★ FORUM ★

More on the Relations among Categorization, Merge and Labeling, and Their Nature

Koji Hoshi

1. Introduction

This opinion piece is intended as a rejoinder to/extension of my own previous proposal in Hoshi (2018), critically reexamining the hypothesis of categorization origin of Merge and labeling in an attempt to rectify mistakes while aiming at theoretical refinement. In the ensuing discussion, I will address the following three claims made in Hoshi (2018):

(i) Categorization has two modes of operation, the interrelational mode of categorization (*IntCat*) and the differentiational mode (*DifCat*); Merge is derived by 'exapting/co-opting' the recursive set-formation sub-component of IntCat.[1]

(ii) Categorization targets either a series of entities or a series of category sets.

(iii) Merge is derived by detaching the labeling sub-component of IntCat.

Concerning (i), I will claim that the two apparently distinct modes of operation of categorization stem from differences of 'extraction patterns' and thus as a precursor of Merge the particular mode of categorization such as IntCat should not be stipulated.

As for (ii), I will make clear the relation between the low-order categorization involving a series of entities and the higher-order categorization involving a series of category sets in humans and non-human animals, in connection with the qualitative difference of the two types of categorization between them in the context of the evolution of human language.

Part of the material in this piece was presented in earlier slightly different forms at the Evolinguistics Workshop 2019 (University of Tokyo, 25 May 2019) and at the International Cognitive Linguistics Conference 15 (Kwansei Gakuin University, 9 August 2019). I would like to express my gratitude to the audiences on those occasions for valuable questions and comments, particularly Koji Fujita, Shin-ichi Tanaka, Daiki Matsumoto, Toshiyuki Kumashiro, and Sotaro Kita. I am also very grateful to editor-in-chief Kleanthes K. Grohmann and other associate editors of *Biolinguistics* for extremely useful suggestions on the revision of the earlier version. All the remaining errors and inadequacies are, of course, solely my own responsibility. This work was supported by MEXT/JSPS Grant-in-Aid for Scientific Research on Innovative Areas #4903 (Evolinguistics), Grant Number JP17H06379.

[1] Note that this pattern indeed fits with *exaptation* in that the recursive set-formation operation that originally had the function for implementing categorization subsequently came to serve another function for implementing Merge in syntax (see Gould & Vrba 1982).

ISSN 1450–3417
Biolinguistics 13: 001–021, 2019
http://www.biolinguistics.eu

Finally, with regard to (iii), I will make a specific claim about the relation between labeling in categorization and labeling in syntax, proposing a possible novel explanation of the binarity of Merge due to the very nature of the origin of labeling for syntactic objects from labeling for categorization.[2]

2. Categorization Origins of Merge and Labeling Reconsidered

2.1. No Need for Specifying the Particular Mode of Categorization IntCat as a Precursor of Merge

The hypothesis of categorization origin of Merge advanced in Hoshi (2018) is based on a reinterpretation of Lenneberg's view of biological evolution of the function of language from a perspective of modern linguistic theorizing. Lenneberg (1967) states that "the cognitive function underlying human language consists of an adaptation of a ubiquitous process (among vertebrates) of categorization and extraction of similarities"[3] (p. 374). Furthermore, with respect to categorization, Lenneberg originally remarks as follows:

> [M]ost animals organize the sensory world by a process of *categorization*, and **from this basic mode of organization two further processes derive: *differentiation* or discrimination, and interrelating of categories or the perception of and tolerance for *transformations*** […]. In man these organizational activities are usually called concept-formation; but it is clear that there is no formal difference between man's concept-formation and animal's propensity for responding to categories of stimuli. There is, however, a substantive difference. The total possibilities for categorization are clearly not identical across species. (Lenneberg 1967: 331; emphasis mine—KH)

First of all, observe *Figure 1* below, which illustrates schematically two patterns of categorization (see also Hoshi 2018: 41). There are two sub-categories, C2 and C3, which are contained in a more comprehensive super-category, C1. One pattern of categorization implicit in this figure is that the two sub-categories are subsumed under the super-category; in turn, the other pattern concealed here is that the more inclusive super-category C1 is subdivided into the two more specific sub-categories C2 and C3.

---

2 In the field of cognitive linguistics, Langacker (1997) proposes the cognitive operation for symbolic assembly called 'Grouping', which is virtually on a par with Merge in the minimalist program (Chomsky 1995 *et seq*.). There are, however, some differences between Merge and Grouping: (i) while Merge is binary, Grouping is not necessarily thus restricted, and (ii) Merge allows for both external and internal variety, whereas Grouping permits only external variety, meaning that it lacks movement (see Langacker 1991). On the other hand, there is a common property in Merge and Grouping: Both are characterized as label-free (see Chomsky 2013, 2015 for Merge and see Langacker 1997 for Grouping). I will put aside a systematic comparison between Merge and Grouping in this opinion piece, leaving it to future research.

3 The term 'adaptation' in the quote does not mean adaptation to the environment but refers to structural innovations on a DNA molecular level (see Lenneberg 1967: 72). An investigation into the genetics of this adaptation is of course beyond the scope of this opinion piece, but I will speculate what kind of qualitative change in the ubiquitous process of categorization could have occurred in our species in the following discussion.

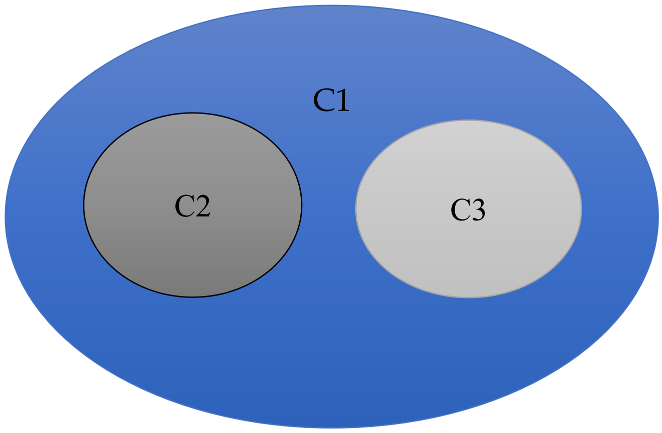

Figure 1: Two Patterns of Categorization

As already stated at the outset, in Hoshi (2018), I called the former 'inter-relational mode of categorization' (IntCat) and the latter 'differentiational mode of categorization' (DifCat). I proposed that Merge was derived from IntCat by 'exapting/co-opting' the set-formation sub-component of IntCat and detaching the labeling sub-component of it, while preserving the ability of IntCat *per se* in the course of the biological evolution of human language in our species (see Hoshi 2018: 47).[4]

In Hoshi (2018), I did not address the relation between the two modes of categorization and two distinct extraction patterns. Recall Lenneberg's (1967) thesis that "the cognitive function underlying human language consists of an adaptation of a ubiquitous process (among vertebrates) of **categorization** and **extraction of similarities**" mentioned above (my emphasis—KH).[5] Strictly speaking, extraction of similarities does not hold for categorization in general. Only IntCat is carried out on the basis of extraction of similarities. Hence, in *Figure 1*, IntCat of C2 and C3 to form C1 will be done by extracting similarities between C2 and C3.

What about the case of DifCat? In order to differentiate an undifferentiated general category into more specific sub-categories, extraction of differences rather than extraction of similarities holds a key role. Thus, in *Figure 1*, DifCat of C1 to form C2 and C3 will be done by extracting differences between C2 and C3. Notice, however, that extraction of differences is just the opposite side of the same coin of extraction of similarities in that, to the extent that one cannot detect similarities between X and Y, one can assume that X and Y are different.

