28th ALTENBERG WORKSHOP IN THEORETICAL BIOLOGY

Origins of Complex Communication and Language: Epigenetic Modeling and Ethological Observation

organized by
D. Kimbrough Oller, Ulrike Griebel, and Rick Dale

July 5-8, 2012

Konrad Lorenz Institute
for Evolution and Cognition Research
Altenberg, Austria

Welcome

to the 28th Altenberg Workshop in Theoretical Biology. The Altenberg Workshops are interdisciplinary meetings organized by the Konrad Lorenz Institute for Evolution and Cognition Research in Altenberg, Austria. The workshop themes are selected for their potential impact on the advancement of biological theory, and leading experts in their fields are asked to invite a group of internationally recognized scientists for three days of open discussion in the relaxed atmosphere of the Lorenz mansion. By this procedure the KLI intends to generate new conceptual advances and research initiatives in the biosciences. We are delighted that you are able to participate in this workshop, and we wish you a productive and enjoyable stay.

Gerd B. Müller
Chairman

The topic

Background: Modeling of communication/language evolution and learning has shown an extraordinarily accelerating publication rate in the last two decades. The work includes both the application of existing statistical methods (Niyogi 2006) and the development of new tools, including options utilizing neural networks of varying architecture (Christiansen & Chater 2008) and implementations in epigenetic robotics (Mugan & Kuipers 2007; Oudeyer & Kaplan 2006; Pezzullo & Calvi 2006). There have been some remarkable successes in this process, with new "existence proofs" illustrating, for example, that associative learning and interactive patterns among simple agents can account for the acquisition and/or evolution of a variety of structures that appear to be related to language (Elman et al. 1996). Yet it remains uncertain how much such modeling has to say about the real processes of language evolution and development. A key factor here concerns the particular "units" of communication to posit as inputs to models, and how to interpret outputs as units of the system being modeled. This is a fundamentally theoretical problem of external validation. One cannot know what models illustrate about the real processes of development or evolution if the models are not grounded in external evidence from the empirical study of language change. Of course the units of communication that were relevant at each stage in the *evolution of language* can only be inferred. On the other hand, the units of communication in the *development of language* in individual infants and children can be observed in longitudinal research. Thus at least in the short run the success of modeling of language change can best be judged by its success in simulating the observed steps of development in real human infants and children.

The problem of "units" in development is not at all trivial for modeling because the units of communication themselves change dramatically in both *forms* (the transmission units) and *functions* (the informational values that can be transmitted) in the process of early development. Consider first the *forms* of natural language: Mature units such as well-formed (or "canonical") syllables, phonemes, or phrases are not produced systematically in the early months of life (Oller 1980;

Stark 1980). Yet infants do “communicate” with precanonical vocalizations including both cries and laughter as well as with a notable set of infant precursor *forms* that appear to have infrastructural significance for speech (Oller 2000). These have been termed “squeals,” “growls,” “vowel-like sounds,” “raspberries,” and so on. Infants engage in vocal turn taking (Papoušek & Papoušek 1989), vocal play, and in expressive acts such as complaint, exultation, and refusal using these vocalizations (Oller 2000). Importantly, these human infant vocalizations of the first months of life can be shown to differ in form, variety, and flexibility of use from vocal *forms* of other primates at any age. So realistic modeling of early vocal development cannot begin with alphabetical units such as phonemes as the units of transmission, because real infants simply do not possess them. Yet most modeling that has been conducted on language evolution or development to date has utilized phonemic level units such as vowels or syllables (Westermann 2008). Clearly it will be necessary to model development such that *forms* (units of transmission) change from stage to stage, and where in the earliest stages *precanonical forms* become the focus.

Consider also the *functions* of natural language: The earliest vocal communications are clearly not symbolic, and thus they do not make explicit reference to objects or events (Bates et al. 1979). Instead, early vocal communications systematically express states (distress, comfort, elation) and can solicit attention or assistance, but the functions that are served by the vocalizations are at best “illocutionary” (Austin 1962) rather than “semantic.” Again however, the range of vocal expressions that is possible in the human infant and the flexibility with which those communications can be presented clearly exceeds that of non-human primates at any age (Griebel & Oller 2008). Thus realistic modeling of the earliest stages of vocal communication cannot begin with semantic (that is, referential) lexical units as the *functions* that are transmitted. Yet modeling that has been conducted on language evolution or development to date has focused precisely on semantic/lexical units (Steels 2004). Clearly it will be necessary to model development such that *functions* of communication change from stage to stage, and where in the earliest stages *illocutionary* (rather than semantic) *functions* become the focus.

