

## The integration hypothesis of human language evolution and the nature of contemporary languages

 Shigeru Miyagawa<sup>1,2\*</sup>,  Shiro Ojima<sup>3</sup>,  Robert C. Berwick<sup>4</sup> and  Kazuo Okanoya<sup>3,5</sup>

- <sup>1</sup>Department of Linguistics and Philosophy, Massachusetts Institute of Technology, Cambridge, MA, USA
- <sup>2</sup>Center for Research and Development of Higher Education, University of Tokyo, Tokyo, Japan
- <sup>3</sup>Department of Life Sciences, The University of Tokyo, Tokyo, Japan
- <sup>4</sup>Department of Electrical Engineering and Computer Science and Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA, USA
- <sup>5</sup>Okanoya Emotional Information Project, Exploratory Research for Advanced Technology, Japan Science and Technology Agency, Tokyo, Japan

How human language arose is a mystery in the evolution of *Homo sapiens*. Miyagawa et al. (2013) put forward a proposal, which we will call the **Integration Hypothesis** of human language evolution, that holds that human language is composed of two components, E for *expressive*, and L for *lexical*. Each component has an antecedent in nature: E as found, for example, in birdsong, and L in, for example, the alarm calls of monkeys. E and L integrated uniquely in humans to give rise to language. A challenge to the Integration Hypothesis is that while these non-human systems are finite-state in nature, human language is known to require characterization by a non-finite state grammar. Our claim is that E and L, taken separately, are in fact finite-state; when a grammatical process crosses the boundary between E and L, it gives rise to the non-finite state character of human language. We provide empirical evidence for the Integration Hypothesis by showing that certain processes found in contemporary languages that have been characterized as non-finite state in nature can in fact be shown to be finite-state. We also speculate on how human language actually arose in evolution through the lens of the Integration Hypothesis.

### Introduction

Human language appears to have developed within the past 100,000 years (Tattersall, 2009). While it is extremely challenging to confirm any hypothesis of the actual process that led to the emergence of language, it is possible to formulate a theory that is broadly compatible with what we find in contemporary systems among mammals, birds, and humans. Miyagawa et al. (2013) put forward such a theory, which we will call the **Integration Hypothesis** of human language evolution. In this article, we will provide empirical evidence from contemporary languages for crucial components of the Integration Hypothesis. We will also speculate on how human language actually arose in evolution through the lens of the Integration Hypothesis.

We will focus on the structures found in human language and compare them to other systems such as those found in monkey alarm calls and birdsong. In recent linguistic theory, it is proposed that there is just one rule for structure building, called Merge, which takes two items and combines them into an unordered set ([Chomsky, 1995](#)). If Merge is what gives human language its unique character for building structures, it is this operation that largely distinguishes human language from other systems ([Hauser et al., 2002](#); [Berwick, 2011](#)). This view of human language leaves open a host of questions including: (i) how did Merge appear?; (ii) why is human language characterizable by a non-finite state grammar ([Chomsky, 1956](#)) while other systems of the animal world are finite-state in nature ([Berwick et al., 2011](#))?; and (iii) why do we find processes such as movement and agreement in human language ([Chomsky, 1995](#); [Miyagawa, 2010](#))? The Integration Hypothesis addresses these questions by advancing a conventional Darwinian view: two pre-adapted systems found elsewhere in the animal world were *integrated* in humans to give rise to the unique system that underlies today's languages. One system, called Type E for *expressive*, is found, for example, in birdsong ([Berwick et al., 2011](#)), which serves to mark mating availability and other “expressive” functions. The second system, Type L for *lexical*, is found in monkey calls ([Seyfarth et al., 1980](#); [Arnold and Zuberbühler, 2006](#)) and honeybee waggle dances ([Riley et al., 2005](#)). Types E and L are the two primary forms of communication found in the animal world. Our view that human language syntax arose from pre-existing systems as found in other species is a conventional mode of evolutionary explanation, and so has been advanced by other researchers. For example, [Fitch \(2011\)](#) suggests that the roots of the core computational capacity of human language may be found in motor control and motor planning, while others such as [Hurford \(2011\)](#) allude to a gradual development from non-human primate call systems. We take no stand on these particular hypotheses regarding language's origin—directly analogizing language motor activity is not at all straightforward, as the recent exchange between [Moro \(2014a,b\)](#) and [Pulvermüller \(2014\)](#) demonstrates. Rather, we approach a different aspect of the origin of language: how a non-context free system emerged by conjoining two antecedent systems that were only finite-state. The Integration Hypothesis is advanced to explore some possibilities; it differs from other accounts like those above in that it is more linguistically detailed and broadly consistent with facts of contemporary languages. At the end, we will speculate on how the E and L systems emerged in humans.

