality in the animal world is at present too sparse to support specific evolutionary trajectories pertaining to compositionality in particular branches, such as, for example, Old World monkeys (Cercopithecidae) or tits (Paridae). Thus, for now at least, Fig. 1 should be taken as a hypothesis about how various types of compositionality might have evolved, independently, in each of those species or larger taxa where compositionality has been observed — a possible guide for further investigations.

The results of this paper suggest that there is nothing that special about compositionality. The main if not only reason for the scarcity of compositionality in animal communication systems is the scarcity of its two primary building blocks, a lexicon of meaning-bearing signs, and a rudimentary syntax enabling two or more such signs to be brought together in juxtaposition. In the presence

of these two building blocks, compositionality would seem to be almost inevitable. Specifically, when two forms  $F_1$  and  $F_2$  bearing meanings  $M_1$  and  $M_2$  respectively are juxtaposed, it is only natural for the combination  $F_1 F_2$  to be assigned a meaning derived from  $M_1$  and  $M_2$ , making use of the Association Operator  $A ( M_1 M_2 )$ . In other words, there is nothing specifically human about compositionality. Moreover, once grammatical items and configurational signs are also present, the development of more complex types of compositionality would also appear to be an automatic consequence. Thus, in order to understand how compositionality evolved, in human language and in animal communication systems, we need to understand the evolution of the respective building blocks of compositionality: the lexicon, elementary syntax, grammatical items and configurational signs.

In conclusion, this paper argues (following Schlenker et al., 2016a, 2016b and others) that the study of animal communication, both synchronic and phylogenetic, would benefit from an increased familiarity with the methods and findings of modern linguistics. In particular, it suggests that the analysis of compositionality in animal communication systems might profit from a better appreciation of the variegated types of compositionality present in human languages. However, at the same time, this paper also suggests that linguistics would do well to expand its own horizons and draw inspiration from aspects of animal communication that present clear analogues in the human domain. Many grammarians, in their quest for the complex and baroque, often lose sight of the simpler structures that human languages may share with other animal communication systems, such as, for example, the Association Operator, or to cite a more specific case, the widespread use of repetition to express intensity and urgency, as evidenced, amongst others, in the human boatman calling out for passengers in (10), the alarmed putty-nosed monkey in (34), and the restless Mexican chickadee in (35). Just as biologists have no problem talking about the mitochondria of mice and men, so, this paper has argued, scholars adopting a unified approach to all forms of language and communication may speak, in one and the same breath, of bare and various types of constructional compositionality, as manifest in human languages and systems of animal communication.

**Abbreviation** The interlinear glosses in the examples cited in this paper make use of the Leipzig Glossing Rules and include the following abbreviations:

*CLF*: classifier; *DIM*: diminutive; *DIR*: directional; *INT*: intensifier; *LNK*: linker; *NEG*: negation; *POL*: politeness; *PROX*: proximal; *PST*: past; *SG*: singular; *YNOQ*: yes/no question; *1*: 1st person

**Acknowledgements** This paper would never have come into existence were it not for Bernard Comrie, who, over two decades ago, brought me to the Max Planck Institute for Evolutionary Anthropology in Leipzig, and in doing so introduced me to some of the issues dealt with in this paper. Although my main research focus was and remains human languages, the bookshelves there opened



my eyes to the field of animal communication, while the lecture halls beckoned with exciting presentations by members of the departments of Christophe Boesch and Michael Tomasello. Doing real interdisciplinary work is challenging and fraught with all kinds of pitfalls and potential misunderstandings, but I am grateful to the entire community of the Max Planck Institute for Evolutionary Anthropology for providing the support and motivation for me to pursue this path. Finally, I would like to express my gratitude to the four guest editors of this volume — Natalie Gontier, Stefan Hartmann, Michael Pleyer and Evelina D. Rodrigues — for facilitating my participation in this volume, and to the editors and two anonymous reviewers for their helpful and constructive comments on earlier drafts of this paper.

