Evolinguistics Workshop 2019

Hierarchy, intention sharing, and language evolution:
Beyond interdisciplinary conceptual barriers

May 25-26, 2019
21KOMCEE East K211 & 212
University of Tokyo, Komaba

Funded by MEXT Grant-in-Aid for Scientific Research
on Innovative Areas #4903:
Evolinguistics: Integrative Studies of Language Evolution for Co-Creative Communication
in collaboration with
Language and Music in Cognition, U of Cologne.

http://evolinguistics.net/en/
http://musikwissenschaft.phil-fak.uni-koeln.de/34666.html?&L=1

Program

May 25 (Sat)

Room K211

13:00 Opening

13:15 Cedric Boeckx (ICREA/U Barcelona)
Language: Confluent evolutionary trajectories

14:15 Break

Room K211

14:25 Ryosuke Tachibana (U Tokyo) & Kazuo Okanoya (U Tokyo)
Exploring syntactical structure of mesoscopic neural activity in songbird brain

14:45 Katsuhiro Sano (Tohoku U)
Emergence of the complex multi-faceted sequences in human tools

15:05 Yuko Yoshimura (Kanazawa U) & Mitsuru Kikuchi (Kanazawa U)
Neural basis of 'intention sharing' in children and adults with autism spectrum disorder

15:45 Junya Morita (Shizuoka U)
Modeling shared communication systems formed with autistic cognition

Room K212

14:25 Daiki Matsumoto (Kyoto U)
Distinctness, recursiveness and hierarchy by Select and Merge

14:45 Masanobu Sorida (independent)
Perceiving syntactic objects

15:05 Masakazu Kuno (Waseda U)
Mathematical exploration of minimalist syntax: A preliminary study

15:45 Koji Hoshi (Keio U)
An exploration into the relation between Merge and categorization in evolinguistics

16:25 Break

Room K211

16:40 Erin Hecht (Harvard U)
Hands, tools, and words: Adaptation and exaptation in human brain evolution

18:00 Reception (at Italian Tomato)

May 26 (Sun)

Room K211

9:30 Julia Uddén (Stockholm U)
Two major complexities to evolve: Sentence structure building in the dorsal pathway and the development of pragmatics around puberty

10:30 Break

Room K211

10:40 Michiru Makuuchi (National Rehabilitation Center for Persons with Disabilities)
Hierarchical structure building in symbol sequence, i.e., language, mathematics and drawing

11:00 Rie Asano (U Cologne)
Syntax in language and music from a perspective of comparative biomusicology

11:40 Kazumi Taniguchi (Kyoto U)
Sharing patterns, sharing intention: A view from dialogic syntax

Room K212

10:40 Haruka Fujita (Kyoto U)
The co-evolution of internalization and externalization in human lexicon

11:00 Tzu-Yin Chen (U Tokyo & RIKEN), Yuki Hirose (U Tokyo) & Takane Ito (U Tokyo)
The facilitative effect of prosodic information in online lexical processing: Evidence from Mandarin Chinese Tone 3 sandhi

11:20 Sandiway Fong (U Arizona), Robert Berwick (MIT) & Jason Ginsburg (Osaka Kyoiku U)
The combinatorics of Merge and workspace tight-sizing

11:40 Genta Toya (JAIST), Rie Asano (U Cologne) & Takashi Hashimoto (JAIST)
The reason of building hierarchical structure: From the view of recursive combination as an internal operation

12:00 Lunch break

Room K211

13:30 John Du Bois (UC Santa Barbara)
Coupling for life: Co-evolution of grammar and prosody in dialogic interaction

14:30 Sabrina Engesser (U Zurich)
Vocal combinations in birds: Implications for language evolution

15:30 Break

15:45 Toshitaka Suzuki (Kyoto U)
Referentiality and compositionality in bird calls

16:25 Kazuo Okanoya (U Tokyo)
"Syntax" in animal vocalizations: Limitations and perspectives

17:05 Discussion

17:30 Closing

May 25 Sat

Room K211			Room K212		
13:00	13:15	Opening			
13:15	14:15	C. Boeckx			
14:15	14:25	*Break*			
14:25	14:45	R.Tachibana & K.Okanoya	14:25	14:45	D. Matsumoto
14:45	15:05	K. Sano	14:45	15:05	M. Sorida
15:05	15:45	Y. Yoshimura & M. Kikuchi	15:05	15:45	M. Kuno
15:45	16:25	J. Morita	15:45	16:25	K. Hoshi
16:25	16:40	*Break*			
16:40	17:40	E. Hecht			
18:00	20:00	Reception (Italian Tomato)			

May 26 Sun

Room K211			Room K212		
9:30	10:30	J. Uddén			
10:30	10:40	*Break*			
10:40	11:00	M. Makuuchi	10:40	11:00	H. Fujita
11:00	11:40	R. Asano	11:00	11:20	T-Y.Chen, Y. Hirose & T. Ito
11:40	12:00	K. Taniguchi	11:20	11:40	S. Fong, R. Berwick & J. Ginsburg
			11:40	12:00	G. Toya, R. Asano & T. Hashimoto
12:00	13:30	*Lunch break*			
13:30	14:30	J. Du Bois			
14:30	15:30	S. Engesser			
15:30	15:45	*Break*			
15:45	16:25	T. Suzuki			
16:25	17:05	K. Okanoya			
17:05	17:30	Discussion			
17:30		Closing			

Online Presentations

1 Dieter Hillert (UC San Diego & San Diego State U)
We were not alone: The gradual evolution of the language capacity

2 Sebastian Klaßmann (U Cologne)
Logical Cells: Combinatory Categorial Grammar as a unified framework for describing melodic improvisation and harmonic analysis

Saturday, May 25th, 2019

Language: Confluent Evolutionary Trajectories

Cedric Boeckx[*1,2,3]

[*]Corresponding Author: cedric.boeckx@ub.edu
[1]ICREA (Catalan Institute for Advanced Studies and Research)
[2]Section of General Linguistics, Universitat de Barcelona, Barcelona, Spain
[3]Universitat de Barcelona Institute for Complex Systems (UBICS)

Everyone these days takes the faculty of language to consist of several parts, and assumes that these parts must have come together at some point in the course of evolution. Disagreements arise regarding the specific identity of these parts and the timing of their emergence. Here, based on work done in my group, I will take a particular stance on these issues, but I will focus more on another, related, question: did these component parts appear in a specific order, and did they begin to cluster in a particular order—can this be rationalized and tested?

Exploring syntactical structure of mesoscopic neural activity in songbird brain

Ryosuke O. Tachibana[1*] and Kazuo Okanoya[1]

*Corresponding Author: rtachi@gmail.com

[1]Department of Life Sciences, Graduate School of Arts & Sciences, The University of Tokyo, Japan

Birdsong is a good model to study neural mechanisms for skilled motor sequences. Bengalese finches' songs consist of a sequence of various sound elements (syllables). Their songs have been reported to show stochastic syllable transitions at branching points in the sequence (Fig. 1A), as results of sequence analyses based on acoustical similarity among syllables. How does this syllable transition rule, or song syntax, reflect the sequential structure of internal premotor activities? The song motor pathway in the bird brain includes two premotor nuclei: HVC (high vocal center) and RA (robust nucleus of arcopallium). HVC neurons projecting to RA (HVC_{RA}) are known to exhibit burst firings sparsely at specific timings in song sequence, as a driving source of syllable productions. Thus, measuring activities of HVC_{RA} neurons during singing will reveal neuronal mechanisms for producing the motor sequence with such complex transition syntax. In the present study, we measured neuronal activities from multiple HVC_{RA} neurons of freely-moving birds by the calcium imaging technique. Fluorescent calcium indicators (GCaMP) were expressed specifically in HVC_{RA} neurons using the adeno-associated viral vectors (Fig. 1B,C). After several weeks of waiting for the expression, we started to image spontaneous song productions by a light-weight miniaturized fluorescent microscope which was mounted onto bird's head. Preliminary results showed that each HVC_{RA} neuron fired at different timings in the song with reflecting the sequential transition pattern. Further analyses on consistency in transition patterns between the neuronal burst sequences and produced songs will be discussed.