### **The Integration Hypothesis of Human Language Evolution ([Miyagawa et al., 2013](#))**

Every human language sentence is composed of two layers of meaning: a *lexical* structure that contains the lexical meaning ([Hale and Keyser, 1993](#)), and an *expression* structure that is composed

of function elements that give shape to the expression (Chomsky, 1995; Miyagawa, 2010). In the question, *Did John eat pizza?*, the lexical layer is composed of the words *John, eat, pizza*; these words are constant across a variety of expressions. The sentence also contains *did*, which has two functions: it marks tense, and by occurring at the head of the sentence, it also signifies a question. Tense and question are two elements that give form to the expression, making it possible to use it in conversation. The two layers of meaning are commonly represented as follows.

(1) Duality of semantics (Chomsky, 1995, 2008; Miyagawa, 2010)

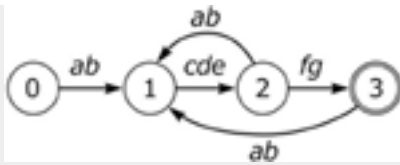


The Integration Hypothesis (Miyagawa et al., 2013) views these two layers as having antecedents in other animal species. The lexical layer is related to those systems that employ isolated uttered units that correlate with real-world references, such as the alarm calls of Vervet monkeys for pythons, eagles, and leopards (Seyfarth et al., 1980). The expression layer is similar to birdsongs; birdsongs have specific patterns, but they do not contain words, so that birdsongs have syntax without meaning (Berwick et al., 2012), thus it is of the E type. Although parallels between birdsong and human language have often been suggested (Darwin, 1871; Jespersen, 1922; Marler, 1970; Nottebohm, 1975; Doupe and Kuhl, 1999; Okanoya, 2002; Bolhuis et al., 2010; Berwick et al., 2012), we believe that the actual link is between birdsong and the expression structure portion of human language.

(2) Human language and the non-human language-like types lexical structure  $\longleftrightarrow$  bee dances/primate calls Type L expression structure  $\longleftrightarrow$  birdsong Type E

Birdsongs can be complex, as in the example of the Bengalese finch. The Bengalese finch song loops back to various positions in the song, which leads to considerable variation (Figure 1). Nevertheless, all known birdsongs can be described as a *k-reversible* finite state automaton (Berwick et al., 2011), a restricted class of automata that are efficiently learnable from examples. The L type also is a simple finite state system. The Integration Hypothesis conjectures that these two major systems in nature that underlie communication, E and L, integrated uniquely in humans to give rise to language.

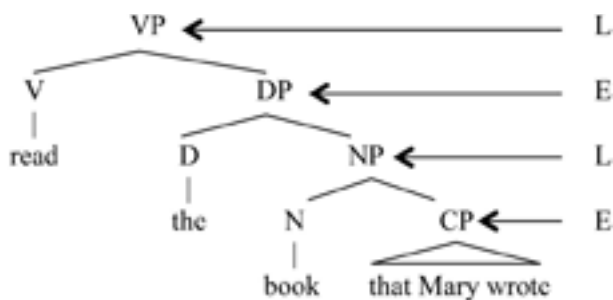
**FIGURE 1**



**Figure 1. Bengalese finch song.**

Some theories of human language are not easily compatible with the views proposed here. For example, Lexical-Functional Grammar (LFG) views words and phrases as having equivalent functions. However, there are the notions of *argument structure* and *expression structure* (Bresnan, 2001, pp. 9–10) that parallel in general terms the design we are assuming. We in fact adopt the term *expression structure* from LFG. Distributed Morphology (Halle and Marantz, 1993; Marantz, 1997; Embick, 2010) denies a division between word and phrasal formation. Nevertheless, DM contains a division reminiscent of the E/L layers. “Words” are listed as category-neutral roots indicated by  $\surd$ , e.g., [ $\surd$ CONSUME]. A category specification head such as D (noun) or *v* (verb) is added to furnish category specification: [D[space]consumption (of water)] [*v*[space]consume (water)]. The “root” layer is something akin to the L system in our proposal. Once a category-specifying item is merged, that structure becomes similar to our E layer—it participates in syntactic processes of merge and labeling, movement, etc. One difference is that in DM, category-less items may combine directly, something we do not believe is possible; L items do not directly combine with each other. This is why we typically find E-L alternations<sup>1</sup>.