**Funding** Open Access funding enabled and organized by Projekt DEAL.

## Declarations

**Conflict of Interest** The author declares that he has no conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Arbib, M. A. & Bickerton, D. (eds.) (2010). *The emergence of protolanguage: Holophrasis vs. compositionality*. Philadelphia and Amsterdam: John Benjamins
- Arnold, K., & Zuberbühler, K. (2006). Language evolution: Semantic combinations in primate calls. *Nature*, *441*, 303.
- Arnold, K., & Zuberbühler, K. (2012). Call combinations in monkeys: Compositional or idiomatic expressions? *Brain and Language*, *120*(3), 303–309.
- Arnold, K., & Zuberbühler, K. (2013). Female putty-nosed monkeys use experimentally-altered contextual information to disambiguate the cause of male alarm calls. *PLoS One*, *8*(6), e65660.
- Baggio, G., Lambalgen, M., & Hagoort, P. (2012). The processing consequences of compositionality. In M. Werning, W. Hinzen, & E. Machery (Eds.), *The Oxford Handbook of Compositionality* (pp. 655–672). Oxford University Press.
- Benítez-Burraco, A., Candy C., Chen, S., Gil, D., Progovac, L., Reifegerste J. & Tatarinova, T. (2022). Cognitive and genetic correlates of a single macro-parameter of crosslinguistic variation, the evolution of language, Proceedings of the 14th International Conference (EVOLANG14)
- Blevins, J. P. (2016). *Word and Paradigm Morphology*. Oxford University Press.
- Boesch, C. (1991). Symbolic communication in wild chimpanzees? *Human Evolution*, *6*, 81–89.
- Boye, K., & Harder, P. (2012). A usage-based theory of grammatical status and grammaticalization. *Language*, *88*(1), 1–44.
- Candiotti, A., Zuberbühler, K., & Lemasson, A. (2012). Context-related call combinations in female Diana monkeys. *Animal Cognition*, *15*, 327–339.
- Chomsky, N. (1995). *The Minimalist Program*. MIT Press.
- Clarke, E., Reichard, U. H., & Zuberbühler, K. (2006). The syntax and meaning of wild gibbon songs. *PLoS One*, *1*(1), e73. <https://doi.org/10.1371/journal.pone.0000073>
- Coye, C., Zuberbühler, K., & Lemasson, A. (2016). Morphologically structured vocalizations in female Diana monkeys. *Animal Behaviour*, *115*, 97–105.
- Croft, W. (2003). *Typology and universals* (2nd ed.). Cambridge University Press.