The requirement that realistic modeling of language development must ultimately reveal stages wherein the very units of communication change with time would appear to be an inevitable implication of EvoDevo, epigenetic reasoning. In the EvoDevo perspective, the products of behavioral development are built upon each other, and upon the changing infant capability to interact with the world using those changing products. The earliest vocal units of communication can be analogized in this theoretical context to the earliest units of hand, arm or leg movement, where it is increasingly clear based on both research and accompanying epigenetic theory that infants begin with very little capability for motor coordination (Smith & Thelen 1993; Smith & Gasser 2005; Thelen 1995). The initial state involves the ability to make discoordinated, exploratory movements. The emergence of more coordinated capabilities appears to depend upon multimodal perception of these exploratory actions and upon computing temporal relations across the modalities across experience (Edelman 1987). Motor systems are seen as self-organizing in this way toward higher levels of coordination. Key questions about how to model these changes concern the precise nature of the initial motoric and sensory capabilities and the precise nature of the initial computational capability. Based on longitudinal research, vocal development appears to proceed in accord with similar exploratory patterns as in the case of limb movements, and it is sensible to propose that just as in other cases of motoric development, theories of epigenetic change are applicable to vocal development. Yet virtually no modeling research has addressed this possibility– instead research thus far has been hampered by the preformationist assumption that human communication development begins with relatively high level units, essentially mature language units that are phonemic on the side of *forms* and semantic/lexical on the side of *functions*.

Ethological observation and theory as the key grounding point for modeling. A central difficulty, then, for modelers of language evolution and development is associated with the fact that the appropriate initial units of analysis have not yet been addressed. These units (both *forms* and *functions*) must be determinable or verifiable by ethological observation. Language model-

ing presents a particularly complex problem (as opposed to modeling of communication systems in other primates) because language units are of extraordinary complexity in terms of both the large numbers of fundamental units and the indefinite size of repertoire after recombination of those fundamental units. Another reason, outlined above, is that the units that are relevant early in development or evolution and the ones that are relevant in later development or evolution are different. Just as a tree growing from a seed does not have the same structures at germination, at the point of a shoot emerging from the ground, at the point of first branching, at the point of first emergence of leaves, and so on, it can be said that language emerges with no less staging of different kinds of structures at different points in time. How to model units at different stages of development or evolution is, then, a critical problem. The fact that the endpoint categories are huge in number makes the problem geometrically more difficult than in cases of the small, fixed repertoires that are common in non-human primates.

In the case of human language, direct ethologically observable evidence of *evolution* is in the main only available for mature languages and only in the form of texts that can illustrate historical changes in languages. Thus the evidence provides only the tip of an iceberg. In *development* on the other hand, direct observation is possible, and consequently development presents a more tractable problem for the coordination of ethological observation and modeling than evolution does. It is reasonable to expect, however, that much of the pattern of change across human development will be broadly similar to the changes of unit types across hominin evolution, owing to the natural logic of how a communication system could possibly evolve or develop in the direction of language-like complexity and power (Oller 2000). Other nonhuman models are also particularly interesting, especially in cases where massive flexibility of signal production (and thus large repertoire sizes) can be seen, as with for example many birds and even some invertebrates (Griebel & Oller 2008).

Aims

This event follows two prior KLI workshops and corresponding volumes in the Vienna Series in Theoretical Biology: *Evolution of Communication Systems: A Comparative Approach* (2004) and *Evolution of Communicative Flexibility: Complexity, Creativity, and Adaptability in Human and Animal Communication* (2008). These workshops helped establish consensus on the ideas that (1) units of communication in human language and non-human communication systems can most effectively be compared at a level of abstraction, allowing focus on *forms* as non-phonemic vocal transmission units and on *functions* as illocutionary rather than semantic units; (2) flexibility of vocal communication in the human case is notably greater than in other primate species even from the first months of human life in both *forms* and *functions*, although considerable light can be shed on the roots of flexibility in vocal signal production (*forms*) from the study of other species, especially birds, marine mammals, and a few invertebrates; and (3) considerable progress has been made over the past twenty years in development of new methods of investigation for language development and evolution through applications of statistical modeling and robotics, approaches that lend themselves particularly to EvoDevo theoretical styles emphasizing epigenetic emergence of communicative system properties. An important implication of the consensus from the prior workshops is that individuals working in modeling and in ethological observation of real communication system development or evolution can profit from working hand in hand. In some cases this has already begun to happen (especially in epigenetic robotics), but the time is ripe for a basic assessment of the possibilities of coordinating modeling efforts with those of ethologically-based theory building. In particular we are ready for a coordinated effort from a variety of theorists and modelers to address the pressing need to implement models that reveal the changing units of communication both in *forms* and *functions* that appear to result from the natural processes of epigenetic emergence. The workshop will bring together a variety of key scholars working in the relevant domains of the evolution and development of communication systems.