(3) E/L hierarchical structure (“D” stands for “Determiner” and is part of the E system for noun phrases)



### Three Challenges for the Integration Hypothesis from Contemporary Languages

We take up three challenges to the Integration Hypothesis from contemporary linguistics: two that ostensibly argue against our proposal that inside E and L we only find finite-state processes; and a third having to do with the assumption that L items cannot combine directly—any combination requires intervention from E.

The first challenge to the Integration hypothesis that E and L are finite state regards the existence of so-called discontinuous word formation. For example, Carden (1983), based on Bar-Hillel and

Shamir (1960) and Langendoen (1975, 1981), argues that sequences involving the prefix *anti-* and a noun such as *missile* are non-finite state in nature (see also Boeckx, 2006; Narita et al., 2014).

(4) a. [anti-missile]

b. [anti-[anti-missile] missile] missile

The ostensible point is that this formation can involve center embedding, which would constitute a non-finite state construction. When additional *anti* is attached to the front of the construction, one or more instances of *missile* must occur at the end (4b), giving the impression of center embedding. However, this is not the correct analysis. When *anti-* combines with a noun such as *missile*, the sequence *anti-missile* is a modifier that would modify a noun with this property, thus, [*anti-missile*]-*missile*, [*anti-missile*]-*defense*. Each successive expansion forms via strict adjacency, as shown by the italicized element below, without the need to posit a center embedding, non-regular grammar.

(5) a. [anti-missile]-*missile*

b. *anti*-[[anti-missile]-*missile*] (modifier)

c. [anti-[[anti-missile]-*missile*]]-*missile* (or, anti-anti-missile-missile-*defense*)

The final construction also led some to claim that when *anti-* is added on the left, two instances of *missile* must occur on the right, which would be a non-regular grammar process. However, that is not the correct way to view this construction. *anti-* is attached to [[*anti-missile*]-*missile*], forming the modifier *anti*-[[*anti-missile*]-*missile*]. To this the additional *missile* is added that is modified by the rest, giving appearance that two instances of *missile* were added.

The second challenge to the finite state nature of E/L is reduplication, often cited as being non-finite state (McCarthy and Prince, 1995, 1999; Urbanczyk, 2007). In reduplication a word is reduplicated in its entirety or in part.

(6) Full reduplication:  $C_1V_1C_2V_2C_3 - C_1V_1C_2V_2C_3$

Partial reduplication:  $C_1V_1 - C_1V_1C_2V_2C_3$ .

Following are actual examples of full and partial reduplication (Moravcsik, 1978).

(7) a. kuuna-kuuna “husbands” (Tohono O’odham plural)

b. tak-takki “legs” (Agta plural)

Contrary to the non-finite state approaches common in the literature, Raimy (2000) provides an analysis of reduplication that, in its most basic form, is similar to the 1 finite state automaton we saw for the song of Bengalese finch. He argues that reduplication is a process of looping back:

(8) 1 Finite State Automaton and Reduplication:



There are cases in which a reduplicant may occur to the right of the base: *erasi-rasi* “he is sick” (Siriono continuative, Key, 1965). Here the reduplicant is a copy that begins in the middle of the base and goes to the end. Right-handed reduplicants always have this property of starting in the middle of the base and copy to the end (Marantz, 1982).

(9) “Suffix” Reduplication:



This copying process is a product of a loop back to the middle of the string.

The third challenge concerns the assumption that the members of L do not directly combine with each other. There are compound words such as *tea:cup*, *brain:power*, that appear to be L-L combinations. However, there is evidence that some E element does occur between the two L's. In German, when two words combine to form a compound, typically an element (/n/ or schwa) is inserted between the two words, as in *Blume-N-wiese* “flower meadow” (Aronoff and Fuhrhop, 2002); this “linking” element has no apparent function, so we can reasonably assume this sequence to be L-E-L. In English, we find a similar linking element in the form of /s/ in: *craftSman*, *markSman*, *spokeSman* (Marchand, 1969). This /s/ has no function other than to link the two L's. These linking elements suggest that there is a slot between the two L's in compound words where we predict an E element to occur. In the case of *teacup*, where there is no overt linker, we surmise that a phonologically null element occurs in that position. As a reviewer notes, languages such as Chinese, where sentences appear to be simple noun-verb-noun sequences, the idea that there are expression items intervening between L items becomes a challenge. Sybesma (2007) argues that there are tests to detect the occurrence of tense in Chinese, hence a T head, despite the fact that it is not pronounced.

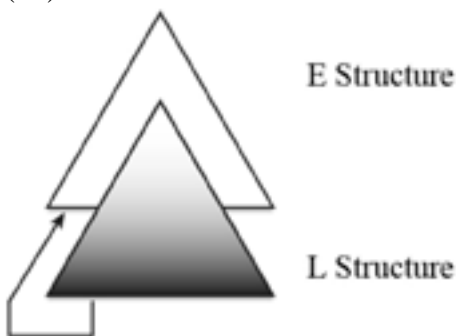
**Movement as a Non-Finite State Process**

An operation that is pervasive in human language is movement.