- Diamond, J. (2012). *The world until yesterday, what can we learn from traditional societies?* Viking.
- Edwards, O. (2020). *Metathesis and unmetathesis in Amarsi*. Language Science Press.
- Engesser, S., Ridley, A. R., & Townsend, S. W. (2016). Meaningful call combinations and compositional processing in the southern pied babbler. *PNAS*, *113*(21), 5976–5981.
- Ficken, M. S., Hailman, E. D., & Hailman, J. P. (1994). The Chick-A-Dee call system of the Mexican chickadee. *The Condor*, *96*, 70–82.
- Fletcher, J., Stirling, L., Mushin, I., & Wales, R. (2002). Intonational rises and dialog acts in the Australian English map task. *Language and Speech*, *45*(3), 229–253.
- Frege, G. (1963). Compound thoughts. *Mind, A Quarterly Review of Psychology and Philosophy*, *72*(285), 1–17. (Translated by R.H. Stoothoff.)
- Gabrić, P. (2021). Overlooked evidence for semantic compositionality and signal reduction in wild chimpanzees (Pan troglodytes). *Animal Cognition*, *25*, 631–643.
- Gardner, R. A., Gardner, B. T., & Van Camfort, T. E. (Eds.). (1989). *Teaching sign language to chimpanzee*. SUNY Press.
- Gentner, T. Q., Fenn, K. M., Margoliash, D., & Nusbaum, H. (2006). Recursive syntactic pattern learning by songbirds. *Nature*, *440*(7088), 1204–1207.
- Gil, D. (2001). Escaping Eurocentrism: Fieldwork as a process of unlearning. In P. Newman & M. Ratliff (Eds.), *Linguistic Fieldwork* (pp. 102–132). Cambridge University Press.
- Gil, D. (2005a). From repetition to reduplication in Riau Indonesian. In B. Hurch (Ed.), *Studies on reduplication (Empirical approaches to language typology 28)* (pp. 31–64). Mouton de Gruyter.
- Gil, D. (2005b). Word order without syntactic categories: How Riau Indonesian does it. In A. Carnie, H. Harley, & S. A. Dooley (Eds.), *Verb first: On the syntax of verb-initial languages* (pp. 243–263). John Benjamins.
- Gil, D. (2007). Creoles, complexity & associational semantics. In U. Ansaldi & S. J. Matthews (Eds.), *Deconstructing Creole: New horizons in language creation* (pp. 67–108). John Benjamins.
- Gil, D. (2008). How complex are isolating languages? In M. Miestamo, K. Sinnemäki, & F. Karlsson (Eds.), *Language Complexity: Typology, Contact, Change* (pp. 109–131). John Benjamins.
- Gil, D. (2012). Where does predication come from? *Canadian Journal of Linguistics*, *57*, 303–333.
- Gil, D. (2014). Sign languages, Creoles, and the Development of Predication. In F. J. Newmayer & L. Preston (Eds.), *Measuring Grammatical Complexity* (pp. 37–64). Oxford University Press.
- Gil, D. (2015). The Mekong-Mamberamo linguistic area. In N. J. Enfield & B. Comrie (Eds.), *Languages of Mainland Southeast Asia, The State of the Art* (pp. 266–355). Pacific Linguistics, DeGruyter Mouton.
- Gil, D. (2017). Isolating-monocategorical-associational language. In H. Cohen & C. Lefebvre (Eds.), *Handbook of categorization in cognitive science* (2nd ed., pp. 471–510). Elsevier.
- Gil, D. (2021). Tense-aspect-mood marking, language family size, and the evolution of predication. *Philosophical Transactions B*, *376*, 20200194.
- Gil, D. (2022). Number in Indonesian. In P. Acquaviva & M. Daniel (Eds.), *Number in the world's languages* (pp. 457–503). De Gruyter.
- Gil, D., & Shen, Y. (2019). How Grammar Introduces Asymmetry into Cognitive Structures: Compositional Semantics, Metaphors and Schematological Hybrids. *Frontiers in Psychology - Language Sciences*. <https://doi.org/10.3389/fpsyg.2019.02275>
- Girard-Buttoz, C., Zaccarella, E., Bortolato, T., Friederici, A. D., Wittig, R. M., & Crockford, C. (2022). Chimpanzees produce diverse vocal sequences with ordered and recombinatorial properties. *Communications Biology*, *5*, 410. <https://doi.org/10.1038/s42003-022-03350-8>
- Gleitman, L. R., Connolly, A. C., & Armstrong, S. L. (2012). Can prototype representations support composition and decomposition? In M. Werning, W. Hinzen, & E. Machery (Eds.), *The Oxford Handbook of Compositionality* (pp. 418–436). Oxford University Press.
- Greenfield, P. M., & Savage-Rumbaugh, S. (1990). Grammatical combination in pan paniscus: Processes of learning and invention in the evolution and development of language. In S. T. Parker & K. R. Gibson (Eds.), *“Language” and Intelligence in Monkeys and Apes, Comparative Developmental Perspective* (pp. 540–578). Cambridge University Press.
- Griesser, M., Wheatcroft, D., & Suzuki, T. N. (2018). From bird calls to human language: Exploring the evolutionary drivers of compositional syntax. *Current Opinion in Behavioural Sciences*, *21*, 6–12.
- Hailman, J. P., & Ficken, M. S. (1986). Combinatorial animal communication with computable syntax: Chick-a-dee calling qualifies as “language” by structural linguistics. *Animal Behaviour*, *34*(6), 1899–1901.