Accordingly, as long as the capacity for extracting both similarities and differences is endowed with, along with the ability for categorization (*Cat* for short), the two different modes of Cat (i.e. IntCat and DifCat) will result in con-

---

4 As briefly touched on in Hoshi (2018), it is well known in biological evolution that, in the case of gene duplication, genes, chromosomes, or whole genomes will be duplicated and the duplicated element/region or its original element/region can drift to (a) new function(s) (see e.g. Ohno 1970, Zhang 2003). I am not claiming that this is the only possibility for this particular hypothesized biological evolutionary event in our species. This issue needs to be pursued in genetics and neurobiology in the future, while always keeping other possibilities in mind.

5 See also Sloutsky (2003) for some discussion on the role of similarity in the development of categorization.

junction with extraction of similarities and extraction of differences, respectively. In a sense, what I argue for is a stricter reinterpretation of Lenneberg's (1967) view that "the cognitive function underlying human language consists of an adaptation of a ubiquitous process (among vertebrates) of categorization and extraction of similarities" (p. 374), regarding extraction of differences as the result of failure to extract similarities.

In light of the consideration thus far, without mentioning the particular mode of IntCat, I will submit (unlike in Hoshi 2018) that Merge as a label-free recursive set-formation operation was derived from the inherently labeled recursive set-formation complex cognitive operation of Cat, by 'exapting/co-opting' the recursive set-formation sub-component operation of Cat, detaching the labeling sub-component operation of it, while preserving the cognitive ability of Cat *per se*, in the course of biological evolution of the language capacity in our species.

2.2. *Low-Order Categorization and Higher-Order Categorization in Non-Human Animals and Humans: A More Fine-Grained Classification of Categorization*

In discussing categorization in Hoshi (2018), in particular *IntCat*, I provided a preliminary definition of the label κ for *IntCat* as a sort of characteristic function, as shown in (1) (see Hoshi 2018: 42):

(1) $$\kappa(x) = \begin{cases} 1 \text{ if } x \in \kappa \\ 0 \text{ if } x \notin \kappa \end{cases}$$

For any element indicated by x, x either 'satisfies' the label κ or not. Then, I formulated *IntCat* as an unordered set-formation under a particular label specified by κ, as shown in (2) (see Hoshi 2018: 42):

(2) $IntCat^{\kappa}(x_1, \ldots, x_n) = \{ x_1, \ldots, x_n \}$ ($x_i \in \kappa$, $1 \leq i \leq n$)

(x_i is a target element for interrelational categorization and κ is a label, where the sequence in the set uniformly contains either a series of entities or a series of sets as the value of x_i)

While I acknowledged the existence of two kinds of *IntCat* in Hoshi (2018), as reflected in the expressions "either a series of entities or a series of sets" in (2), I did not provide any empirical rationale for postulating the two kinds of *IntCat* in this fashion and did not fully develop a more fine-grained classification of categorization. Hence, I will aim to fill this gap in the ensuing discussion.

Since, as already argued above, reference to the particular mode of categorization *IntCat* is not necessary, I will keep using the general term *categorization* (*Cat*) in what follows in addressing a more fine-grained classification of Cat in terms of the degree of order in categorization in non-human animals and humans, linking it to the context of biological evolution of language capacity.[6]

---

[6] As Lenneberg (1967) already pointed out clearly, there are closely interrelated but in principle separate goals for the investigation into the phylogeny of human language: (i)

By way of the degree of order in categorization, I would like to differentiate between *low-order categorization* (*low-order Cat*) and *higher-order categorization* (*higher-order Cat*) as a first approximation. For ease of exposition, I will look at cases where only two individual instances or two sub-categories are involved, but as the formulation of Cat in (2) above shows, in principle, Cat is not restricted to being binary unlike Merge (see Hoshi 2018 for some relevant discussion).

First, observe the following *Figure 2* on low-order Cat.

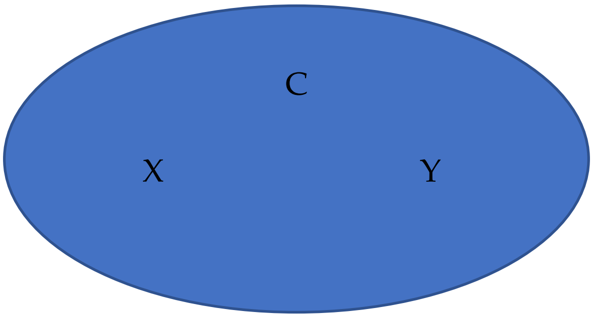

Figure 2: Low-order Cat (X, Y = individual instances/tokens; C= category label)

What I mean by 'low-order Cat' refers to the point that there are no categories of categories involved and only categorization targeting individual objects (i.e. instances/tokens), including events, is at work. Two different individual instances/tokens X and Y are interrelated under the category label C in *Figure 2*. For example, suppose X and Y correspond to two dogs as individual instances/ tokens, which are categorized under the label C (= the concept *DOG* or the proto-concept *DOG*).[7]

Hurford (2007: 12) remarks that "the advent of language does change the nature and range of concepts, but I will use the term concept broadly enough to attribute concepts to some animals"—differentiating the following three stages of concepts in terms of an evolutionary succession: proto-concepts → pre-linguistic concepts → (post-)linguistic concepts. I will embrace this kind of distinction of 'levels' of concepts in the following discussion. Notice, however, that the distinction between low-order Cat and higher-order Cat, which is relevant to both non-human animals and humans, is independent of such different 'levels' of concepts in Hurford (2007).

Lenneberg (1967) already suggested that low-order Cat seems to be shared among (at least) all the vertebrates, and it has been suggested in the literature

---

biological evolution of *language capacity* (see also Chomsky 2017), and (ii) cultural evolution of *language* (see also Tomasello 1999). Obviously, ontogeny of language enters into both the biological evolution of language capacity in (an) individual(s) among our hominin ancestors and the subsequent cultural evolution of language responsible for linguistic particularization and diversification in their own language community and those of their descendants (e.g., development of particular phonemic contrasts, conventional use of particular forms of grammatical constructions, grammaticalization, vocabulary expansion, etc.). I will set aside this important aspect of evolution of language in this contribution.

7 Regarding the term *proto-concept*, I will follow Hurford (2007) in assuming that it refers to whatever is instantiated in the brain that will enable animals to display "regular and systematic behavior in connection with some class of entities from the environment" (Hurford 2007: 16) and is related to "reflex actions" (Hurford 2007: 18).

that non-human animals are only capable of engaging themselves to low-order Cat, by which they will only categorically respond to individual objects (i.e. instances/tokens), including events, on the basis of sensory input stimuli from the immediate environment.

For instance, Bickerton (1990) addresses a typical example of low-order Cat in non-human animals. In animals such as the vervet monkey with its alarm calls, the spots, roar, and smell of different perceptions of individual leopards as its predators all provoke a similar result of flight, which leads to the formation of a category 'leopard', which is a cognitive linking of the three different sensory modalities for the same behavioral response of flight (vision, audition, olfaction) (see also Bouchard 2013: Chaps. 4–5). I believe that this cognitive linking in the brain can be regarded as serving as the 'label' for low-order Cat in non-human animals, and presumably our non-linguistic hominin ancestors as well as modern humans, for that matter (see Mareschal et al. 2010 for a comprehensive discussion on concepts and categorization in both humans and non-human animals).