(10) What did you eat \_\_\_?

The question word *what* is the object of *eat*, yet it has evidently been displaced from this position of thematic interpretation after the verb to where it is actually pronounced, at the head of the sentence. This is clearly a non-finite state operation. When we look at a typical syntactic movement, it is from the L structure to the E structure: *what* begins in the L position of object, then moves to the E position of Question (e.g., Chomsky, 2001, 2008; Miyagawa, 2010).

(11) Movement



Agreement is another process that crosses E and L (Miyagawa et al., 2013). Movement and agreement are processes that, by connecting E and L, tie the two structures together. Hence, while we find finite state grammar processes inside E and L, thus reflecting their antecedents in the non-human animal world, non-finite state procedure is introduced to link the two structures. It is only in crossing from one structure to another that something other than a finite state operation is required.

Theories that do not posit movement nevertheless have operations that cross E and L. For example, Head-driven Phrase Structure Grammar (HPSG) constructs “pointers” between “what” at the head of sentences to the position after “eat,” via the propagation of information from “what” to this thematic argument point. Although there is no explicit “movement,” the effect is the same (Sag et al., 2003). Similarly, LFG reconstructs such pairings by means of information structure pairings that cross E-L boundaries, using a base context-free grammar that is composed from two finite-state systems in just the manner suggested above. To be sure, given the wide range of current syntactic theories, in other cases it is simply not possible to mimic the E-L account—an unsurprising outcome, since such theories are often incompatible with each other, as noted by Jackendoff (2010).

### **Speculation on the Integration of E and L**

Given the evolutionary proximity between humans and other primates, the lexical structure in human language can plausibly be traced to non-human primates and their alarm calls and similar L

systems. However, the same cannot be said of expression structure and birdsong. The ancestors of present-day birds and mammals split 300 million years ago (Benton, 1990), an evolutionary divide of 600 million years that suggests convergent evolution—independent evolution of E systems in birds and humans, rather than descent from a common ancestor that possessed this trait. Further, even within the *Aves* lineage, vocal learning in songbirds has been independently evolved; for example, there are closely related bird species, such as Ruby Throated hummingbird and Anna's hummingbird, where the former possesses vocal learning but the latter does not—a concrete example of convergent evolution. The other evolutionary possibility is that E systems were present in the common ancestors of humans and non-human primates, or even the rest of the mammalian lineage, in which case humans would have E in virtue of common descent, although the E system would not necessarily be expressed as part of a communication system.

Some behavioral patterns of non-human mammals can be described by finite-state grammars. Examples include the food-hoarding behavior of Syrian golden hamsters (Jones and Pinel, 1990) and the facial grooming actions of rats (Berridge et al., 1987). However, the finite-state nature of rodents' action sequences does not, in itself, make them Type-E systems. Individual action units in such sequences are relatively independent of each other, while song elements in birdsong are produced rapidly in succession, creating a sustained pattern when seen as a whole. In rodents, each action unit also has a functional meaning, while individual song elements of birds are meaningless.

The two requirements for an E system are:

(12) E System

- (i) It creates a sustained pattern;
- (ii) It holistically expresses an internal state of the singer.

E systems may be present to a limited extent in the singing behavior of non-human primates. Most non-human primates do not sing, but there is an exception: gibbons (Hylobatidae) (Marshall and Marshall, 1976; Haimoff, 1984). They sing long, complex songs. The gibbon song, as a whole, has functions such as territory advertisement, mate attraction, the strengthening of pair and family bonds (Brockelman and Srikosamatara, 1984; Raemaekers et al., 1984; Mitani, 1985; Geissmann and Orgeldinger, 2000). This is analogous to birdsong, a Type E system, which holistically expresses the singer's internal state.

In most gibbon species, male songs can be flexible in the order of notes (song elements) (Raemaekers et al., 1984; Haimoff, 1985; Mitani, 1988). For example, the male song of the Javan silvery gibbon (*Hylobates moloch*) contains 14 distinct note types, which can be assembled into a

song in various orders (Geissmann et al., 2005). The transition from one note type to another appears to be probabilistic (see Figure 7 of Geissmann et al., 2005). The gibbon song, characterized by probabilistic transitions among different note types but lacking internal syntactic hierarchy, may be analogous in its grammatical structure to certain birdsong.