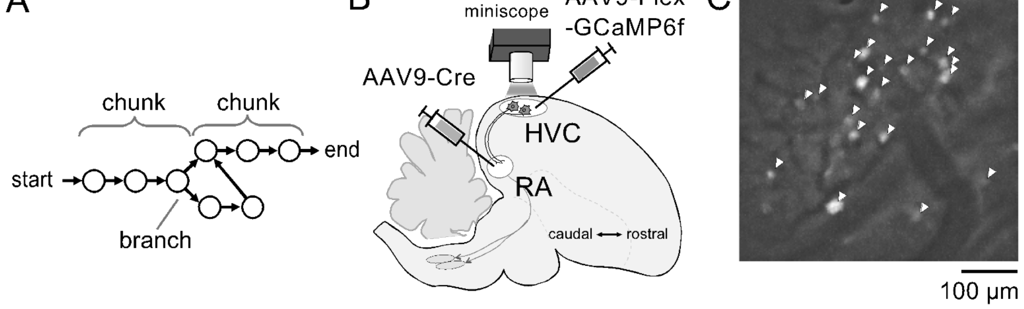


Figure 1. In vivo Ca2+ imaging of premotor neuronal activities in the Bengalese finch's brain. **A**. Schematic drawing of expected syntactic structure of Bengalese finches' songs. **B**. Viral injection for expressing calcium indicator (GCaMP) in the premotor nucleus HVC. **C**. Example of detected fluorescent responses from premotor neurons.

Emergence of the complex multi-faceted sequences in human tools

Katsuhiro Sano[*1]

[*]Corresponding Author: sano.k@tohoku.ac.jp

[1] Center for Northeast Asian Studies, Tohoku University, Sendai, Japan

The oldest stone tools appeared at ~2.6 Ma in East Africa and then the first genus *Homo*, *Homo habilis*, emerged at ~2.4 Ma. *Homo habilis* then evolved into *Homo erectus*, which coincides with the emergence of the first designed stone tools, such as handaxes, at around 1.75 Ma (Beyene et al., 2015). Although advanced handaxes show three-dimensional symmetricity and refined workmanship, they were used without any handles or shafts. The procedures of the production and use of these tools are quite simple. They were hold by a hand and used for cutting or chopping.

On the other hand, Neanderthals in Europe and an almost contemporaneous *Homo* group in East Africa, as late *Homo heidelbergensis* or early *Homo sapiens*, started to haft a stone tool onto a wooden shaft between 300 ka and 200 ka (Mazza et al., 2006; Brooks et al., 2018). They used birch bark tar, bitumen, or resin as adhesives to connect two different materials. So, the procedures of the production and use of the composite tools are more complex than those of choppers or handaxes. However, their composite tools do not share the same hierarchy with the syntax of language.

Nevertheless, the complex projectile technology using a spearthrower or a bow involves an intricate procedure and requires multistage planning (Lombard and Haidle, 2012). Therefore, the emergence of the projectile technology may reflect the emergence of the cognitive ability for controlling merge in language. Hence, we are trying to detect the origin and dispersal process of the complex projectile technology. The paper presents current results of our studies.

References

Beyene, Y., Sano, K., Asfaw, B., Suwa, G., 2015. Technological and Cognitive Advances Inferred from the Konso Acheulean Assemblages. In: Beyene, Y., Asfaw, B., Sano, K., Suwa, G. (Eds.), *Konso-Gardula Research Project Volume 2. Archaeological Collection: Background and the Early Acheulean Assemblages*. The University Museum, The University of Tokyo, Bulletin No. 48, pp. 65–81.

Brooks, A.S., Yellen, J.E., Potts, R., Behrensmeyer, A.K., Deino, A.L., Leslie, D.E., Ambrose, S.H., Ferguson, J.R., d'Errico, F., Zipkin, A.M., Whittaker, S., Post, J., Veatch,

E.G., Foecke, K., Clark, J.B., 2018. Long-distance stone transport and pigment use in the earliest Middle Stone Age. *Science*. 360, 90–94.
Lombard, M., Haidle, M.N., 2012. Thinking a bow-and-arrow set: cognitive implications of Middle Stone Age bow and stone-tipped arrow technology. *Cambridge Archaeological Journal*. 22, 237–264.
Mazza, P.P.A., Martini, F., Sala, B., Magi, M., Colombini, M.P., Giachi, G., Landucci, F., Lemorini, C., Modugno, F., Ribechini, E., 2006. A new Palaeolithic discovery: tar-hafted stone tools in a European Mid-Pleistocene bone-bearing bed. *Journal of Archaeological Science*. 33, 1310–1318.

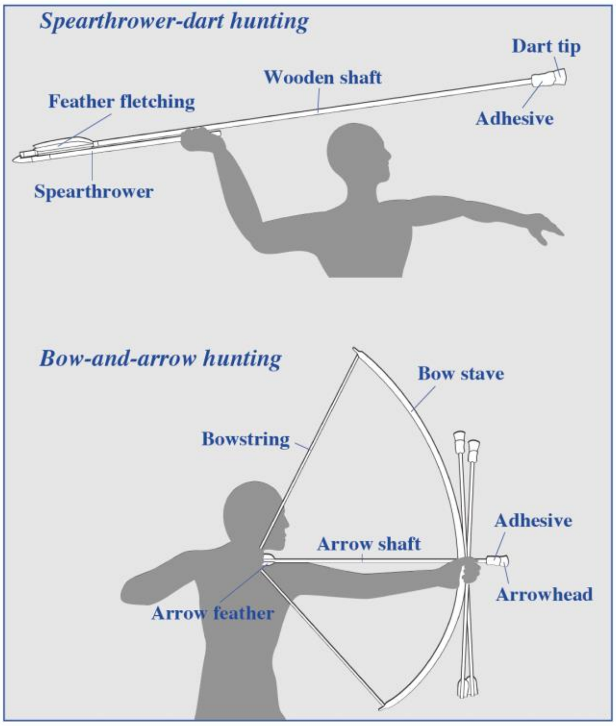


Figure 1. Spearthrower-dart and bow-and-arrow technologies. Both hunting weapons comprise several parts and require multistage planning.

Neural basis of 'intention sharing' in children and adults with autism spectrum disorder

Yuko Yoshimura[*1,2] , Mitsuru Kikuchi[2,3]

[*]Corresponding Author: yukuchen@staff.kanazawa-u.ac.jp
[1] Institute of Human and Social Sciences, Kanazawa University, Kanazawa, Japan
[2]Research center for child mental development, Kanazawa University, Kanazawa, Japan
[3] Department of Psychiatry & Behavioral Science, Kanazawa University, Kanazawa, Japan

In the process of human ontogeny, brain changes in early childhood are related not only to their own survival, but also to socializing with the society and prospering of species. In recent years, advances in brain imaging technology have led to an understanding of the maturation of human brain structure and white matter fibers from the fetal stage. However, functional activities of the brain related to language development and sociality in early childhood are still unclear because of the difficulty of physiological examination at the time of observation. Autism spectrum disorder (ASD), which has been reported to have a prevalence rate of 1% or more in recent years, is a neurodevelopmental disorder mainly related to sociality and communication. Language developmental delay, pragmatic language impairment are observed in many individuals with ASD. We hypothesize that various perceptual processing features, such as speech processing in the brain cortex in early childhood, are associated with difficulties in acquiring typical language development and communication skills. In other words, it is thought that the atypical processing at the perceptual level, which is considered to be low-order processing in the brain, creates diverse personalities of children defined as autism spectrum and typical development, and influences the interaction with society.
Our purpose is to elucidate the relationship between brain activity and development evoked by auditory stimuli (especially 'intention sharing' voice) using child- customized magnetoencephalography (MEG) system. We introduce the differences in brain response to intention sharing voice between children with ASD and typical developing children (Yoshimura et al., 2013; Yoshimura et al., 2016; Yoshimura et al., 2017). Furthermore, we report the results of examining the brain regions involved in the processing of intention sharing by examining the relationship with serotonin transporter using Positron Emission Tomography (PET) in adults with ASD (Figure 1).