Note that non-human animals do not and presumably our non-linguistic hominin ancestors did not possess the word *leopard per se* in their communication systems. At this point, in order to avoid confusion, let me hasten to add a caveat. As claimed in Hoshi (2018), I will keep distinguishing between ***labeling*** of categories and ***naming*** of categories. In fact, the distinction between these two processes can be traced back to Lenneberg (1967), who submits that the formation of a concept for categorization, which is equivalent to labeling of categories in my term, is prior to and more primitive than the attachment of 'words' (sound/sign patterns) to certain types of conceptualization, which is equivalent to naming of categories in my terms.[8] In human language, concepts as the labels of a subset of all the possible categories in conceptualization are also 'named' and used for thought or communication.

Thus, inasmuch as categories are shared and used for communication among conspecifics, it is expected that ***naming*** of categories is also to be observed among non-human animals to some degree. Alarm calls in monkeys in stimulus–response behavior in animal communication systems might be a case in point (see, e.g., Seyfarth et al. 1980).

Next, look at *Figure 3* on higher-order Cat. Two separate categories C1 and C2 are subsumed under the super-category label C3. This is a case of higher-order Cat, which implicates categorization of categories rather than categorization of individual objects/events.

[8] See also Murphy (2010) and the other papers in Mareschal et al. (2010) for a detailed discussion on human and non-human concepts.

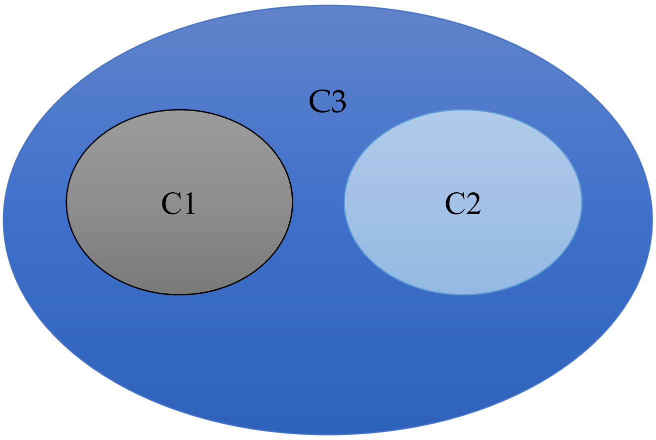

Figure 3: Higher-order Cat (C1, C2 = category labels for sub-categories; C3 = category label for super-category)

Let me illustrate this by using a simple example of part of taxonomic categories of foods (see also Hoshi 2018). Suppose you have formed a low-order category with the label 'lemon' indicated by C1, and likewise formed another low-order category with the label 'tangerine' indicated by C2. Then you will form a super-category with the label 'citrus fruits' indicated by C3 by interrelating the two sub-category sets on the basis of extraction of similarities between C1 and C2. This is an example of higher-order Cat, because here only categorization of categories is in operation.

Spinozzi et al. (1999), Conway & Christiansen (2001), Penn et al. (2008), and Bouchard (2013), among others, state that only humans can engage in such higher-order Cat, that is, categorization of categories, beyond categorization of individual objects including events. On the other hand, Roberts & Mazmanian (1988) and Vonk & McDonald (2002, 2004) report that pigeons, monkeys, gorillas, and orangutans are also capable of dealing with higher-order Cat to some extent in addition to low-order Cat (see also Vauclair 2002 for more discussion on both low-order and higher-order Cat in non-human primates).

Furthermore, as a concrete illustration of higher-order Cat in non-human animals, Blumstein (2007) and Rundus et al. (2007) report that mammals such as California ground squirrels not only distinguish between their predator snakes and the other stimuli in the external world but also differentiate rattlesnakes and gopher snakes, displaying different defensive behaviors to the two kinds of snakes.

Yang (2013) provides a statistically robust demonstration that, while young children use a rule-based combinatorial power of grammar even at the earliest stages of language acquisition, primates lack the expected range of such a combinatorial grammar, which is based on a close examination of the sign combinations used by an ASL-taught chimpanzee named Nim Chimpsky, reported in Terrace et al. (1979). This can be naturally interpreted as suggesting that primates lack Merge unlike human children. At the same time, however, he acknowledges that it is well-documented that non-human primates have the ability to learn word-like symbolic units, citing Savage-Rumbaugh et al. (1998) and Lyn et al.

(2011). This seems to suggest that although primates lack Merge, they have the ability of Cat with the power of both low-order and some limited degree of higher-order. Note that to the extent that primates can make use of Cat, they can associate word-like units such as signs or artificial physical objects like plastic chips with the available (proto-)concepts as labels of Cat in a symbolic fashion in the process of naming of categories, although the link between such symbols for externalization and their associated (proto-)concepts is presumably not mediated by Merge and the syntactic objects (SOs) generated by it.

Given that the possibility of higher-order Cat in non-human animals is not flatly denied, as discussed above, I will tentatively take the position that in principle at least some degree of higher-order Cat is available for non-human animals as well.

Let us next turn to low-order and higher-order Cat in humans. At least as far as our species is concerned, insofar as the emergence of Merge was critical in our language and cognition, it is plausible to infer that it also had a great impact on Cat (both low-order and higher-order) in 'post-Merge' humans. In particular, I surmise that it made a significant contribution to diversifying a variety of concepts (both simplex and complex) to be employed as labels for Cat in them.

One such contribution by Merge was its generating SOs as an infinite variety of combinations of concepts as lexical items, which were/have been employed as novel labels not only for low-order Cat but also for higher-order Cat in our species. For instance, the higher-order Cat in *Figure 3*, the label/name for the super-category 'citrus fruits' C3 that subsumes the sub-categories such as 'lemon' C1 and 'tangerine' C2 was created by merging a concept corresponding to 'citrus' (i.e. the property of producing juicy fruits with a slightly sour taste) and another concept corresponding to 'fruit'.

Another possible form of contribution by Merge was its making available the formation of a variety of label-free root elements and categorizers as conceptual atoms as target elements of Merge, which lies behind the rich human language lexicon.[9] For example, in *Figure 3*, suppose that C1 is the category with the label/name 'mother' and C2 is the category with the label/name 'father', where 'mother' and 'father' can be analyzed as something like {*n*, {√*female*, √*parent*}} and {*n*, {√*male*, √*parent*}}, respectively (*n* is a nominal categorizer; √*female*, √*male*, and √*parent* are category-free root elements). Then one can form the super-category C3 with the label/name 'parent'. Note that if *n* and √*parent* are merged, one would obtain {*n*, √*parent*}, which would give us the nominal 'parent' to be used for the label/name 'parent' for the higher-order Cat under discussion.

I speculate that this kind of 'lexical decomposition' is possible due to the label-free nature of the 'Merge formula': Merge (X, Y) = {X, Y}. If you flip-flop it, you would obtain the 'inverted Merge formula': {X, Y} = Merge (X, Y). In analyzing primary linguistic data (PLD), human children could employ the schema expressed by the inverted Merge formula in forming lexical items X and Y, where X and Y are either root elements or categorizers.

---

[9] See, e.g., Marantz (1997), Embick & Noyer (2007), Embick & Marantz (2008), Nóbrega & Miyagawa (2015), and Fujita (2017) for the treatment of word formation in the framework of Distributed Morphology (DM).

Accordingly, in this way, the evolutionary way to human-unique Cat and cognition was opened and paved. Thus, this affected the very nature of Cat in our species in a fundamental fashion in such a way that Merge-related higher-order as well as low-order Cat became available in our species thanks to the advent of Merge, leading to what Tattersall (2017: 64) calls "the qualitative distinctness of both modern symbolic cognition and language."