Hence, non-human primates, our close relatives, may have the latent potential to vocalize continuously in a finite state fashion to convey a holistic message. What prevents most of them from doing so is not entirely clear. It may be difficult for them to coordinate various articulation apparatuses rhythmically, which is required in singing and speech-like vocalizations. Non-human primates' ability to produce rhythmic orofacial movements has only recently begun to be reported. The gelada, a non-human primate, can vocalize during the action of “lip-smacking” (rapid opening and closing of the mouth and lips), which shares rhythmic features with orofacial movements involved in human speech (Ghazanfar et al., 2012; Bergman, 2013). Further searches for E-like systems should be continued in both vocal and non-vocal domains. We also need to understand the neural mechanisms underlying Type-L and Type-E systems, in evolutionary contexts.

Rauschecker's work (e.g., Rauschecker, 2012) suggests that auditory regions of the brain are hierarchically organized in both humans and non-human primates, with more anterior portions of the ventral auditory stream responding to more complex auditory objects such as spoken words in humans and calls in monkeys. It might be tempting to link Type-L systems to the ventral auditory stream, but we must await future research before accepting such a view.

## Conflict of Interest Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Acknowledgments

We would like to thank the two reviewers and the associate editor for numerous helpful suggestions. We also thank Yoichi Inoue for comments on an earlier draft. Finally, we thank the assistance of Edward Flemming, Junko Ito, Armin Mester, Hiroki Nomoto, and Donca Steriade. This study was partially supported by MEXT Grants-in-Aid for the Scientific Research (No. 23240033 to Kazuo Okanoya and No. 23520757 to Shiro Ojima) and ERATO, Japan Science and Technology Agency, and by internal funding from MIT.

## Footnotes

1. <sup>^</sup>As a reviewer notes, a recent approach called nanosyntax (e.g., [Starke, 2009](#)) appears to be fundamentally in conflict with the Integration Hypothesis.

Nanosyntax posits that morphemes may consist of several terminal nodes, thus, syntactic in nature. We leave any attempt to compare this with our approach for future research.

## References

Arnold, K., and Zuberbühler, K. (2006). Language evolution: semantic combinations in primate calls. *Nature* 441, 303. doi: 10.1038/441303a

[Pubmed Abstract](#) | [Pubmed Full Text](#) | [CrossRef Full Text](#)

Aronoff, M., and Fuhrhop, N. (2002). Restricting suffix combinations in German and English: closing suffixes and the monosuffix constraint. *Nat. Lang. Linguist. Theory* 20, 451–490. doi: 10.1023/A:1015858920912

[CrossRef Full Text](#)

Bar-Hillel, Y., and Shamir, E. (1960). Finite-state languages: formal representations and adequacy problems. *Bull. Res. Council. Isr.* 8F, 155–166.

Benton, M. J. (1990). Phylogeny of the major tetrapod groups - morphological data and divergence dates. *J. Mol. Evol.* 30, 409–424. doi: 10.1007/BF02101113

[Pubmed Abstract](#) | [Pubmed Full Text](#) | [CrossRef Full Text](#)

Bergman, T. J. (2013). Speech-like vocalized lip-smacking in geladas. *Curr. Biol.* 23, R268–R269. doi: 10.1016/j.cub.2013.02.038

[Pubmed Abstract](#) | [Pubmed Full Text](#) | [CrossRef Full Text](#)

Berridge, K. C., Fentress, J. C., and Parr, H. (1987). Natural syntax rules control action sequence of rats. *Behav. Brain Res.* 23, 59–68. doi: 10.1016/0166-4328(87)90242-7

[Pubmed Abstract](#) | [Pubmed Full Text](#) | [CrossRef Full Text](#)

Berwick, R. C. (2011). “All you need is merge: biology, computation, and language from the bottom up,” in *The Biolinguistic Enterprise: New Perspectives on the Evolution and Nature of the Human Language Faculty*, eds A. M. Di Sciullo and C. Boeckx (Oxford: Oxford University Press), 461–491.

Berwick, R. C., Beckers, G. J., Okanoya, K., and Bolhuis, J. J. (2012). A bird's eye view of human language evolution. *Front. Evol. Neurosci.* 4:5. doi: 10.3389/fnevo.2012.00005

[Pubmed Abstract](#) | [Pubmed Full Text](#) | [CrossRef Full Text](#)

Berwick, R. C., Okanoya, K., Beckers, G. J. L., and Bolhuis, J. J. (2011). Songs to syntax: the linguistics of birdsong. *Trends Cogn. Sci.* 15, 113–121. doi: 10.1016/j.tics.2011.01.002

[Pubmed Abstract](#) | [Pubmed Full Text](#) | [CrossRef Full Text](#)

Boeckx, C. (2006). *Linguistic Minimalism: Origins, Concepts, Methods, and Aims*. New York, NY: Oxford University Press.