Yoshimura Y, Kikuchi M, et al., (2013). Atypical brain lateralisation in the auditory cortex and language performance in 3- to 7-year-old children with high-functioning autism spectrum disorder: a child-customised magnetoencephalography (MEG) study. *Molecular autism.* 4(1) 38. doi: 10.1186/2040-2392-4-38

Yoshimura Y, Kikuchi M, et al., (2016). Atypical development of the central auditory system in young children with Autism spectrum disorder. *Autism Research.* 9(11):1216-1226.
Yoshimura Y, Kikuchi M, et al., (2017). Altered human voice processing in the frontal cortex and a developmental language delay in 3- to 5-year-old children with autism spectrum disorder. *Scientific reports.* 7(1) 17116. doi: 10.1038/s41598-017-17058-x

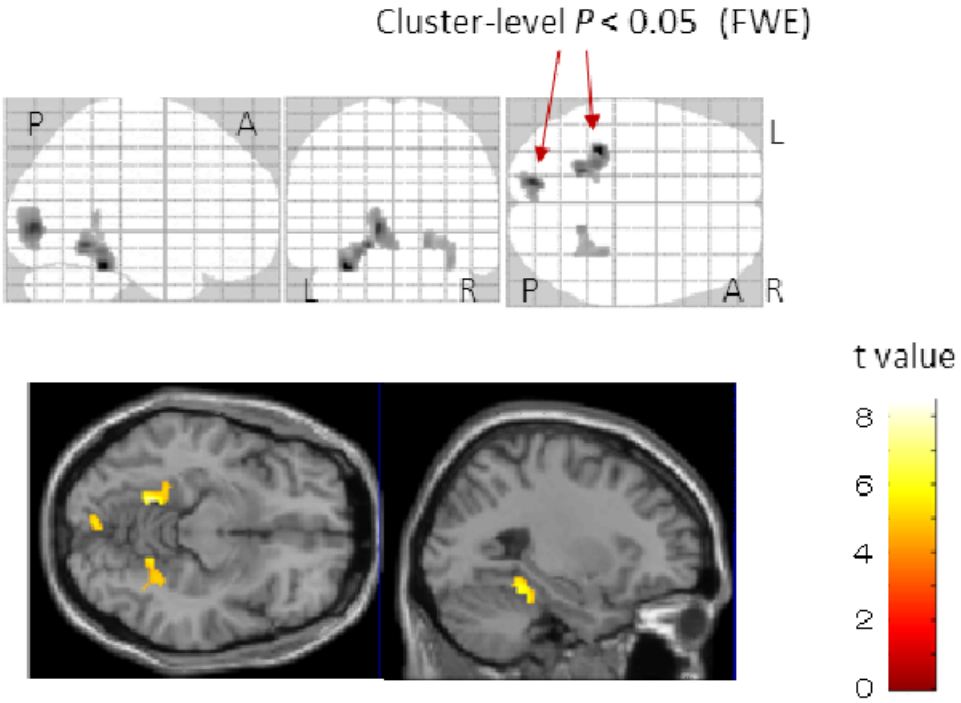


Figure 1. Correlation between serotonin transporter and brain response evoked by 'intention sharing' voice in the right hemisphere.

Modeling Shared Communication Systems formed with Autistic Cognition

Junya Morita [*1]

[*]Junya Morita: j-morita@inf.shizuoka.ac.jp
[1]Faculty of Informatics, Shizuoka University, Hamamatsu, Japan

What is the biological, cognitive and social mechanisms leading to shared communication systems in the society? In this research, we focus on autistic cognition as a personal characteristic related to the question above. This characteristic may lead to difficulties in social life, but it is also suggested that it has been involved in important innovation leading to the present age. Therefore, rather than considering the autism spectrum as just an obstacle, we see it as an individual characteristic that plays a role in forming a new communication system. In this research, we use a communication game that generates simple artificial languages as the experimental environment (Konno, et al., 2013), and associate autism spectrum index (AQ: Autism Spectrum Quotient) as individual characteristics with data obtained from games. As a result of analyzing the obtained data, we found marginally positive correlations between the AQ and the success of this communication game. The detailed analysis also revealed the significant correlation between the patterned thinking style measured with AQ and the success of communication in the early and final stage of the game. From these result, we hypothesize that the AQ plays the role of scaffolding to develop a system used to share intentions, and also it contributes to maintain the constructed communication systems in the society. In the future work, we will conduct a computational modeling extending our previous work (Morita, et al., 2017) to represent and test the hypothesis above.

Konno, T., Morita, J., and Hashimoto, T. (2013). Symbol communication systems integrate implicit information in coordination tasks. *Advances in cognitive neurodynamics* (iii), 453-459.

Morita, J.,Konno, T., Okuda, J., Samejima, K., Li, G., Fujiwara, M., and Hashimoto, T. (2017). Implicit memory processing in the formation of a shared communication system. In *Proceedings of 15th International Conference on Cognitive Modeling*, pp.19—24.

Distinctness, Recursiveness and Hierarchy by Select and Merge

Daiki MATSUMOTO[1]

*matsumoto.daiki.24x@st.kyoto-u.ac.jp

[1]Graduate School of Human and Environmental Studies, Kyoto University, Kyoto, Japan

In the field of Generative Grammar, it is widely accepted that the only syntactic computation *Merge*, which takes two distinct items as its inputs and creates a set out of them, always yields a hierarchically-structured expression (Chomsky 1995). As Boeckx (2014) insists, *Merge* in this sense has no way of referring to any kind of semantic/phonological properties of its inputs. Simply put, *Merge* is blind. It just combines two items and yields a hierarchy. One critical fact here is that generative grammarians have never asked one crucial question: how can the distinctness of the inputs to *Merge* be secured? One of the main reasons for this ignorance is, I think, that despite the caveat aptly made by Boeckx, they still subconsciously rely on the inner properties of its inputs. Put differently, the distinctness in question is just presupposed because, for example, the words *kick* and *John* are both phonologically and semantically different. It is true that these two are distinct to each other from our eyes. However, as Boeckx argues, *Merge* itself has no way of knowing this. Based on these observations, I argue that other than *Merge*, another operation which I call *Select* is necessary. It maps the copy of a concept stored in the Lexicon LEX onto the workspace on which *Merge* operates (Chomsky et al. 2019), securing one necessary condition for hierarchy: distinctness. I further insist that *Select*, together with Boeckx's definition of *phases*, forces us to look at the role of the interface between syntax and LEX as a memory savor for syntax, securing another condition for hierarchy: recursiveness. Finally, I propose a(n alternative) language system in which I assume that syntax is inside the conceptual-intentional (CI) system, and the interface between CI and sensorimotor (SM) systems is what we call LEX.

Boeckx, C. (2014). *Elementary syntactic structures: Prospects of a feature-free syntax*. Cambridge, UK: Cambridge University Press.

Chomsky, N. (1995). *The minimalist program*. Cambridge, MA, US: MIT Press.

Chomsky, N., Gallego, Á. J., and Ott, D. (2019). Generative grammar and the faculty of language: Insights, questions and challenges. Unpublished manuscript. https://ling.auf.net/lingbuzz/003507

Perceiving Syntactic Objects

Masanobu Sorida[*1]

[*]Corresponding Author: masanobusorida@gmail.com
[1]Independent Scholar, Japan

In this talk, I consider the origin of *labeling*, an essential component of the language faculty, and argue that it originates from a domain-general cognitive capacity that serves to distinguish *figure* and *ground*. A common feature of figure-ground segmentation (FGS) and labeling can be stated as follows: Given a structured object, the most prominent element within it, which can be found with least efforts (minimal search), is chosen and is regarded as the central element for interpretation (figure for perception, label for language). Concrete data make the parallelism clear. FGS and labeling sometimes show ambiguity. The image (1) is Rubin's vase, which is ambiguous between a vase and two faces, a celebrated illustration of figure-ground reversal. A parallel example can be found in language. Embedded *what*-clauses can be interpreted in two ways: an indirect question (2a) or a free relative (2b). This is explained in terms of labeling. The structure of the *what*-clause is in (3), where minimal search is ambiguous between D (for the free relative) and C (for the indirect question). Thanks to the selectional properties of the main verbs, however, the ambiguity does not arise here. A real ambiguity is observed in (4). The *what*-clause can be interpreted either way. Thus "Labeling reversal" takes place just as in Rubin's vase. A further parallelism can be shown by observing that FGS and labeling obey the same constraint. In (5), two images of Rubin's vase are juxtaposed, and you are forced to interpret them in parallel: If you see, say, a vase on the left, you must see another vase on the right. Labeling observes the same constraint. When two *what*-clauses are apposed, they are forced to be interpreted in parallel: When the first conjunct is interpreted, say, as an indirect question, the second one should also be.

Boeckx, Cedric. (2009). *Language in Cognition: uncovering mental structures and the rules behind them.* Wiley-Blackwell, Malden.

Cecchetto, Carlo and Caterina, Donati. (2015). *(Re) labeling*. MIT Press, Cambridge, MA.

Chomsky, Noam. (2013). Problems of Projection. *Lingua* 130, 33-49.

Donati, Caterina. (2006). On wh-head movement. In *Wh-movement: Moving on*, ed. by Lisa Lai-Shen Cheng and Norbert Corver, 21-46. Cambridge, MA: MIT Press.

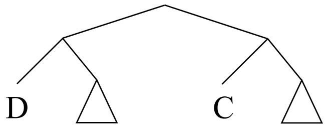
Hoffman, Donald. (1998). *Visual intelligence*. New York. Norton.

(1)

(2)

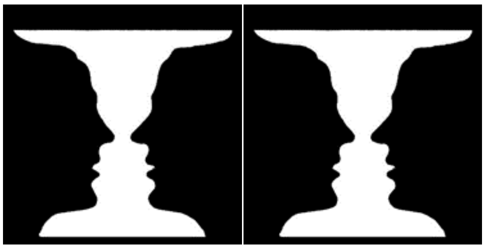
a. John wonder [what Sue ate for lunch].
b. John devoured [what Sue had given him].