- Hailman, J. P., Ficken, M. S., & Ficken, R. W. (1985). The “chick-a-dee” calls of *Parus atricapillus*: A recombinant system of animal communication compared with written English. *Semiotica*, 56(3/4), 191–224.
- Hale, K. (1983). Warlpiri and the grammar of non-configurational languages. *Natural Language and Linguistic Theory*, 1, 5–47.
- Hampton, J. A., & Jönsson, M. L. (2012). Typicality and compositionality: The logic of combining vague concepts. In M. Werning, W. Hinzen, & E. Machery (Eds.), *The Oxford handbook of compositionality* (pp. 385–402). Oxford University Press.
- Haspelmath, M. (2021). Bound forms, welded forms, and affixes: Basic concepts for morphological comparison. *Voprosy Jazykoznanija*, 2021(1), 7–28.
- Heine, B., Claudi, U., & Hünnemeier, F. (1991). *Grammaticalization: A conceptual framework*. Chicago University Press.
- Heine, B., & Kuteva, T. (2002). *World lexicon of grammaticalization*. Cambridge University Press.
- Heine, B., & Reh, M. (1984). *Grammaticalization and reanalysis in African languages*. Buske
- Jackendoff, R., & Wittenberg, E. (2014). What you can say without syntax: A hierarchy of grammatical complexity. In F. J. Newmeyer & L. B. Preston (Eds.), *Measuring linguistic complexity* (pp. 65–82). Oxford University Press.
- Jackendoff, R., & Wittenberg, E. (2017). Linear grammar as a possible steppingstone in the evolution of language. *Psychonomic Bulletin and Review*, 24(1), 219–224.
- Kayne, R. (1984). *Connectedness and binary branching*. Foris.
- Kess, J. F. & Anita M. Copeland (1984). The structure and function of Nootkan baby talk. *Working Papers of the Linguistic Circle of the University of Victoria* 4(1), 8–22
- Kuhn, J., Sumir Keenan, S., Arnold, K., & Lemasson, A. (2014). On the /-oo/ ‘suffix’ of Campbell’s monkeys (*C. Campbelli*). Ms. <http://jeremykuhn.net/papers/Kuhn-oo-suffix-10.2014.pdf>.
- Lakoff, R. (1973). Language and women’s place. *Language in Society*, 2(1), 45–80.
- Lehmann, C. (2002). *Thoughts on Grammaticalization* (2nd ed.). Seminar für Sprachwissenschaft der Universität.
- Löbner, S. (2012). Rules in Compositional Semantics. In M. Werning, W. Hinzen, & E. Machery (Eds.), *The Oxford handbook of compositionality* (pp. 220–241). Oxford University Press.
- Lowry, O. (2011). Belfast intonation and speaker gender. *Journal of English Linguistics*, 39(3), 209–232.
- Miles, H. L. (1990). The Cognitive Foundations for Reference in a Signing Orangutan. In S. T. Parker & K. R. Gibson (Eds.), *“Language” and Intelligence in Monkeys and Apes, Comparative Developmental Perspectives* (pp. 511–539). Cambridge University Press.
- Outtara, K., Lemasson, A., & Zuberbühler, K. (2009a). Campbell’s monkeys use affixation to alter call meaning. *PLoS One*, 4(11), e7808.
- Outtara, K., Lemasson, A., & Zuberbühler, K. (2009b). Campbell’s monkeys concatenate vocalizations into context-specific call sequences. *PNAS*, 106(51), 22026–22031.
- Pagin, P. (2011). Compositionality, complexity, and evolution. PERILUS 2011, Symposium on Language Acquisition and Language Evolution. The Royal Swedish Academy of Sciences and Stockholm University, pp. 51–62.
- Partee, B. H. (1984). Compositionality. In F. Landman & F. Veltman (Eds.), *Varieties of formal semantics (number 3 in ‘GRASS’)* (pp. 281–311). Foris.
- Preece, K., & Beekman, M. (2014). Honeybee waggle dance error: Adaption or constraint? unravelling the complex dance language of honeybees. *Animal Behavior*, 94, 19–26. <https://doi.org/10.1016/j.anbehav.2014.05.016>
- Riley, J. R., Greggers, U., Smith, A. D., Reynolds, D. R., & Menzel, R. (2005). The flight paths of honeybees recruited by the waggle dance. *Nature*, 435, 205–207. <https://doi.org/10.1038/nature03526>
- Sabaté, L. N., Mesbahi, G., Dezechache, G., César, C., Zuberbühler, K., & Berthet, M. (2022). Animal linguistics in the making: The Urgency Principle and titi monkeys’ alarm system. *Ethology, Ecology and Evolution*, 34(3), 378–394.
- Schel, A. M., Candiotti, A., & Zuberbühler, K. (2010). Predator-detering alarm call sequences in *Guereza colobus* monkeys are meaningful to conspecifics. *Animal Behaviour*, 80(5), 799–808.
- Schlenker, P., Chemla, E., Arnold, K., Lemasson, A., Outtara, K., Keenan, S., Stephan, C., Ryder, R., & Zuberbühler, K. (2014). Monkey semantics: Two “dialects” of Campbell’s monkey alarm calls. *Linguistics and Philosophy*, 37, 439.
- Schlenker, P., Chemla, E., Arnold, K., & Zuberbühler, K. (2016a). Pyow-Hack revisited: Two analyses of putty-nosed monkey alarm calls. *Lingua*, 171, 1–23.