In closing this section, modifying and generalizing Hoshi's (2018) preliminary definition of IntCat, I will define Cat as a sort of recursive set-formation operation based on a characteristic function κ for a label of Cat. The category label is a characteristic function because it will sort out potential target objects, depending upon whether they 'satisfy' relevant concepts as labels or not. Also, Cat is a recursive set-formation operation, given that its output element (i.e. a category set) can serve as its input element(s) as well to the extent that an appropriate label for the Cat in question is available.

Suppose that κ is the variable of a label for Cat, then κ can be taken as a sort of characteristic function that applies to any element indicated by x or X (where x is the variable for an entity and X is the variable for a category set of entities or a category set of category sets) that either 'satisfies' the label or not, as defined as follows:[10]

(3)

$$\kappa(x) \text{ or } \kappa(X) = \begin{cases} 1 \text{ if x or X satisfies } \kappa \\ 0 \text{ if x or X does not satisfy } \kappa \end{cases}$$

On the basis of this characteristic function of the label κ, Cat for low-order and higher-order can be defined as a sort of recursive set-formation operation that takes two or more entities as in (4) or two or more category sets of entities or two or more category sets of category sets as in (5) and that yields one super-set category with an appropriate category label:

(4) ***Low-order Cat***$^{\kappa}$ $(x_1, \ldots, x_n) = X = \{x_i \mid \kappa(x_i)\}$ $(1 \leq i \leq n)$

[informally $\{x_1, \ldots, x_n\}$ $(1 \leq i \leq n)$]

(The sequence in the resultant category set uniformly contains a series of different entities as values of x_i.)

(5) ***Higher-order Cat***$^{\kappa}$ $(X_1, \ldots. X_n) = \boldsymbol{X} = \{X_i \mid \kappa(X_i)\}$ $(1 \leq i \leq n)$

[informally $\{X_1, \ldots, X_n\}(1 \leq i \leq n)$]

(The sequence in the resultant category set uniformly contains a series of different category sets as values of X_i.)

---

10 Here, I am making a sort of idealization with respect to the label variable κ that serves as a characteristic function for categorization (Cat). In reality, as is well-known in the fields of psychology and cognitive linguistics, the criteria for membership determination/identification for categories on the basis of extraction of similarities among category set members is flexible, nuanced and complicated (see among others Rosch 1973, Lakoff 1987, and Taylor 2003). See also Lenneberg (1967: Chap. 8) for some discussion on this sort of characteristics of the classification criteria for categorization.

In summary, I will draw as *Figure 4* the picture of the evolutionary relation between Cat and Merge that I have argued for in this section.

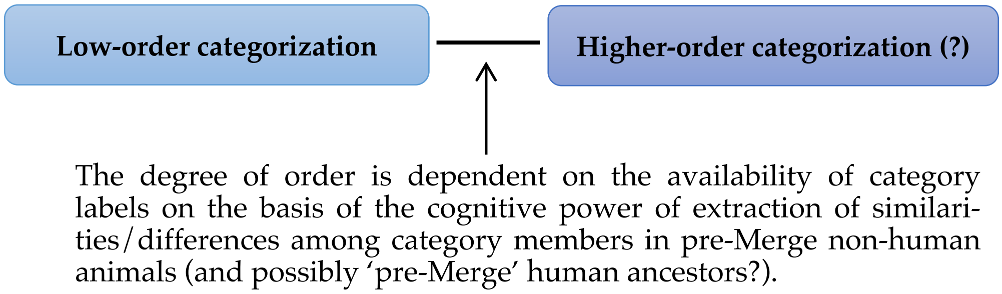

Figure 4: Low-order Cat & higher-order Cat in non-human animals

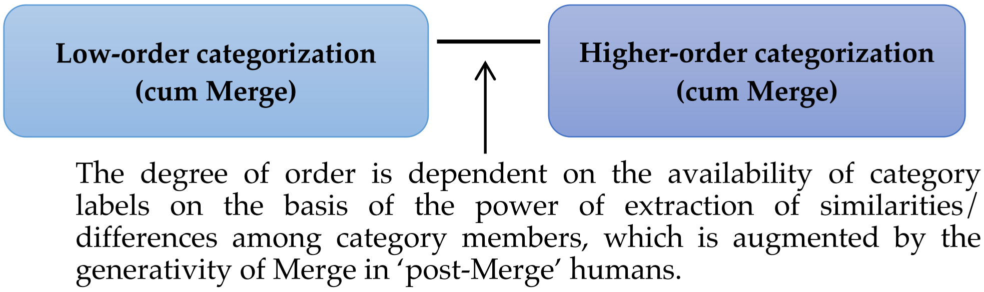

Figure 5: Low-order Cat & higher-order Cat (cum Merge) in humans

Low-order Cat is shared among at least all the vertebrates and higher-order categorization is possibly possessed by some of them as well. If my hypothesis of the categorization origin of Merge and the prevalent assumption that Merge is a label-free recursive set-formation operation are on the right track, it follows that there is a crucial difference between non-human animals and humans with respect to the properties of low-order categorization and higher-order categorization, given that SOs generated by Merge can be put to use as labels for Cat only in our species (see Hoshi 2018).[11]

2.3. *The Relation between Labeling in Categorization and Labeling in Syntax*

Based on neurocognitive considerations, Goucha et al. (2017) claim that labeling and not Merge is the evolutionary novelty that gave rise to the faculty of language (FL) in its present form. To the extent that Merge as a label-free recursive set-formation operation was exapted/co-opted from the recursive set-formation sub-component of Cat, as I argued in the previous section, it is indeed the case that Merge in this sense is not the evolutionary novelty in line with Goucha et al. (2017). On the other hand, as long as labeling is also involved in Cat (Hoshi 2018), the real evolutionary novelties are the nature of labeling in syntax

[11] It has been reported in the literature that non-human primates such as Old World monkeys are capable of combining two items but never more than two for their calls in communication (see Miyagawa & Clarke 2019 and references cited therein). To account for this fact, I will follow Miyagawa & Clarke (2019) in thinking that such two-item combination is not carried out by Merge but by the "dual-compartment frame into which each of the calls can fit" (p. 1).

in comparison with that of labeling in Cat and the constraint of binarity imposed on Merge.[12] Ideally, at least part of the nature of labeling in syntax and the binarity condition on Merge input should be somehow related to the nature of labeling in Cat, if the hypothesis of categorization origin of Merge put forth in Hoshi (2018) and refined in this opinion piece is on the right track.[13]

Goucha et al. (2017) submit that "the labeling of the outcome of the operation *Merge* is a necessary cognitive prerequisite for a complete account of the uniqueness of human language processing. It is through labeling that asymmetrical hierarchical structures can originate, thus distinguishing language from other communication systems" (p. 14).

However, labeling of a set *per se* is not a novelty in Merge at all, given that categorization (Cat) also inherently implicates labeling. Furthermore, strictly speaking, asymmetric hierarchical structures are not derived from labeling in and of itself, but originates from relative inclusion/containment relations holding among generated sets and elements within such sets regardless of labeling of those sets, interpreting such relative inclusion/containment relations in question as equivalent to asymmetric hierarchical structures.[14] This is indeed the case, given that the current formulation of Merge is independent of labeling and syntactic hierarchical structuring is implemented via unbounded application of Merge under the bare phrase structure system in the current minimalist program (Chomsky 2013, 2015).