(3)



(4) Jim saw [what Sue was holding in her hand].

(5)



(6) Jim saw [what Sue was holding in her hand] and [what Bill was clutching in his teeth].

Mathematical Exploration of Minimalist Syntax: A Preliminary Study

Masakazu Kuno*[1]

*Corresponding Author: kuno@waseda.jp

[1]Department of Education, Waseda University, Tokyo, Japan

This paper begins with the following two questions about mapping between syntax and mathematics.

(1) What is the set G such that G is manipulated by syntactic operations?

(2) What syntactic operations apply to G?

Our proposal is as follows:

(3) G is the set of lexical items and phrases, call them *Syntactic Objects SOs*, (phrases result from Multiplication Merge defined in (4)).

(4) G is manipulated by Multiplication Merge (MM).

MM (a,b) = a*b, for arbitrary elements a, b $\in$ G.

MM corresponds to multiplication operation but unlike multiplication of numbers, MM of SOs does not satisfy the associative law, i.e., not always (a*b)*c = a*(b*c). This is clear as phrases like *black taxi driver* exhibit semantic ambiguity, depending on which SO undergoes MM with which one, viz., (black*(taxi*driver)) or ((black*taxi)*driver).

MM is not the sole type of Merge. Language manifests a structure that indicates the distributive property (i.e., a*(b+c) = a*b+a*c), one straightforward case being ATB movement, as in (5).

(5) John bought fish and cooked it (= John bought fish and John cooked it.)

Given this property, mathematical consideration suggests (6).

(6) Language has a type of Merge corresponding to addition, call it Addition Merge (AM).

AM (a,b) = a+b, for arbitrary elements a, b $\in$ G.

Thus, ATB example in (5) can be represented as the formula in (6), which undergoes formula deformation by factorization.

(6) (John*bought*fish) + (John*cooked*it) *and* is taken as the application of AM.

= John*{(1*bought*fish) + (1*cooked*it)} Factorization by *John*

We extend this analysis to DP movement to Spec-TP and Wh-movement to Spec-CP. They can be analysed as factorization by DP/Wh-phrase or one by T/C; the former corresponds to overt movement and the latter covert movement (Chomsky 1995). We will show that most, if not all, of the syntactic principles naturally follow.

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

An Exploration into the Relation between Merge and Categorization in Evolinguistics

Koji Hoshi*[1]

*khoshi@a7.keio.jp

[1]Faculty of Economics, Keio University, Yokohama, Japan

Providing an evolutionarily adequate account of how the particular form of hierarchical structuring in human language (characterized by Merge) evolved has been one of the most significant and fundamental issues in evolinguistics in gaining a better understanding of origins and evolution of our species (Fujita, 2009 inter alia). With this backdrop in mind, extending and refining the hypothesis of categorization origin of Merge in Hoshi (2018), which is an elaboration of Lenneberg's (1967) idea on the evolution of human language in modern theoretical linguistic terms, I will claim that Merge as the hierarchical structure-building operation was derived from the categorization operation, which is independently and widely observed in not only humans but non-human animals as well (Murphy, 2010 inter alia). More specifically, I will propose that Merge emerged in Homo sapiens as part of the evolutionary recombination of the "old traits" in the categorization operation, while preserving the latter process per se, viz., an array of "concepts" and a set-formation operation, affecting the nature of categorization in our species in a fundamental fashion. Given that categorization is at work both in the sensori-motor (SM) system (perceptual categorization and action categorization (Jackendoff, 2007 inter alia)) and in the conceptual-intentional/thought system (conceptual categorization (Lenneberg, 1967 inter alia)), I will also consider implications of my hypothesis in connection with other approaches to the evolutionary origin of Merge. In particular, I will point out a possibility that Merge was of both SM and C-I origins simultaneously at the stage of protolanguage in our human ancestors, which is in line with the motor control origin of Merge extensively discussed in the literature (Fujita, 2009 et seq. and references cited therein).

Fujita, K. (2009). A prospect for evolutionary adequacy: Merge and the evolution and development of human language. *Biolinguistics,* 3, 128-153.

Hoshi, K. (2018). Merge and labeling as descent with modification of categorization: A neo-Lennebergian approach. *Biolinguistics,* 12, 39-54.

Jackendoff, R. (2007). *Language, Consciousness, Culture: Essays on Mental Structure.* Cambridge, MA: MIT Press.

Lenneberg, E. H. (1967). *Biological Foundations of Language.* New York: John Wiley & Sons.

Murphy, G. L. (2010). What are categories and concepts? In D. Mareschal, P. C. Quinn & S. E.G. Lea (eds.), *The Making of Human Concepts*, 11-28. Oxford: Oxford University Press.

Hands, Tools, and Words: Adaptation and Exaptation in Human Brain Evolution

Erin Hecht [*1]

[*]Corresponding Author: erin_hecht@fas.harvard.edu
[1] Department of Human Evolutionary Biology, Harvard University, Cambridge, USA

Chimpanzee stone tool use has remained essentially unchanged for at least 4300 years. In that time, humans have gone from the Bronze Age to Pluto. How have we so drastically outpaced our closest living relatives? What is it about the human brain that makes us faster and more prolific social learners, better at innovating upon socially learned actions, and capable of producing rapidly-evolving cumulative culture with products like complex technology and language? This talk will discuss adaptations to the neural circuitry for observing and reproducing others' hand-object actions, and how the elaboration of this basic functionality may have supported the emergence of more complex capacities. Comparative neuroimaging studies have uncovered some properties of frontoparietal circuits that are likely common to all primates, some shared by humans and chimpanzees, and some unique to our species. Experimental archaeology research provides a window on human evolution after we diverged from our primate relatives, including how individual brains are forced to change in order to acquire the skills that were important selective pressures in our evolutionary history, like toolmaking and symbolic communication. Together, these lines of research suggest an evolutionary trend toward elaboration in dorsal-stream, "vision-for-action" circuitry. These changes seem tuned to support increased integration of hierarchical conceptualizations of action goals with concrete kinematic, proprioceptive, and spatio-temporal details – a function which would be increasingly important during the evolution of intentional, instrumental behaviors where complex sequences of fine motor actions are acquired via social transmission, as occurs in gesture, tool use, and language.

Sunday, May 26th, 2019

Two major complexities in language evolution: sentence structure building in the dorsal pathway and the development of pragmatics around puberty

Julia Udden*[1]

*Corresponding Author: julia.udden@psychology.su.se
[1]Department of Psychology, Stockholm University, Stockholm, Sweden

Human neurobiology distinguishes itself in specific ways from the non-human primate neurobiology and the discrepancy is in no small part localized to language related networks, including the arcuate fasciculus (left dorsal language pathway). I will present experiments taking a closer look on the function of this pathway. It is generally assumed that structure building processes during spoken and written comprehension are subserved by modality-independent lexical, morphological, grammatical, and conceptual processes. I will mainly present a large-scale neuroimaging study (fMRI, N=204) on whether the unification of sentence structure is supramodal in this sense, testing if observations replicate across written and spoken sentence materials. The activity in the unification network should increase when it is presented with a challenging sentence structure, irrespective of the input modality. In order to make computationally specific suggestion of challenges during sentence comprehension, we build on well-established findings, inspired by experiments using the artificial grammar learning paradigm. Artificial language processing, just like natural language processing, may be hypothesized to rely on common neural mechanisms for structured sequence processing, for instance localized to the left dorsal language pathway. I build on the well-established finding that multiple non-local dependencies, overlapping in time, are challenging. In other words, this challenge during sentence comprehension relate to the need for keeping some words (more specifically so called "non-attached constituents") online when building a sentence. More generally put, these challenges mark the presence of important aspects of hierarchical processing. We show that the added load of a challenging sentence structure leads to an increased supramodal neural response in the left dorsal language pathway.

Are there other distinguishable brain networks relevant for communication that have undergone recent evolutionary changes? To answer this question, I am developing a new research line on pragmatics, with an evolutionary and developmental perspective. What role does human domestication and increased/prolonged parental investment play for language development? This question suggest that *adolescent* development of language and communication might be a good area for further research. I will present a recent study on the adolescent ability to comprehend and produce requests for feedback through prosodic information.