Bolhuis, J. J., Okanoya, K., and Scharff, C. (2010). Twitter evolution: converging mechanisms in birdsong and human speech. *Nat. Rev. Neurosci.* 11, 747–759. doi: 10.1038/nrn2931

[Pubmed Abstract](#) | [Pubmed Full Text](#) | [CrossRef Full Text](#)

Bresnan, J. (2001). *Lexical-Functional Syntax*. Oxford: Blackwell.

Brockelman, W. Y., and Srikosamatar, S. (1984). "Maintenance and evolution of social structure in gibbons," in *The Lesser Apes. Evolutionary and Behavioural Biology*, eds H. Preuschoft, D. J. Chivers, W. Y. Brockelman, and N. Creel (Edinburgh: Edinburgh University Press), 298–323.

Carden, G. (1983). The non-finite-state-ness of the word formation component. *Linguist. Inq.* 14, 537–541.

Chomsky, N. (1956). Three models for the description of language. *IRE Trans. Inf. Theory* 2, 113–124. doi: 10.1109/TIT.1956.1056813

[CrossRef Full Text](#)

Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: The MIT Press.

Chomsky, N. (2001). "Derivation by phase," in *Ken Hale: A Life in Language*, ed M. Kenstowicz (Cambridge, MA: MIT Press), 1–52.

Chomsky, N. (2008). "On phases," in *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud*, eds R. Freidin, C. Otero, and M.-L. Zubizarreta (Cambridge, MA: MIT Press), 133–166.

Darwin, C. (1871). *The Descent of Man, and Selection in Relation to Sex*. London: John Murray. doi: 10.1037/12293-000

[CrossRef Full Text](#)

Doupe, A. J., and Kuhl, P. K. (1999). Birdsong and human speech: common themes and mechanisms. *Annu. Rev. Neurosci.* 22, 567–631. doi: 10.1146/annurev.neuro.22.1.567

[Pubmed Abstract](#) | [Pubmed Full Text](#) | [CrossRef Full Text](#)

Embick, D. (2010). *Localism versus Globalism in Morphology and Phonology*. Cambridge, MA: MIT Press. doi: 10.7551/mitpress/9780262014229.001.0001

[CrossRef Full Text](#)

Fitch, W. T. (2011). The evolution of syntax: an exaptationist perspective. *Front. Evol. Neurosci.* 3:9. doi: 10.3389/fnevo.2011.00009

[Pubmed Abstract](#) | [Pubmed Full Text](#) | [CrossRef Full Text](#)

Geissmann, T., Bohlen-Eyring, S., and Heuck, A. (2005). The male song of the Javan silvery gibbon (*Hylobates moloch*). *Contrib. Zool.* 74, 1–25.

Geissmann, T., and Orgeldinger, M. (2000). The relationship between duet songs and pair bonds in siamangs, *Hylobates syndactylus*. *Anim. Behav.* 60, 805–809. doi: 10.1006/anbe.2000.1540

[Pubmed Abstract](#) | [Pubmed Full Text](#) | [CrossRef Full Text](#)

Ghazanfar, A. A., Takahashi, D. Y., Mathur, N., and Fitch, W. T. (2012). Cineradiography of monkey lip-smacking reveals putative precursors of speech dynamics. *Curr. Biol.* 22, 1176–1182. doi: 10.1016/j.cub.2012.04.055

[Pubmed Abstract](#) | [Pubmed Full Text](#) | [CrossRef Full Text](#)

Haimoff, E. H. (1984). “Acoustic and organizational features of gibbon songs,” in *The Lesser Apes. Evolutionary and Behavioural Biology*, eds H. Preuschoft, D. J. Chivers, W. Y. Brockelman, and N. Creel (Edinburgh: Edinburgh University Press), 333–353.

Haimoff, E. H. (1985). The organization of song in Mueller's gibbon (*Hylobates muelleri*). *Int. J. Primatol.* 6, 173–192. doi: 10.1007/BF02693652

[CrossRef Full Text](#)

Hale, K., and Keyser, J. (1993). “On argument structure and the lexical expression of syntactic relations,” in *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, eds K. Hale and J. Keyser (Cambridge, MA: MIT Press), 53–108.

Halle, M., and Marantz, A. (1993). “Distributed morphology and the pieces of inflection,” in *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, eds K. Hale and S. J. Keyser (Cambridge, MA: MIT Press), 111–176.