Labeling of SOs is required to identify the nature of SOs, not to determine the hierarchical structuring of SOs, at the CI and SM interfaces. Rather, the hierarchical nature of syntactic structures of human language arises due to the specific mode of application of Merge *per se*. Note that Merge is a binary recursive unordered set-formation operation, which allows for taking in not only conceptual atoms such as lexical items but also its own output elements as its input elements. Crucially, this in turn requires that any unordered set generated by Merge must be in the state of being closed off once each application of Merge has been completed. Suppose X and Y are merged to yield an unordered set, hence Merge (X, Y) = {X, Y}. If this set is closed off, then merging of Z with this set would produce the following 'structured' unordered set {Z, {X, Y}}:

(6) Merge (Z, {X, Y}) = {Z, {X, Y}} (← a hierarchical structure)

If {X, Y} were not closed off, then merging of Z with this set would give the following 'unstructured' unordered set {Z, X, Y}:

---

12 Strictly speaking, there is at least one more evolutionary novelty: The target elements of Merge are either lexical items or SOs generated by Merge. I will not address the important issue of evolutionary origins of lexical items in human language in this opinion piece.

13 See also Boeckx (2009), Hornstein (2009), and Murphy (2015), *inter alia*, for other decompositional approaches to the evolution of Merge, which are different from mine.

14 Note that exactly the same asymmetrical hierarchical structures obtain among categories 'generated' by Cat as well (see Hoshi 2018 for some discussion). Note also that those inclusion/containment relations among sets and elements within such sets are created by the recursive nature of the relevant set-formation operation, which is shared by Merge and Cat, *pace* Hauser et al. (2002), if the hypothesis of categorization origin of Merge is correct.

(7) Merge (Z, {X, Y}) = {Z, X, Y} (← a flat structure)

As such, the hierarchical structuring of SOs as exemplified by (6) can arise thanks to the inclusion/containment relations holding among 'layered sets and elements within such sets' generated by unbounded recursive application of Merge with its 'close-off' property. In the pre-minimalist generative framework, dominance relations are defined against two-dimensional tree-diagrammatic phrase structures or X-bar structures, which virtually determine hierarchical structural relations among elements in them. On the other hand, in the bare phrase structure approach, such dominance relations cannot be directly defined but must be indirectly defined via set inclusion/containment relations (see, e.g., Chomsky 2000).

While labeling of a set is not unique to syntax proper as discussed so far, the nature of labeling seems to differ between syntax and categorization. What seems to be unique to the labeling of SOs and different from the labeling of categories is that, in the latter, members in category sets are equal and symmetric in status, and so extraction of similarities with respect to the label is required to determine membership; whereas, in the former, members in SOs are not equal and asymmetric in status, and so the label is determined on the basis of asymmetric nature of members of SOs (e.g., categorizers vs. label-free root elements; see the discussion below).

Now, I will reformulate and elaborate on my hypothesis from Hoshi (2018) concerning the categorization origin of Merge and labeling, as depicted in *Figure 6*. (Note that the requirement that SOs generated by Merge will be subject to labeling stems from the fact that they will be employed for labeling of Cat (both low-order and higher-order) in our species.)

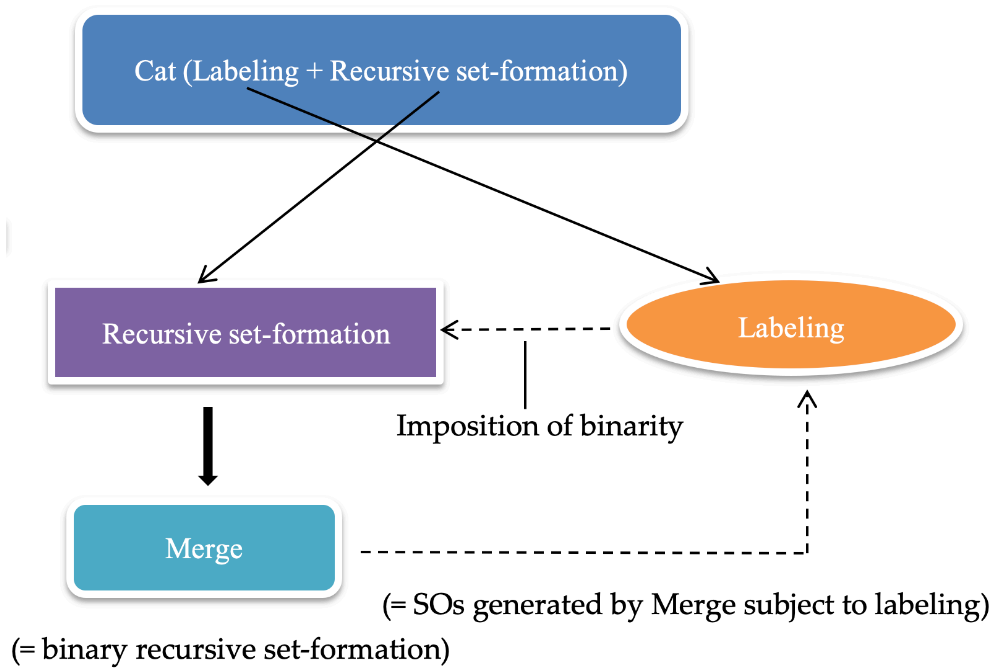

Figure 6: Evolutionary Relations among Cat, Merge, and Labeling

In the minimalist program, the binary nature of Merge has been one of the recalcitrant puzzles of human language. Chomsky (2008), for instance, tries to account for the binarity of Merge by appealing to the third factor principle of minimal computation, also suggesting other possibly conspiring conditions such as requirements of linearization at the SM interface along the lines of Kayne's (1994) Linear Correspondence Axiom, and conditions of predicate–argument structure at the CI interface.

However, as Fujita (2017) points out correctly, explanation of the nature of binarity of Merge by the third-factor principle of minimal computation does not seem to carry much force, because nothing in principle would prevent it from applying to other cognitive domains equally as well, so it could not provide the FL-specific property in Merge. Recall that, for example, the cardinality of the target of categorization is not restricted to two, as discussed in comparison with Merge in Hoshi (2018).

If my hypothesis of categorization origin of Merge and labeling illustrated in *Figure 6* is on the right track, it would provide an alternative account for the binarity of Merge on the basis of the nature of labeling, which I claim to be originally due to the sub-component of Cat. First of all, keep in mind that the labeling procedure for both Merge and Cat can be taken as a procedure for the purpose of identifying what kind of set in question is for Merge and Cat. Recall from section 2.2 that the identity of a category will be determined by the labeling procedure for characterizing a category set which is specified by the form κ(x)/κ(X), where κ is a characteristic function that will apply to any element x/X, in conjunction with the (non-)extraction of similarities related to κ among any member x/X of that category. Thus, there are two elements involved in the form κ(x)/κ(X), viz. one is a function and the other is its argument.

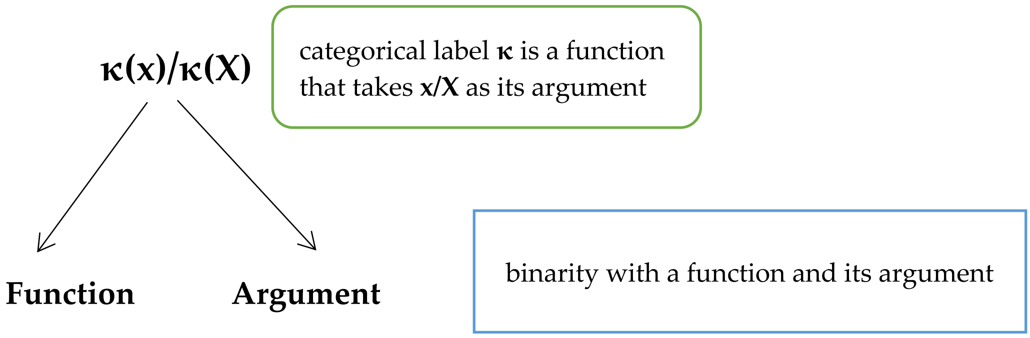

Figure 7: Binarity of Merge as Imposition of the Function–Argument Pattern of Categorial Labeling Schema