Hierarchical structure building in symbol sequence, i.e., language, mathematics and drawing

Michiru Makuuchi*[1]

*Corresponding Author: makuuchi-michiru@rehab.go.jp
[1]The National Rehabilitation Centre for Persons with Disabilities, Tokorozawa City, Japan

Hierarchical structure is a significant feature of language, but it is found in other cognitive domain such as music and mathematics. Using fMRI, we have demonstrated Broca's area builds hierarchical structures in sentence processing in German and Japanese, and arithmetic calculation given in the reverse Polish notation. Recently, we found that drawing has hierarchical structure as well based on analysis of composition order of parts in complex object drawing. We believe that semeiotic is the appropriate cover term for these symbol sequence with hierarchical structure. This theoretical foundation will allow us to study human unique cognitive abilities such as language, mathematics, drawing, music, etc. with unified framework.

Syntax in language and music from a perspective of comparative biomusicology

Rie Asano*[1]

*Corresponding Author: rie.asano@uni-koeln.de

[1]Department of Systematic Musicology, University of Cologne, Cologne, Germany

Music is hierarchically structured and its complexity in terms of Chomsky hierarchy corresponds to that of context-free language (e.g., Rohrmeier, Zuidema, Wiggins, & Scharff, 2015). My talk discusses ways to link such abstract theoretical models of cognitive capacities to cognitive and neural processes within a comparative biological framework. First, I introduce different levels to investigate language and music as neurocognitive systems, and elucidate problems of explanatory gap (e.g., Embick & Poeppel, 2015; Marr, 1982). Second, by focusing on syntax, I suggest how language and music can be comparatively investigated at each level as well as across levels. I then discuss the results of two ALE meta-analysis of musical syntactic processing and their relation to neural correlates of language syntactic processing, and introduce a hypothesis to explain the relationship between syntax in language and music. The hypothesis states that language and music share a set of basic computational and neural principles, but differ in their degree of expressions on the motor to cognitive gradient. Finally, I propose a comparative approach putting more focus on cognitive and neural processes by focusing on musical rhythm and the cortico-basal ganglia-thalamocortical circuits, and draw implications for the evolution of capacities accounting for hierarchical complexity in music, language, and action.

Embick, D., & Poeppel, D. (2015). Towards a computational(ist) neurobiology of language: correlational , integrated and explanatory neurolinguistics. *Language, Cognition and Neuroscience*, *30*(4), 357–366. https://doi.org/10.1080/23273798.2014.980750

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. Cambridge, MA: The MIT Press.

Rohrmeier, M., Zuidema, W., Wiggins, G. A., & Scharff, C. (2015). Principles of structure building in music, language and animal song. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *370*(1664), 20140097-. https://doi.org/10.1098/rstb.2014.0097

Sharing Patterns, Sharing Intention: A View from Dialogic Syntax

Kazumi Taniguchi

taniguchi.kazumi.8n@kyoto-u.ac.jp
Graduate School of Human and Environmental Studies, Kyoto University, Kyoto, Japan

This paper aims to offer a brief introduction of "dialogic syntax" (Du Bois, 2014) which points to an intriguing phenomenon called "resonance:" some aspects of utterances are reproduced and shared in spontaneous conversations to give rise to affinity across utterances. This paper focuses on resonance of partially schematic syntactic patterns that emerge as "ad hoc constructions" in discourse (Brône and Zima, 2014), showing conversational data observed in English and Japanese. Such sharing of ad hoc constructions can be seen as cooperative synchronization at the syntactic level, which presupposes (and enhances) sharing of intention among participants of conversations. The phenomenon treated in this paper will shed light on mechanisms by which a language system is shared and established through interactions in a given community, suggesting possibilities of "dialogic bootstrapping" for language evolution.

Brône, G. and E. Zima (2014). "Towards a dialogic construction grammar: Ad hoc routines and resonance activation." *Cognitive Linguistics* 25(3), 457-495.

Du Bois, J. W. (2014). "Towards a dialogic syntax." *Cognitive Linguistics* 25(3), 411-441.

The Co-Evolution of Internalization and Externalization in Human Lexicon

Haruka Fujita*[1]

*fujita.haruka.42w@st.kyoto-u.ac.jp

[1]Graduate School of Human and Environmental Studies, Kyoto University, Kyoto, Japan

Although Generative Grammar (GG) is compatible with evolutionary research in that it regards human language as a modular system and makes it clear what to investigate, most researches in GG (e.g. Berwick & Chomsky, 2016; Chomsky, Gallego, & Ott 2019) focus on the hierarchical properties of internal thought and consider communicative functions involving intention sharing as a secondary phenomenon. However, given that language serves the two major functions of internal thought and external communication, the co-evolutionary relation between internalization and externalization needs to be considered. I argue that this relation is especially evident in the emergence of human lexicon. As cognitive semantics has shown (Lakoff 1987, Lakoff & Johnson 1980 inter alia), the concepts underlying human lexicon are classified into concrete concepts and abstract concepts. While concrete concepts have common physical bases, abstract concepts do not have such universal standards and need a specific idealized cognitive model (Lakoff 1987). The first step of the co-evolution of internalization and externalization in human lexicon can be attributed to this property: concrete concepts can be internally combined into complex linguistic forms without externalization because of their concreteness, but abstract concepts first need to be shared through externalization. As a second step of the co-evolution, after the expressive power of language involving syntax and intention sharing has developed enough, various concepts including complex concrete concepts as well as abstract ones will be finally externalized and used for communication. This makes it easier to manipulate them as syntactic objects, for now they have physical entities such as auditory or visual stimuli. Consequently, more complex internal combination of lexical concepts is promoted, and these new concepts will be externalized again. It is this circular interaction of internalization and externalization that has provided human language with a rich and complex lexicon.

Berwick, R. C. & Chomsky, N. (2016). *Why only us: Language and evolution*. Cambridge, MA: MIT Press.

Chomsky, N., Gallego, A. J., & Ott, D. (2019). Generative grammar and the faculty of language: Insights, questions, and challenges (3rd version). Available online. https://ling.auf.net/lingbuzz/003507.

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.

Lakoff, G. & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.

The facilitative effect of prosodic information in online lexical processing: Evidence from Mandarin Chinese Tone 3 sandhi

Tzu-Yin CHEN[*1,2], Yuki HIROSE[1], and Takane ITO[1]

[*]Email: tess5566779@yahoo.co.jp

[1] Language and Information Sciences, The University of Tokyo, Tokyo, Japan

[2]Laboratory for Language Development, Center for Brain Science, RIKEN, Saitama, Japan

Mandarin tone 3 sandhi (T3S) is a phenomenon where a T3 syllable becomes a T2 syllable when it precedes another T3 syllable (e.g., *ni3* 'you' + *hau3* 'good'→*ni2 hau3* 'hello'). The necessary conditions of T3S are that a T3S syllable must be followed by another syllable and that the following syllable must be of the T3 type. First, two visual world paradigm experiments were employed to examine whether a T3S syllable helps to predict a following component's (i) lexical structure and (ii) tone type. Experiment 1 found that when hearing input including a T3S syllable (e.g., *zhu2 sun3* 'bamboo-shoot' → *zhu2 sun2*…), listeners looks more to compound objects (e.g., a picture of a bamboo-shoot and cow combination; *zhu2 sun3* + *ru3 niu2*), suggesting that T3S information anticipates lexical structure. Experiment 2 found no predictive effect of a T3S syllable on upcoming tone type. When a T3S modifier was input (e.g., *zhu2 sun3* →*zhu2 sun2*…), no increase of looks was observed to visual candidates with T3-type head nouns (e.g., a picture of a bamboo-shoot and cow combination; *zhu2 sun3* + *ru3 niu2*) compared to looks to non-T3-type head nouns (e.g., a picture of a bamboo-shoot and butterfly combination; *zhu2 sun3* + *hu2 die2*). These results suggest that T3S does not help predict the phonological form of the following component. Furthermore, T3S application is optional in native use (i.e. a T3 syllable is allowed to remain unchanged even when followed by another T3 syllable). Experiment 3 examined whether T3S/ unchanged T3 syllable are linked to different structures. The results showed that listeners associated T3S syllable with N-N compound structure (e.g., a combination of bamboo-shoot and cow) but unchanged T3 syllable with N-N coordination structure (e.g., single bamboo-shoot and cow), suggesting that the non-application of T3S facilitates the presence of a word boundary.

Hirose, Y., & Mazuka, R. (2015). Predictive processing of novel compounds: Evidence from Japanese. *Cognition*, 136, 350-358.

Hirose, Y., & Mazuka, R. (2017). Exploiting Pitch Accent Information in Compound Processing: A Comparison between Adults and 6-to 7-Year-Old Children. *Language Learning and Development*, 13(4), 375-394.