Hauser, M. D., Chomsky, N., and Fitch, W. T. (2002). The faculty of language: what is it, who has it, and how did it evolve? *Science* 298, 1569–1579. doi: 10.1126/science.298.5598.1569

[Pubmed Abstract](#) | [Pubmed Full Text](#) | [CrossRef Full Text](#)

Hurford, J. R. (2011). *The Origins of Grammar: Language in the Light of Evolution II*. Oxford: Oxford University Press.

Jackendoff, R. (2010). “Your theory of language evolution depends on your theory of language,” in *The Evolution of Human Language*, eds R. Larson, V. Déprez, and H. Yamakido (Cambridge: Cambridge University Press), 63–72.

Jespersen, O. (1922). *Language: Its Nature, Development, and Origin*. London: George Allen and Unwin Ltd.

Jones, C. H., and Pinel, J. P. J. (1990). Linguistic analogies and behavior - the finite-state behavioral grammar of food-hoarding in hamsters. *Behav. Brain Res.* 36, 189–197. doi: 10.1016/0166-4328(90)90056-K

[Pubmed Abstract](#) | [Pubmed Full Text](#) | [CrossRef Full Text](#)

Key, H. (1965). Some semantic functions of reduplication in various languages. *Anthropol. Linguist.* 7, 88–101.

Langendoen, D. T. (1975). Finite-state parsing of phrase-structure languages and the status of readjustment rules in grammar. *Linguist. Inq.* 6, 533–554.

Langendoen, D. T. (1981). The generative capacity of word-formation components. *Linguist. Inq.* 12, 320–322.

Marantz, A. (1982). Re Reduplication. *Linguist. Inq.* 13, 435–482.

Marantz, A. (1997). “No escape from syntax: don't try morphological analysis in the privacy of your own lexicon,” in *University of Pennsylvania Working Papers in Linguistics*, Vol. 4.2 (Philadelphia, PA), 201–225.

Marchand, H. (1969). *The Categories and Types of Present-Day English Word-Formation: A Synchronic-Diachronic Approach*. Munich: Verlag C. H. Beck.

Marler, P. (1970). Birdsong and speech development: could there be parallels? *Am. Sci.* 58, 669–673.

[Pubmed Abstract](#)

Marshall, J. T., and Marshall, E. R. (1976). Gibbons and their territorial songs. *Science* 193, 235–237. doi: 10.1126/science.193.4249.235

[Pubmed Abstract](#) | [Pubmed Full Text](#) | [CrossRef Full Text](#)

McCarthy, J. J., and Prince, A. (1995). “Faithfulness and reduplicative identity,” in *Papers in Optimality Theory*. *University of Massachusetts Occasional Papers in Linguistics 18*, eds J. Beckman, L. W. Dickey, and S. Urbanczyk (Amherst, MA: Graduate Linguistic Student Association), 249–384.

McCarthy, J. J., and Prince, A. (1999). “Prosodic morphology (1986),” in *Phonological Theory: The Essential Readings*, ed J. Goldsmith (Malden, MA: Blackwell), 238–288.

Mitani, J. C. (1985). Gibbon song duets and intergroup spacing. *Behaviour* 92, 59–96. doi: 10.1163/156853985X00389

[Pubmed Abstract](#) | [Pubmed Full Text](#) | [CrossRef Full Text](#)

Mitani, J. C. (1988). Male gibbon (*Hylobates agilis*) singing behavior: natural history, song variations and function. *Ethology* 79, 177–194. doi: 10.1111/j.1439-0310.1988.tb00710.x

[CrossRef Full Text](#)

Miyagawa, S. (2010). *Why Agree? Why Move?: Unifying Agreement-Based and Discourse-Configurational Languages*. Cambridge, MA: The MIT Press.

Miyagawa, S., Berwick, R. C., and Okanoya, K. (2013). The emergence of hierarchical structure in human language. *Front. Psychol.* 4:71. doi: 10.3389/fpsyg.2013.00071

[Pubmed Abstract](#) | [Pubmed Full Text](#) | [CrossRef Full Text](#)

Moravcsik, E. (1978). "Reduplicative constructions," in *Universals of Human Language*, Vol. 3: *Word Structure*, ed J. H. Greenberg (Stanford, CA: Stanford University Press), 297–334.

Moro, A. (2014a). On the similarity between syntax and actions. *Trends Cogn. Sci.* 18, 109–110. doi: 10.1016/j.tics.2013.11.006

[Pubmed Abstract](#) | [Pubmed Full Text](#) | [CrossRef Full Text](#)

Moro, A. (2014b). Response to Pulvermüller: the syntax of actions and other metaphors. *Trends Cogn. Sci.* 18, 221. doi: 10.1016/j.tics.2014.01.012

[Pubmed Abstract](#) | [Pubmed Full Text](#) | [CrossRef Full Text](#)

Narita, H., Iijima, K., and Sakai, K. (2014). Ningen-gengo-no kiso-ha fukuzatsu nanoka? (Is the basis of human language complex?). *Brain Nerve* 66, 276–279.