Notice that a function can be regarded as the label of any syntactic object SO under this conception. Even {XP, YP} structures such as the subject–predicate configuration {*n*P, *v*P}, either XP or YP can be taken as a function, hence the label of the whole set {XP, YP}. Thus, in {*n*P, *v*P}, for instance, *v*P is the function and *n*P is its argument, so *v*P will serve as the label for the whole set {*n*P, *v*P}.[15]

[15] Notice that this view of labeling is incompatible with Chomsky's (2013, 2015) thesis that the subject-predicate {XP, YP} structure and the intermediate *wh*-movement configuration {XP,

Suppose that, in the process of biological evolution of Merge illustrated in *Figure 6*, the 'function–argument binary' nature of the labeling procedure of Cat was imposed upon the recursive set-formation operation 'exapted/co-opted' from the recursive set-formation sub-component of Cat. Presumably, this would yield the binary requirement of Merge. At the same time, this claim has a significant consequence on the very nature of label in syntax. Notice that the label *per se* for Cat corresponds to a function rather than its argument. Thus, it is expected that the label, say, X in the SO {X, Y} generated by Merge, should be a function rather than its argument for determining the identity of the SO by the labeling procedure.

The typical case of X and Y as a categorizer (*n, v, a, p*) and a root element √*R*, respectively, can be regarded as a categorizer being a function that takes a root element as its argument to yield a particular label of the whole set {X, Y}, such as 'nominal', 'verbal', 'adjectival', 'prepositional/postpositional'. By the same token, when the set generated by Merge is {T, *v*P} or {C, TP}, the traditional functional category head T and C would be taken as a function that receives an eventuality expressed by *v*P as its argument, yielding a tensed eventuality (i.e. a situation), and a function that accepts a tensed eventuality/situation expressed by TP as its argument, giving out a tensed eventuality/situation with a force (i.e. a proposition), respectively, along the lines of Ramchand & Svenonius (2014).

At this point, the reader might wonder about the status of the labeling such as <ϕ, ϕ> under ϕ-feature agreement and <Q, Q> under Q-feature agreement implicating internal Merge of the subject and a *wh*-element in languages like English (see Chomsky 2013, 2015). Notice that this labeling pattern is quite different from the labeling pattern observed in the categorizer–root element structure on the basis of the categorizer as the label due to its status of a function.

In fact, if the line of analysis developed in this section is correct, neither <ϕ, ϕ> nor <Q, Q> can be a proper label for any SO {X, Y} in narrow syntax (NS) in the first place. In my view, labeling like <ϕ, ϕ> and <Q, Q> should be excluded from NS and the mapping from NS to the CI interface. The claim that agreement phenomena, including both ϕ-feature agreement and Q-feature agreement, are morpho-phonological in nature and do not belong to NS does not sound far-fetched, if we extend Marantz's (1991) and McFadden's (2004) idea that case/Case is a purely morphological phenomenon to both ϕ-feature agreement and Q-feature agreement.

One possibility is that agreement is a kind of 'extraction of similarities' which was 'exapted/co-opted' from Cat as a potential device in natural language and came to be employed in a subset of the whole set of natural language in the mapping from NS to the SM interface (note that East Asian languages like Japanese, Korean, and Chinese do not display ϕ-feature agreement, as is widely known in the literature; see, e.g., Fukui 1986 *et seq*. for Japanese). Thus, under the

YP} structure will necessitate obligatory extraction of either XP or YP to guarantee the labeling of the {XP, YP} structure. Hence, the reason for such extraction must be sought elsewhere. Note, also, that this view of labeling is not necessarily incompatible with Moro's (2000) analysis of small clauses, given that the small clause {XP, YP} is still ambiguous under my view in that both XP and YP could be either a predicate or an argument here as it stands, which may necessitate extraction of either XP or YP.

present line of analysis, <ϕ,ϕ> and <Q,Q> at {*n*P/DP, TP} and {*wh*-*n*P/DP, CP} in the traditional term seem to represent results of extraction of certain similarities (i.e. having ϕ-feature and Q-feature) by the minimal search of ϕ-bearing and Q-bearing elements in the SO set for the purpose of interpretation at the SM interface system (phonology/morphology) in FL. Consequences of this stand on labeling are far-reaching and I would like to leave investigation of them on the basis of concrete analysis of relevant empirical data in natural language to another occasion.

3. Parallelism of Laterality between Syntactic Processing and Categorization cum Extraction of Similarities

Given the hypothesis of categorization origin of Merge and the well-established fact that syntactic processing is implemented by the dorsal pathway connecting the pars opercularis (BA 44) in the posterior inferior frontal lobe and the posterior superior temporal gyrus/middle temporal gyrus in the temporal lobe in the left hemisphere (see Friederici 2017, Goucha et al. 2017, *inter alia*), it is expected that categorization (Cat) should also be implemented neutrally in the left hemisphere as well.

The interaction between the frontal cortex and the temporal cortex (both lateral and medial) is at work in categorization processes (Goucha et al. 2017; see Freedman et al. 2003 for such interaction between the frontal and temporal cortex for categorization in primates). Rogers et al. (2013: 131) point out that there is a close relation between the left hemisphere and categorization, citing Kosslyn et al. (1992), who demonstrate that "quantitative judgments by humans of the distance between objects are faster in the left visual field (right hemisphere), whereas at least some categorical judgments (e.g., above/below or connected/unconnected) are faster in the right visual field (left hemisphere)."

Vallortigara et al. (1999: 168) note that the fact that the left hemisphere is specialized for language and speech "could, in a sense, be regarded as a manifestation of a more general 'categorizing' ability of this hemisphere" and that this left hemisphere asymmetrical bias for the function of categorization holds in both birds and mammals (see also Denenberg 1981). On the basis of this observation, they suggest that this lateralization "emerged early in vertebrate evolution" from a phylogenetic point of view.

Reviewing the literature on the biology and behavior of brain asymmetries in humans and non-human animals, Rogers et al. (2013) note that the left hemisphere control is required for assessment of a stimulus for its categorization by focusing attention to local features of the environment and that "the left hemisphere is specialized to attend to similarities or invariances between stimuli, in order to allocate stimuli into categories following rules established through experience or biological predispositions" (p. 27).

In light of the parallelism of left hemisphere lateralization on the function of language, particularly syntax, and the function of categorization, the same left-side lateralizational bias for the two cognitive functions does not seem to be a mere accident. In fact, citing Vallortigara et al. (1999) and Rogers et al. (2013), Corballis (2002) suggest that the parallelism between the left hemisphere asymmetrical bias for the function of categorization observed among vertebrates and the one for the function of language in our species indicates that categorization shared among vertebrates might be a precursor to language. However, since, strictly speaking, language as a whole is not lateralized to the left hemisphere (see, e.g., Hickok & Poeppel 2007, Friederici 2017), it is more appropriate to state that categorization shared among vertebrates might be a precursor to Merge, as claimed by the hypothesis of categorization origin of Merge.

Using transcranial direct current stimulation on human subjects, Lupyan et al. (2012) demonstrate that the selection of properties relevant for categorization is related to the regulatory function of the inferior frontal gyrus including Broca's area in the left hemisphere. This can be interpreted as implying that the inferior frontal gyrus is involved in labeling of categorization.