Table. 1 Auditory stimuli used in Experiment 1

a.	T3/ single	[*zhu2* 竹 *sun3* 筍]	"bamboo-shoot"
b.	T3/ compound (T3S is applied)	[*zhu2* 竹 *sun3* 筍]+ [*ru3* 乳 *niu2* 牛]→ 2-2-3-2	"bamboo-shoot cow"
c.	non-T3/single	[*xiang1* 香 *jiao1* 蕉]	"banana"
d.	non-T3/compound	[*xiang1* 香 *jiao1* 蕉] [*ru3* 乳 *niu2* 牛]	"banana cow"

Figure. 1 An Example of visual display used in Experiment 1.
Wan3 浣* is the pronounciation in Taiwan Mandarin.

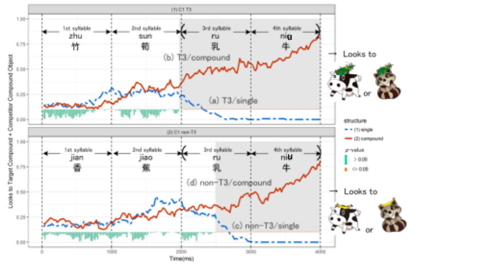


Figure. 2 Proportion of fixation for [Target Compound] + [Competitor Compound] objects in Experiment 1.

Table. 2 Auditory stimuli used in Experiment 2

a.	T3 + T3 (T3S is applied)	[*zhu2* 竹 *sun3* 筍] [*ru3* 乳 *niu2* 牛]→ 2-2-3-2	"bamboo-shoot cow"
b.	T3 + non-T3	[*zhu2* 竹 *sun3* 筍] [*hu2* 蝴 *die2* 蝶]	"bamboo-shoot butterfly"
c.	non-T3 + T3	[*xiang1* 香 *jiao1* 蕉] [*ru3* 乳 *niu2* 牛]	"banana cow"
d.	non-T3 + non-T3	[*xiang1* 香 *jiao1* 蕉] [*hu2* 蝴 *die2* 蝶]	"banana butterfly"

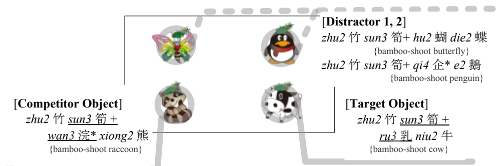


Figure. 3 An Example of visual display used in Experiment 2.
Qi 企* is the pronounciation in Taiwan Mandarin.

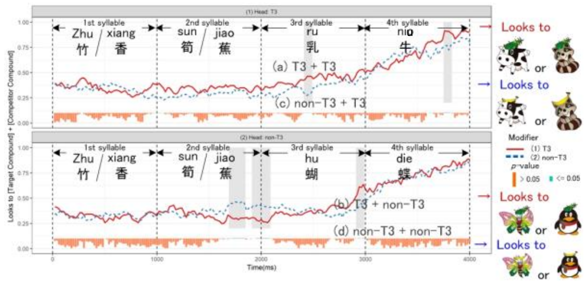


Figure. 4 Proportion of fixation on [Target Compound] + [Competitor Compound] objects in Experiment 2.

Table. 3 Auditory stimuli and paired visual displays used in Experiment 3

a. unchanged T3 + T3
Zhu2 sun3 ru3 niu2
(bamboo shoot cow,竹筍乳牛)

b. T3S + T3
Zhu2 sun2 ru3 niu2
(bamboo shoot cow,竹筍乳牛)

c. lexical T3 + non-T3
Zhu2 sun3 wu1 quei1
(bamboo shoot turtle,竹筍烏龜)

d. baseline non-T3 + T3
Qing1 jiao1 ru3 niu2
(green pepper cow,青椒乳牛)

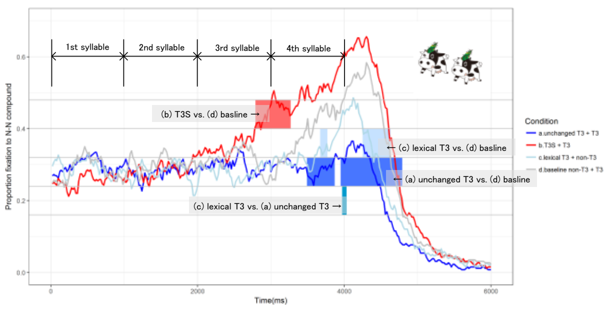


Figure. 5 Proportion of fixation on N-N compound interpretation in Experiment 3.

The Combinatorics of Merge and Workspace Right-Sizing

Sandiway Fong[1], Robert Berwick[2], Jason Ginsburg[*3]

[*]jginsbur@cc.osaka-kyoiku.ac.jp

[1]Department of Linguistics, University of Arizona, Tucson, USA

[2]LIDS MIT, Cambridge, USA

[3]Department of Educational Collaboration, Osaka Kyoiku University, Kashiwara, Japan

Contrary to e.g. Gazzaniga (2008), we believe the marvel we call the human brain is actually the weak link in our cognitive apparatus. There is ample evidence for this in the biological domain. Our sensory apparatus far outstrips the brain's capacity to process high-resolution input. Human eyes are capable of both extreme sensitivity, i.e. single photon level (Tinsley et al. 2016), and peak acuity of 77 cycles/degree (Curcio et al. 1990), all unnecessary for scene analysis. Olfactory thresholds can be of the order of parts per billion (ppb) (Wackermannová et al., 2016). Human eardrums can detect vibrations smaller than the diameter of a hydrogen atom (Fletcher & Munson, 1933). In case after case, the brain does not make use of the full resolution of our sensory inputs. We believe, for language, that the brain also (must) economize where it can: Chomsky (2005) terms the pressure for computational efficiency a Third Factor consideration. This applies to the reduction of Merge to simplest (binary) Merge (cf. Komachi et al., 2019). By considering a novel assumption about term accessibility with respect to Minimal Search, we show simplest Merge may follow directly from Workspace (WS) sizing constraints. Although simplest Merge (by itself) has demonstrably undesirable combinatorics (e.g. from an initial WS of just two lexical items, about 8 million distinct sets can be formed in just 8 Merges), language does not make full use of this resolution. As with our sensory apparatus, the brain cannot process all Merge possibilities, many of which are cases of iterated vacuous movement, unattested in data. We demonstrate that the same WS constraints that give us simplest Merge also rule out many such cases of infinite looping. This is a significant result because it means no dedicated "wetware" need be evolved to block potentially infinite loops.

Chomsky, N.A. (2005). Three Factors in Language Design. *Linguistic Inquiry*, 36, 1–22.

Curcio, C. A., Sloan, K. R., Kalina, R. E., & Hendrickson, A. E. (1990). Human Photoreceptor Topography. *The Journal of Comparative Neurology*, 292, 497–523.

Fletcher, H. and Munson, W. A. (1933). Loudness, its definition, measurement and calculation. *The Journal of the Acoustical Society of America*, 5, 82–108.

Gazzaniga, M. (2008). *Human: The Science Behind What Makes Us Unique*. New York, Harper Collins.

Komachi, M., Kitahara, H., Uchibori, A., & Takita, K. (2019). Generative procedure revisited. *Reports of the Keio Institute of Cultural and Linguistic Studies,* 50, 269–283.

Tinsley, J. N., Molodstov, M. U., Prevedel, R., Wartmann, D., Espigulé-Pons, J., Lauwers, M., & Vaziri, A. (2016). Direct detection of a single photon by humans. *Nature Communications,* 7, 12172.

Wackermannová, M., Pinc, L., & Jebavý, L. (2016). Olfactory sensitivity in mammalian species. *Physiological Research*, 65, 369-390.

The Reason of Building Hierarchical Structure: From the View of Recursive Combination as an Internal Operation

Genta Toya[*1], Rie Asano[2], and Takashi Hashimoto[1]

[*]Corresponding Author: toyagenta@gmail.com

[1]School of Knowledge Science, Japan Advanced Institute of Science and Technology, Nomi, Japan

[2]Department of Systematic Musicology, University of Cologne, Cologne, Germany

There are two proposals for explaining the emergence of hierarchical structures underlying human language, music, arithmetic, and action sequences. On the one hand, it is proposed that the cost of combining element give rise to the hierarchical structures (Mengitsu et al., 2016). On the other hand, it is suggested that the hierarchical structures are produced by recursive application of operation combining two items (Everaert et al., 2015), i.e., recursive combination (RC). The former claim was confirmed by computer simulation. Though the latter claim was also confirmed by computer simulation (Toya & Hashimoto, 2018), it is the simulation of physical manipulations related to the combination of elements. So as to prove the plausibility of the latter claim, it is necessary to confirm that RC evolves as an internal operation of an individual. In the current study, we aimed to clarify the adaptive significance and evolutionary conditions of RC in brain processing.