Nottebohm, F. (1975). Continental patterns of song variability in *Zonotrichia capensis*: some possible ecological correlates. *Am. Nat.* 109, 605–624. doi: 10.1086/283033

[CrossRef Full Text](#)

Okanoya, K. (2002). "Sexual display as a syntactical vehicle: The evolution of syntax in birdsong and human language through sexual selection," in *The Transition to Language*, ed A. Wray (Oxford: Oxford University Press), 46–63.

Pulvermüller, F. (2014). The syntax of action. *Trends Cogn. Sci.* 18, 219–220. doi: 10.1016/j.tics.2014.01.001

[Pubmed Abstract](#) | [Pubmed Full Text](#) | [CrossRef Full Text](#)

Raemaekers, J. J., Raemaekers, P. M., and Haimoff, E. H. (1984). Loud calls of the gibbon (*Hylobates lar*): repertoire, organization and context. *Behaviour* 91, 146–189. doi: 10.1163/156853984X00263

[CrossRef Full Text](#)

Raimy, E. (2000). *The Phonology and Morphology of Reduplication*. Berlin: Mouton de Gruyter. doi: 10.1515/9783110825831

[CrossRef Full Text](#)

Rauschecker, J. P. (2012). Ventral and dorsal streams in the evolution of speech and language. *Front. Evol. Neurosci.* 4:7. doi: 10.3389/fnevo.2012.00007

[Pubmed Abstract](#) | [Pubmed Full Text](#) | [CrossRef Full Text](#)

Riley, J., Greggers, U., Smith, A., Reynolds, D., and Menzel, R. (2005). The flight paths of honeybees recruited by the waggle dance. *Nature* 435, 205–207. doi: 10.1038/nature03526

[Pubmed Abstract](#) | [Pubmed Full Text](#) | [CrossRef Full Text](#)

Sag, I. A., Wasow, T., and Bender, E. M. (2003). *Syntactic Theory: A Formal Introduction*. Stanford, CA: CSLI Publications.

Seyfarth, R. M., Cheney, D. L., and Marler, P. (1980). Monkey responses to three different alarm calls: evidence of predator classification and semantic communication. *Science* 210, 801–803. doi: 10.1126/science.7433999

[Pubmed Abstract](#) | [Pubmed Full Text](#) | [CrossRef Full Text](#)

Starke, M. (2009). “Nanosyntax: A short primer to a new approach to language,” in *Nordlyd 36.1, special issue on Nanosyntax*, eds P. Svenonius, G.

Ramchand, M. Starke, and K. T. Taraldsen (Tromsø: CASTL), 1–6.

Sybesma, R. (2007). Whether we tense-agree overtly or not. *Linguist. Inq.* 38, 580–587. doi: 10.1162/ling.2007.38.3.580

[CrossRef Full Text](#)

Tattersall, I. (2009). “Language and the origin of symbolic thought,” in *Cognitive Archaeology and Human Evolution*, eds S. A. De Beaune, F. L. Coolidge, and T. G. Wynn (New York, NY: Cambridge University Press), 109–116.

Urbanczyk, S. (2007). “Reduplication,” in *The Cambridge Handbook of Phonology*, ed P. De Lacy (Cambridge: Cambridge University Press), 473–494.

**Keywords:** biolinguistics, language evolution, linguistics, birdsong, agreement, movement in language

**Citation:** Miyagawa S, Ojima S, Berwick RC and Okanoya K (2014) The integration hypothesis of human language evolution and the nature of contemporary languages. *Front. Psychol.* 5:564. doi: 10.3389/fpsyg.2014.00564

**Received:** 24 January 2014; **Accepted:** 21 May 2014;

**Published online:** 09 June 2014.

Edited by:

[Andrea Moro](#), Institute for Advanced Study IUSS Pavia, Italy

Reviewed by:

[Itziar Laka](#), University of the Basque Country, Spain

[Ina Bornkessel-Schlesewsky](#), University of Marburg, Germany

**Copyright** © 2014 Miyagawa, Ojima, Berwick and Okanoya. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

\***Correspondence:** Shigeru Miyagawa, Department of Linguistics and Philosophy, Massachusetts Institute of Technology, 32D-808/14N-305, Cambridge, MA 02139, USA e-mail: miyagawa@mit.edu