If syntactic processing consists of both Merge and labeling and binarity of Merge is related to the categorial labeling schema as claimed above, it might suggest that the neural connection between BA44 and the posterior superior/middle temporal gyri is responsible for labeling and binarity of Merge and Merge as a pure recursive set-formation operation is implemented by a different brain region or network. Without further discussion in this opinion piece, on the basis of the 'basal ganglia grammar' model defended by Lieberman (2000, 2002, 2006), Balari & Lorenzo (2013), and Balari et al. (2013), I tentatively regard the basal ganglia as being mainly responsible for implementing the label-free recursive set-formation operation in Merge as part of its multi-purpose operation in conjunct-tion with different neural connections for different functions in the brain. In fact, Teichmann et al. (2015) discovered a crucial cortical-subcortical 'syntax' pathway linking Broca's area and the striatum in the basal ganglia. More specifically, they identified the BA45-left caudate head pathway in combination with the dorsal arcuate-BA44 pathway as responsible for phrasal level syntactic processing (see also references cited therein).

While a comprehensive investigation into this issue goes well beyond the scope of this piece, I would like to make a remark on this issue in light of the hypothesis of categorization origin of Merge that I proposed here. Although the emergence of Merge in the faculty of language seems to be a relatively recent adaptation of our species, if my hypothesis of categorization origin of Merge is on the right track, it does not necessarily indicate that Merge must be implemented solely by some neural network located in the neocortex, as suggested here.[16]

4. Concluding Remarks

In this opinion piece, I have addressed the three issues that I touched on but did not sufficiently discuss in Hoshi (2018). Those are (i) the relation between two modes of operation of categorization, viz., the interrelational mode of categorization (*IntCat*) and the differentiational mode of categorization (*DifCat*); (ii) the

---

[16] See Lenneberg (1967) for the view that the function of language should be regarded as the result of not only the horizontal interaction within the neocortex but also the vertical interaction between the neocortex and the subcortical structures. See also Theofanopoulou & Boeckx (2016) for recent discussion on a similar view.

relation between two levels of categorization, viz. low-order and higher-order categorization in humans and non-human animals; and (iii) the evolutionary relation between labeling in categorization and labeling in syntax.

Concerning (i), I claimed that the two apparently distinct modes of operation of categorization (IntCat and DifCat) stem from differences of 'extraction patterns' and thus as a precursor of Merge the particular mode of categorization such as IntCat should not be stipulated. As for (ii), I clarified the relation between low-order categorization, involving a series of entities, and higher-order categorization, involving a series of category sets, in humans and non-human animals, in connection with the qualitative difference of the two types of categorization between them in the context of evolution of human language. Finally, with regard to (iii), I made a specific claim about the relation between labeling in categorization and labeling in syntax, proposing a possible novel explanation of the binary nature of Merge with respect to its input cardinality on the basis of the very nature of the origin of labeling for SOs from labeling for categorization.

References

Balari, Sergio & Guillermo Lorenzo. 2013. *Computational Phenotypes: Towards an Evolutionary Developmental Biolinguistics.* Oxford: Oxford University Press.

Balari, Sergio, Antonio Benítez-Burraco, Víctor M. Longa, & Guillermo Lorenzo. 2013. The fossils of language: What are they, who has them, how did they evolve? In Cedric Boeckx & Kleanthes K. Grohmann (eds.), *The Cambridge Handbook of Biolinguistics*, 489–523. Cambridge: Cambridge University Press.

Bickerton, Derek. 1990. *Language & Species*. Chicago: University of Chicago Press.

Blumstein, Daniel T. 2007. Feeling the heat: Ground squirrels heat their tails to discourage rattlesnakes attack. *PNAS* 104, 14177–14178.

Boeckx, Cedric. 2009. The nature of Merge: Consequences for language, mind, and biology. In Massimo Piattelli-Palmarini, Juan Uriagereka, & Pello Salaburu (eds.), *Of Minds & Language: A Dialogue with Noam Chomsky in the Basque Country*, 44–57. Oxford: Oxford University Press.

Bouchard, Denis. 2013. *The Nature and Origin of Language*. Oxford: Oxford University Press.

Chomsky, Noam. 1995. *The Minimalist Program.* Cambridge, MA: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels, & Juan Uriagereka (eds.), *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, 89–155. Cambridge, MA: MIT Press.

Chomsky, Noam. 2008. On phases. In Robert Freidin, Carlos P. Otero, & Maria Luisa Zubizarreta (eds.), *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud*, 133–166. Cambridge, MA: MIT Press.

Chomsky, Noam. 2013. Problems of projection. *Lingua* 130, 33–49.

Chomsky, Noam. 2015. Problems of projection: Extensions. In Elisa Di Domenico, Cornelia Hamann, & Simona Matteini (eds.), *Structures, Strategies and Beyond: Studies in Honour of Adriana Belletti*, 3–16. Amsterdam: John Benjamins.

Chomsky, Noam. 2017. The language capacity: architecture and evolution. *Psycho-nomic Bulletin & Review* 24, 200–203.
Conway, Christopher M. & Morten H. Christiansen. 2001. Sequential learning in non-human primates. *Trends in Cognitive Sciences* 5, 539–546.
Corballis, Michael C. 2002. *From Hand to Mouth: The Origins of Language*. Princeton, NJ: Princeton University Press.
Denenberg, Victor H. 1981. Hemispheric laterality in animals and the effects of early experience. *Behavioral and Brain Sciences* 4, 1–49.
Embick, David & Rolf Noyer. 2007. Distributed morphology and the syntax–morphology interface. In Gillian Ramchand & Charles Reiss (eds.), *The Oxford Handbook of Linguistic Interfaces*, 289–324. Oxford: Oxford University Press.
Embick, David & Alec Marantz. 2008. Architecture and blocking. *Linguistic Inquiry* 39, 1–53.
Freedman, David J., Maximilian Riesenhuber, Tomaso Poggio, & Earl K. Miller. 2003. A comparison of primate prefrontal and inferior temporal cortices during visual categorization. *Journal of Neuroscience* 23, 5235–5246.
Friederici, Angela D. 2017. *Language in Our Brain: The Origins of a Uniquely Human Capacity*. Cambridge, MA: MIT Press.
Fujita, Koji. 2017. On the parallel evolution of syntax and lexicon: A Merge-only view. *Journal of Neurolinguistics* 43B, 178–192.
Fukui, Naoki. 1986. A Theory of Category Projection and Its Applications. Cambridge, MA: MIT doctoral dissertation.
Goucha, Tomas, Emiliano Zaccarella, & Angela D. Friederici. 2017. A revival of the *Homo loquens* as a builder of labeled structures: Neurocognitive considerations. *Neuroscience and Biobehavioral Reviews* 81B, 213–224.
Gould, Stephen Jay & Elisabeth S. Vrba. 1982. Exaptation—A missing term in the science of form. *Paleobiology* 8, 4–15.
Hauser, Marc D., Noam Chomsky, & W. Tecumseh Fitch. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science* 298, 1569–1579.
Hickok, Gregory & David Poeppel. 2007. The cortical organization of speech processing. *Nature Reviews Neuroscience* 8, 393–402.
Hornstein, Norbert. 2009. *A Theory of Syntax: Minimal Operations and Universal Grammar*. Cambridge: Cambridge University Press.
Hoshi, Koji. 2018. Merge and labeling as descent with modification of categorization: A neo-Lennebergian approach. *Biolinguistics* 12, 39–54.
Hurford, James R. 2007. *The Origins of Meaning: Language in the Light of Evolution*. Oxford: Oxford University Press.
Kayne, Richard S. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.
Kosslyn, Stephen M., Christopher F. Chabris, Chad J Marsolek, & Olivier Koenig. 1992. Categorical versus coordinate spatial relations: Computational analyses and computer simulations. *Journal of Experimental Psychology: Human Perception and Performance* 18, 562–577.
Lakoff, George. 1987. *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. Chicago, IL: University of Chicago Press.