We designed a model with a system of statistical action generation using reinforcement learning and a system of constructive action generation using combining action representations. We hypothesize that the function of the current model corresponds to that of the cortico-basal ganglia-thalamocortical (CBGT) circuits. In order to search the conditions under which RC evolves in the system of constructive action generation, we conducted the evolutionary simulation of agent that explores and learns the physical rewards in making tools. As a result, RC of action representation evolved only when a reward function for learning changes according to tools made by oneself. RC of action possibly solves the problem "exploration or exploitation," which is taken as the task of reinforcement learning. RC which generates hierarchical structures evolves without the combining cost. The CBGT circuit which realizes RC possibly evolved as a solution to explore novel resource.

Everaert, M. B. H., Huybregts, M. A. C., Chomsky, N., Berwick, R. C., & Bolhuis, J. J. (2015). Structures, not string: Linguistics as part of the cognitive sciences. *Trends. Cogn. Sci.* 19, 729–743. doi: 10.1016/j.tics.2015.09.008

Mengistu, H., Huizinga, J., Mouret, J-B, & Clune, J. (2016). The Evolutionary Origins of Hierarchy", *PLoS Comput. Biol.* 12, e1004829. doi:10.1371/journal.pcbi.1004829

Toya, G., & Hashimoto, T. (2018). Recursive Combination Has Adaptability in Diversifiability of Production and Material Culture. *Front. Psychol.* 9, 1–17. doi: 10.3389/fpsyg.2018.01512

Coupling for Life:

Co-evolution of grammar and prosody in dialogic interaction

John W. Du Bois [*1]

[*]Corresponding Author: dubois@ucsb.edu

[1]Department of Linguistics, University of California, Santa Barbara, USA

Every living organism faces ultimately the same problem: how to avoid decay into thermodynamic equilibrium (death). According to Schrodinger (1943), the organism solves this by "continually sucking orderliness from its environment". Evolution has discovered many different ways of life, each defined by a strategy for coupling with the environment to extract order and information from it. Yet the human way of life is unique in the variety, novelty, and rapidity of its coupling strategies, as reflected in the diversity of languages, cultures, and practices. Cultural accumulation has given us unprecedented power to shape our own adaptive niche, as we collaborate to solve the novel coupling problems posed by each new environment. To learn the "secret of our success" (Henrich, 2016), we must ask how human hypersociality (Tomasello, Carpenter, Call, Behne & Moll, 2005) motivates us to couple our attention, emotions, and thoughts with others, as we enact the "extended mind" (Clark, 2008) that makes cultural accumulation possible. But an extended mind must have efficient ways of distributing its cognitive work across collaborating individuals (Hutchins & Johnson, 2009; Kolodny & Edelman, 2018). How does language help humans create new ways of coupling for life?

In this talk I argue that what makes the human way of life possible is our unique capacity for rapid, reversible "soft" coupling at two levels: (1) Conversational partners couple their words, thoughts, emotions, and actions in their sociocultural environment; (2) this then mediates their soft coupling to the external environment. Based on evidence from a corpus of naturally occurring conversation (Du Bois, Chafe, Meyer, Thompson, Englebretson & Martey, 2000-2005), I document the strategies interlocutors use to construct resonance, defined as the "activation of affinities across utterances" (Du Bois, 2014). Prosody provides critical infrastructure for building the structure of resonance. Specifically, the intonation unit serves as a prelinguistically available "prosodic niche", providing the emotional motivation for joint attention that is a precondition for dialogic coupling. The prosodic niche provides a predictable locus for unpredictable work, creating a workspace for cognitive integration within and across the extended mind. In the long term, behavioral strategies for coupling in the dialogic moment are learned and retained for reuse, yielding enduring consequences for the emergence and evolution of linguistic structure. Evolved grammars in turn support greater complexity of collaborative interaction, allowing new plasticity and creativity, all mediated by coupling through resonance in the prosodic niche.

Clark, Andy (2008). *Supersizing the mind: Embodiment, action, and cognitive extension*. Oxford: Oxford University Press.

Du Bois, John W. (2014). Towards a dialogic syntax. *Cognitive Linguistics* 25(3): 359–410.

Du Bois, John W., Chafe, Wallace L., Meyer, Charles, Thompson, Sandra A., Englebretson, Robert & Martey, Nii (2000-2005). *Santa Barbara corpus of spoken American English, Parts 1-4*. Philadelphia: Linguistic Data Consortium.

Henrich, Joseph (2016). *The secret of our success: How culture is driving human evolution, domesticating our species, and making us smarter*. Princeton: Princeton University Press.

Hutchins, Edwin & Johnson, Christine M. (2009). Modeling the emergence of language as an embodied collective cognitive activity. *Topics in Cognitive Science* 1(3): 523–546.

Kolodny, Oren & Edelman, Shimon (2018). The evolution of the capacity for language: The ecological context and adaptive value of a process of cognitive hijacking. *Philosophical Transactions of the Royal Society B* 373(1743): 20170052.

Schrödinger, Erwin (1943). *What is life?* Cambridge: Cambridge University Press.

Tomasello, Michael, Carpenter, Malinda, Call, Josep, Behne, Tanya & Moll, Henrike (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences* 28: 675–735.

Vocal Combinations in Birds: Implications for Language Evolution

Sabrina Engesser[*1]

[*]Corresponding Author: sabrina.engesser@outlook.com
[1] Department of Comparative Linguistics, University of Zurich, Zurich, Switzerland

Language's generative capacity is one of its key characterising features. A finite set of meaningless sounds can be combined to create meaningful words (phonology / combinatoriality), which can then be assembled into higher-order structures with derived meaning (syntax / compositionality). Comparative work on non-human animals investigating the evolutionary origin of such combinatorial abilities has mainly focused on the vocal repertoires of singing species or primates. Whilst these studies have demonstrated rudimentary combinatorial capacities outside of humans, evidence for basic phoneme-like or semantically compositional structures in non-human communication systems is rare. In this talk, I will present data on combinatorial structures in the discrete vocal repertoire of two cooperatively breeding birds: chestnut-crowned babblers (*Pomatostomus ruficeps*) and southern pied babblers (*Turdoides bicolor*). Furthermore, I will discuss the value of the comparative approach as a means of investigating similarities and differences among non-human and human communication systems. Ultimately, comparative data, particularly from species distantly related to humans, can help to unveil both the selective drivers promoting combinatorial capacities and potential precursors of language's combinatorial layers.

Referentiality and compositionality in bird calls

Toshitaka Suzuki[*1]

[*]Corresponding Author: toshi.n.suzuki@gmail.com
[1]The Hakubi Center for Advanced Research, Kyoto University, Kyoto, Japan

In human speech, words often convey independent meanings and syntax allows combining multiple words into more complex, compositional expressions. In contrast, animal communication signals have typically been considered as motivational: vocalizations merely reflect emotion or arousal of signalers and do not provide compositional messages even when multiple units are combined. However, recent field studies have challenged this assumption by showing that several species of birds and nonhuman primates may be able not only to assign independent meanings to acoustically discrete vocalizations, but also to combine these signals into higher lexical sequences. In this talk, I introduce my recent studies on vocal communication in a small bird species, the Japanese tit (*Parus minor*). Japanese tits produce acoustically discrete calls in a variety of contexts, such as when encountering a predator and when facilitating group cohesion. Field experiments have revealed that these calls may not merely reflect arousal of signalers, but also convey information about external referents, such as the presence of a particular type of predator. In addition, these birds are able to combine meaningful calls into higher structured sequences according to an ordering rule. Playback experiments revealed that receiver tits are able to use an ordering rule to extract compound meanings from call sequences even if these sequences are composed of novel combinations of calls. These findings demonstrate interesting parallels between bird calls and human language, opening new avenues for exploring the origins and evolution of linguistic capabilities.

"Syntax" in animal vocalizations: limitations and perspectives

Kazuo Okanoya*

*Corresponding Author: cokanoya@mail.ecc.u-tokyo.ac.jp
[1]Department of Life Sciences, The University of Tokyo, Japan

A part of motivation to study "syntax" in animal vocalizations is to detect a key to understand emergence of human language. The use of the term "syntax" to refer to a sequential characteristic in animal vocalization became popular when Peter Marler (1977) gave a distinction between phonological syntax and lexical syntax. Former refers to combinations of vocal items that changes sound patterns, but that do not change meanings. Later refers to combinations of sound items that also change referential meanings. Hurford (2011) suggested combinatorial syntax to refer to the former and compositional syntax to refer to later. The search for combinatorial syntax has been more directed towards the study of animal songs as songs have combinatorial richness, while that for combinatorial syntax is targeted in animal calls as calls are context specific. General design in such an experiment is to test for production and perception.