- Schlenker, P., Chemla, E., Schel, A. M., Fuller, J., Gautier, J.-P., Kuhn, J., Veselinović, D., Arnold, K., Cēsar, C., Keenan, S., Lemasson, A., Ouattara, K., Robin Ryder, R., & Zuberbühler, K. (2016b). Formal monkey linguistics. *Theoretical Linguistics*, 42(1–2), 1–90.
- Schlenker, P., Chemla, E., Cēsar, C., Ruder, R., & Zuberbühler, K. (2017). Titi semantics: Context and meaning in Titi monkey call sequences. *Natural Language and Linguistic Theory*, 35, 271–298. <https://doi.org/10.1007/s11049-016-9337-9>
- Shen, Y., & Gil, D. (2017). How language influences the way we categorize hybrids. In H. Cohen & C. Lefebvre (Eds.), *Handbook of categorization in cognitive science* (2nd ed., pp. 1177–1200). Elsevier.
- Stephan, C., & Zuberbühler, K. (2008). Predation increases acoustic complexity in primate alarm calls. *Biology Letters*, 4, 641–644.
- Suzuki, R., Buck, J., & Tyack, P. (2006). Information entropy of humpback whale songs. *Journal of the Acoustical Society of America*, 119(3), 1849–1866.
- Suzuki, T. N. (2014). Communication about predator type by a bird using discrete graded and combinatorial variation in alarm calls. *Animal Behaviour*, 87, 59–65.
- Suzuki, T. N. (2021). Animal linguistics: Exploring referentiality and compositionality in bird calls. *Ecological Research*, 36(2), 221–231.
- Suzuki, T. N., Griesser, M., & Wheatcroft, D. (2019). Syntactic rules in avian vocal sequences as a window into the evolution of compositionality. *Animal Behaviour*, 151, 267–274.
- Szabó, Z. G. (2020). Compositionality. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Fall 2000 edition). Stanford, Metaphysics Research Lab
- Tanner, J. E. (2004). Gestural phrases and gestural exchanges by a pair of zoo-living Lowland gorillas. *Gesture*, 4(1), 1–24.
- Tanner, J. E., & Perlman, M. (2016). Moving beyond meaning: Gorillas combine gestures into sequences for creative display. *Language and Communication*, 54, 56–72.
- Templeton, C. N., Greene, E., & Davis, K. (2005). Allometry of alarm calls: Black-capped chickadees encode information about predator size. *Science*, 308(5730), 1934–1937.
- Trask, R. L. (1993). *A dictionary of grammatical terms in linguistics*. Routledge.
- Veselinović, D., Candiotti, A., & Lemasson, A. (2014). *Female Diana monkeys (Cercopithecus Diana) have complex calls*. New York University.
- Von Frisch, K. (1967). *The dance language and orientation of bees*. Harvard University Press.
- Warren, P. (2016). *Uptalk: The phenomenon of rising intonation*. Cambridge University Press.
- Werning, M. (2012). Non-symbolic compositional representation and its neuronal foundation: Towards an emulative semantics. In M. Werning, W. Hinzen, & E. Machery (Eds.), *The Oxford handbook of compositionality* (pp. 633–654). Oxford University Press.
- Wheatcroft, D. (2015). Repetition rate of calls used in multiple contexts communicates presence of predators to nestlings and adult birds. *Animal Behaviour*, 103, 35–44.
- Zimmermann, T. E. (2012). Compositionality problems and how to solve them. In M. Werning, W. Hinzen, & E. Machery (Eds.), *The Oxford handbook of compositionality* (pp. 81–106). Oxford University Press.