Langacker, Ronald. W. 1991. *Foundations of Cognitive Grammar,* vol. II: *Descriptive Application.* Stanford, CA: Stanford University Press.
Langacker, Ronald. W. 1997. Constituency, dependency, and conceptual grouping. *Cognitive Linguistics* 8.1, 1–32.
Lenneberg, Eric H. 1967. *Biological Foundations of Language.* New York: John Wiley & Sons.
Lieberman, Philip. 2000. *Human Language and Our Reptilian Brain: The Subcortical Bases of Speech, Syntax, and Thought.* Cambridge, MA: Harvard University Press.
Lieberman, Philip. 2002. On the nature and evolution of the neural bases of human language. *Yearbook of Physical Anthropology* 45, 36–62.
Lieberman, Philip. 2006. *Toward an Evolutionary Biology of Language.* Cambridge, MA: Harvard University Press.
Lupyan, Gary, Daniel Mirman, Roy Hamilton, & Sharon L. Thompson-Schill. 2012. Categorization is modulated by transcranial direct current stimulation over left prefrontal cortex. *Cognition* 124, 36–49.
Lyn, Heidi, Patricia Greenfield, Sue Savage-Rumbaugh, Kristen Gillespie-Lynch, & William D Hopkins. 2011. Nonhuman primates do declare! A comparison of declarative symbol and gesture use in two children, two bonobos, and a chimpanzee. *Language and Communication* 31, 63–74.
McFadden, Thomas. 2004. The Position of Morphological Case in the Derivation: A Study on the Syntax–Morphology Interface. Philadelphia, PA: University of Pennsylvania doctoral dissertation.
Marantz, Alec. 1991. Case and licensing. In German Westphal, Benjamin Ao, & Hee-Rahk Chae (eds.), *Proceedings of ESCOL '91*, 234–253. Ithaca, NY: Cornell Linguistics Club.
Marantz, Alec. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. *University of Pennsylvania Working Papers in Linguistics* 4, 201–225.
Mareschal, Denis, Paul C. Quinn, & Stephen E. G. Lea (eds.). 2010. *The Making of Human Concepts*. Oxford: Oxford University Press.
Miyagawa, Shigeru & Esther Clarke. 2019. Systems underlying human and old world monkey communication: One, two, or infinite. *Frontiers in Psychology* 10: 1911.
Moro, Andrea. 2000. *Dynamic Antisymmetry*. Cambridge, MA: MIT Press.
Murphy, Elliot. 2015. Labels, cognomes, and cyclic computation: An ethological perspective. *Frontiers in Psychology* 6: 715.
Murphy, Gregory L. 2010. What are categories and concepts? In Mareschal et al., 11–28.
Nóbrega, Vitor A. & Shigeru Miyagawa. 2015. The precedence of syntax in the rapid emergence of human language in evolution as defined by the integration hypothesis. *Frontiers in Psychology* 6: 271.
Ohno, Susumu. 1970. *Evolution by Gene Duplication*. New York: Springer.
Penn, Derek. C., Keith J. Holyoak, & Daniel J. Povinelli. 2008. Darwin's mistake: Explaining the discontinuity between human and nonhuman minds. *Behavioral and Brain Sciences* 31, 109–178.

Ramchand, Gillian & Peter Svenonius. 2014. Deriving the functional hierarchy. *Language Sciences* 46, 152–174.
Roberts, William A., & Mazmanian, Dwight S. 1988. Concept learning at different levels of abstraction by pigeons, monkeys, and people. *Journal of Experimental Psychology*: *Animal Behavior Processes* 14, 247–260.
Rogers, Lesley J., Giorgio Vallortigara, & Richard J. Andrew. 2013. *Divided Brains: The Biology and Behavior of Brain Asymmetries*. Cambridge: Cambridge University Press.
Rosch, Eleanor H. 1973. Natural categories. *Cognitive Psychology* 4, 328–350.
Rundus, Aaron S., Donald H. Owings, Sanjay S. Joshi, Erin Chinn, & Nicolas Giannini. 2007. Ground squirrels use an infrared signal to deter rattlesnake predation. *PNAS* 104, 14372–14376.
Savage-Rumbaugh, Sue, Stuart G. Shanker, & Talbot J. Taylor. 1998. *Apes, Language, and the Human Mind*. New York: Oxford University Press.
Seyfarth, Robert M., Dorothy L. Cheney, & Peter R. Marler. 1980. Monkey responses to three different alarm calls: Evidence of predator classification and semantic communication. *Science* 210, 801–803.
Sloutsky, Vladimir M. 2003. The role of similarity in the development of categorization. *Trends in Cognitive Sciences* 7, 246–251.
Spinozzi, Giovanna, Francesco Natale, Jonas Langer, & Karen E. Brakke. 1999. Spontaneous class grouping behavior by bonobos (Pan paniscus) and common chimpanzees (P. troglodytes). *Animal Cognition* 2, 157–170.
Tattersall, Ian. 2017. How can we detect when language emerged? *Psychonomic Bulletin & Review* 24, 64–67.
Taylor, John R. 2003. *Linguistic Categorization*. Oxford: Oxford University Press.
Teichmann, Marc, Charlotte Rosso, Jean-Baptiste Martini, Isabelle Bloch, Pierre Brugières, Hugues Duffau, Stéphane Lehéricy, & Anne-Catherine Bachoud-Lévi. 2015. A cortical–subcortical syntax pathway linking Broca's area and the striatum. *Human Brain Mapping* 36, 2270–2283.
Terrance, Herbert S., Laura-Ann Petitto, Richard J. Sanders, & Thomas G. Bever. 1979. Can an ape create a sentence? *Science* 206, 891–902.
Theofanopoulou, Constantina & Cedric Boeckx. 2016. The central role of the thala-mus in language and cognition. In Koji Fujita & Cedric Boeckx (eds.), *Advances in Biolinguistics: The Human Language Faculty and Its Biological Basis*, 230–255. London: Routledge.
Tomasello, Michael. 1999. *The Cultural Origins of Human Cognition*. Cambridge, MA: Harvard University Press.
Vauclair, Jacques. 2002. Categorization and conceptual behavior in nonhuman primates. In Marc Bekoff, Colin Allen, & Gordon M. Burghardt (eds.), *The Cognitive Animal: Empirical and Theoretical Perspectives on Animal Cognition*, 239–245. Cambridge, MA: MIT Press.
Vallortigara, Giorgio, Lesley J. Rogers, & Angelo Bisazza. 1999. Possible evolutionary origins of cognitive brain function. *Brain Research Reviews* 30, 164–175.
Vonk, Jennifer & Suzanne E. McDonald. 2002. Natural concepts in a juvenile gorilla (*gorilla gorilla gorilla*) at three levels of abstraction. *Journal of the Experimental Analysis of Behavior* 78, 315–332.

Vonk, Jennifer & Suzanne E. McDonald. 2004. Levels of abstraction in orangutan (*Pongo abelii*) categorization. *Journal of Comparative Psychology* 118, 3–13
Yang, Charles. 2013. Ontogeny and phylogeny of language. *PNAS* 110, 6324–6327.
Zhang, Jianzhi. 2003. Evolution by gene duplication. *Trends in Ecology and Evolution* 18, 292–298.

Koji Hoshi
Keio University
Faculty of Economics
4-1-1 Hiyoshi Kohoku-ku
223-8521 Yokohama, Kanagawa
Japan
khoshi@a7.keio.jp