In combinatorial syntax, search for structural complexity based on Chomsky hierarchy have been a popular procedure. On perceptual side, the use of an artificial grammar to test the degree of structural complexity animals can learn is tested. To test production complexity, cautions must be paid not to over interpret occasional occurrence of context free grammar. To test perceptual ability, stimuli must carefully be designed (ten Cate, 2017). To show animal vocalizations bear compositional generativity, the threshold must be set even higher. However, Engesser and Townsend (2019) suggest more finer distinction of combinatoriality that could also change holistic meaning of vocalizations.

My suggestion for combinatory syntax study is to assume animal vocalizations are at finite-state level and find proximate and ultimate causes that produced the sequential system. My suggestion for compositional syntax study is to test animal cognitive ability in the lab and observe natural use of such ability in the field. Both requires long way to go, but I believe these are one of few procedures that leads us to the emergence of language.

Engesser, S., & Townsend, S. W. (2019). Combinatoriality in the vocal systems of nonhuman animals. Wiley Interdisciplinary Reviews: Cognitive Science, e1493.

Hurford, J. R. (2011). The origins of grammar: Language in the light of evolution II. OUP Oxford.

Marler, P. (1977). The structure of animal communication sounds. Recognition of complex acoustic signals: report of Dahlem workshop.

ten Cate, C. (2017). Assessing the uniqueness of language: Animal grammatical abilities take center stage. Psychonomic bulletin & review, 24(1), 91-96.

Online Presentation

We Were Not Alone: The Gradual Evolution of the Language Capacity

Dieter Hillert*[1,2]

*Corresponding Author: dhillert@ucsd.com
[1] ACTRI, UC San Diego, La Jolla, USA
[2] SHSL, San Diego State University, San Diego, USA

Modern humans are unique in the world of extant species as they use a finite set of rules to create and express an infinite amount of new meanings. Here, we discuss how this neurally implemented language capacity might have emerged in the hominin lineage. Genetic (copy-code mechanism of SRGAP2), neuroanatomical (cranial capacity and connectivity), and cognitive-behavioral data (development of tool and art production) show that the emergence of the human language capacity does not coincide with the beginning of behavioral modernity (BM) ca. 100-50 ka, but much earlier (see Figure 1). These findings support the view that the brain anatomy of modern humans and of our known sister species Neanderthals and Denisovans (*Homo sapiens sensu lato*) were language-ready (Dediu & Levinson, 2013; Hillert, 2015). The data show, moreover, that the roots of the modern language capacity can be traced back to a precursor language stage in our immediate ancestor Homo erectus. This account is consistent with Bickerton's (1990) protolinguistic analysis of pidgin and creole languages as well as with Jackendoff's (1999) linear grammar approach based on the analysis of less common languages and linguistic phenomena in acquisition and disorders. In considering fossil records along with evidence of symbolic-cognitive behavior, we conclude here that Homo erectus' language capacity was presumably premodern. The evolution of a language capacity supporting morpho-syntactic structures of a fully-fledged modern language may be therefore triggered by a co-evolving exponential accumulation of cultural behavior and cognitive workspace expansion in our species. In contrast to recent claims that the human language capacity suddenly emerged with the appearance of BM in our species (Berwick & Chomsky, 2016), we find overwhelming evidence that *we were not alone* and that this capacity evolved gradually, reasonable during the late Homo erectus epoch 1 ma or even earlier.

Berwick, R.C. & Chomsky, N. (2016). *Why Only Us: Language and Evolution.* Cambridge, MA: MIT Press.

Bickerton, D. (1990). *Language and Species.* Chicago: University of Chicago Press.

Dediu, D. & Levinson, S. (2013). On the antiquity of language: the reinterpretation of Neandertal linguistic capacities and its consequences. *Frontiers in Psychology 4* (397). doi: 10.3389/fpsyg.2013.00397

Hillert, D. G. (2015). On the Evolving Biology of Language. *Frontiers in Psychology 6* (1796). doi: 10.3389/fpsyg.2015.01796.

Jackendoff, R. (1999). Possible stages in the evolution of language. *Trends in Cognitive Sciences 3*, 272-279.

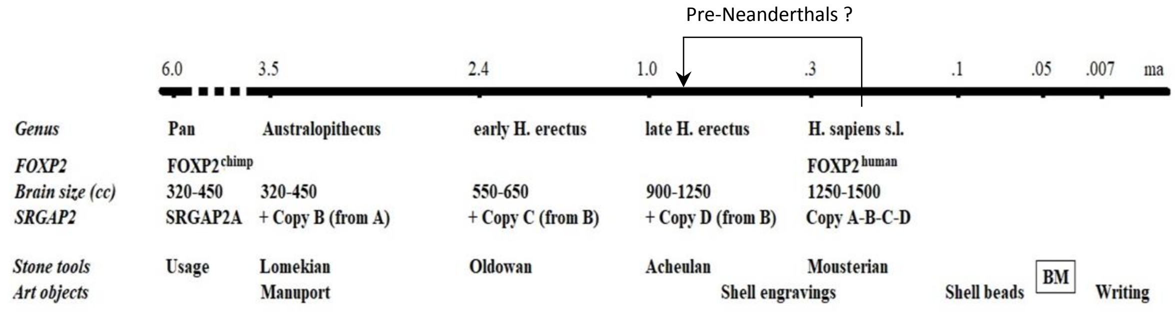


Figure 1. Timeline of genetic, neuroanatomical, cognitive-behavioral conditions indicating or contributing to the gradual evolution of the human language capacity (BM = behavioral modernity).

Logical Cells: Combinatory Categorial Grammar as a unified framework for describing melodic improvisation and harmonic analysis

Sebastian Klaßmann[*1]

[*]Corresponding Author: sebastian.klassmann@gmail.com
[1]Department of Systematic Musicology, Institute of Musicology, Faculty of Arts and Humanities, University of Cologne, Cologne, Germany

Philip Johnson-Laird (2002) suggests that tonal, melodic improvisation in jazz music must be directly linked to a harmonic consensus of interacting musicians that is in turn based on a representation of the underlying compositional constituents of the piece being performed. Modelling harmonic structures by means of using formal methods from computational linguistics is nothing new. However, when specifically looking at jazz from this perspective, melodic improvisation is a major constituant of performance and yet remains mostly unreflected upon – there appears to be no real methodological interface between it and syntactic models of harmonic progressions. Larson (2002) receives linear improvisation as being based on expectation, resolution and surprise which are in turn dependent upon a contextual function of a given melodic unit. From a harmonic perspective, a *combinatory categorial grammar* (CCG, cf. Steedman, 1996; Granroth-Wilding, 2013; Granroth-Wilding & Steedman, 2014) formally integrates this notion. This presentation focusses on early ideas regarding potential structural interfaces between harmonic consensus and melodic extrapolation which form the basis of my PhD research. This approach seeks to incorporate the notions that improvised melodies in jazz mostly make *sense* on a cellular level – i.e. in terms of local combinations of short note sequences as proposed by Frieler et al. (2016) – and that by recombining archetypical melodic cells, convincing lines in the jazz idiom can be generated (cf. Ligon, 1996; Vincent, 2015). The main goal of this presentation is to discuss the possibility of understanding these melodic cells as belonging to complex syntactic categories which are inherited from and thus directly interface with a structural harmonic representation that can be described by means of the CCG formalism. These concepts will be discussed based on my own analysis and annotation of melodic cells offered by Vincent (2015), as well as melodic archetypes taken from Bert Ligon's (1996) work.

Frieler, K., Pfleiderer, M., Zaddach, W. G., & Abeßer, J. (2016). Midlevel analysis of monophonic jazz solos: A new approach to the study of improvisation. *Musicae Scientiae*, *20*(2), 143-162.

Granroth-Wilding, M. (2013). *Harmonic analysis of music using combinatory categorial grammar*. Ph.D. dissertation, University of Edinburgh.

Granroth-Wilding, M., & Steedman, M. (2014). A robust parser-interpreter for jazz chord sequences. *Journal of New Music Research*, *43*(4), 355-374.

Johnson-Laird, P. N. (2002). How jazz musicians improvise. *Music Perception: An Interdisciplinary Journal*, *19*(3), 415-442.

Larson, S. (2002). Musical forces, melodic expectation, and jazz melody. *Music Perception: An Interdisciplinary Journal*, *19*(3), 351-385.

Ligon, B. (1996). *Connecting chords with linear harmony*. Hal Leonard Corporation.

Steedman, M. (1996). The blues and the abstract truth: Music and mental models. In A. Garnham, & J. Oakhill (Eds.) *Mental models in cognitive science: Essays in honour of Phil Johnson-Laird*, 305-318.

Vincent, R. (2015). *Jazz Guitar Soloing - The Cellular Approach*. Sher Music Corporation.