The format

The workshop will be run in a seminar/discussion format. There are 16 presentations, with 45 minutes allotted for each—roughly 30 minutes for each talk, followed by 15 minutes for discussion and questions on that talk. This means that we will have seven presentations each on Friday and Saturday and two on Sunday morning. Each day the presentations will be followed by an extended general discussion session. On Sunday there will also be a discussion of publication plans. The three opening presentations on Friday morning will be by the organizers, Ulrike Griebel (University of Memphis, Institute for Intelligent Systems), D. Kimbrough Oller (University of Memphis, School of Communication Sciences and Disorders), and Rick Dale (University of California, Merced, Cognitive and Information Sciences). These papers will address both goals and framework of the workshop.

To support discussion during the sessions, we encourage each participants to send a rough draft of their presentation to the organizers in advance of the workshop, to be circulated among the participants. The targeted date for receipt of the drafts is Friday, 29 June.

Manuscript preparation and publication

The Altenberg Workshops in Theoretical Biology are fully sponsored by the KLI. The organizers plan to publish a volume in the *Vienna Series in Theoretical Biology* (The MIT Press). Each participant will supply a manuscript for the publication. The volume will further develop the novel ideas and concepts generated as a result of the meeting. The contributors are not necessarily limited to the original participants; they may be complemented by experts on those topics that emerge as important and may include co-authors invited at the discretion of the participants. This procedure is intended to generate new conceptual advances in the area of modeling for communication development and evolution, and because of the explicit interdisciplinary nature of the effort, the outcome should be attractive to a wide range of experts in the human sciences and neighboring disciplines.

We hope to circulate draft presentations before the workshop (see above). We expect that participants will revise their drafts as a result of discussion at the KLI and (assuming the participants agree to the review procedures that were utilized in our two prior volumes) the ensuing round-robin review process during which commentaries will be elicited for each paper from two selected members of the workshop. We are aiming for a December 31, 2012, date for receipt of finished manuscripts for publication. The length of the contributions should be approximately 8,000 words (30 or so double-spaced pages of text). The use of figures and photographs is highly encouraged. All contributions will be edited for style and content, and the figures, tables, and the like will be drafted in a common format. The editors will send specific instructions after the workshop.

D. Kimbrough Oller, Ulrike Griebel, and Rick Dale

Participants

BENJAMIN K. BERGEN **bkbergen@ucsd.edu**

Department of Cognitive Science
University of California, San Diego
9500 Gilman Drive, La Jolla, CA 92093-0515
USA

CYNTHIA BREAZEAL **cynthiab@media.mit.edu**

Personal Robots Group
MIT Media Lab
Massachusetts Institute of Technology
20 Ames St, E15-468
Cambridge, MA 02139
USA

MORTEN CHRISTIANSEN **mhc27@cornell.edu**

Department of Psychology
Cornell University
228 Uris Hall
Ithaca NY 14853-7601
USA
&
Santa Fe Institute
1299 Hyde Park Road
Santa Fe, NM 87501
USA

RICK DALE **rdale@ucmerced.edu**

Cognitive and Information Sciences
University of California, Merced
5200 North Lake Rd, Merced, CA 95343
USA

BART DE BOER **B.G.deBoer@uva.nl**

Vrije Universiteit Brussel
Dienst ARTI
Pleinlaan 2
1050 Brussel
Belgium

DAN DEDIU **Dan.Dediu@mpi.nl**

Max Planck Institute for Psycholinguistics
PO Box 310
6500 AH Nijmegen
The Netherlands

ULRIKE GRIEBEL **ulrikegriebel@gmail.com**

Institute for Intelligent Systems
The University of Memphis
FedEx Institute of Technology
365 Innovation Drive, Suite 303
Memphis, TN 38152
USA
&
Konrad Lorenz Institute for Evolution and Cognition Research
Adolf Lorenz Gasse 2
A-3422 Altenberg
Austria

CARLOS GUSSENHOVEN **C.Gussenhoven@let.ru.nl**

Department of Linguistics
Radboud University Nijmegen
PO Box 9103
NL-6500 HD Nijmegen
The Netherlands

VITTORIO LORETO **vittorio.loreto@roma1.infn.it**

Physics Department
Sapienza University of Rome
Piazzale A. Moro 2
00185 Roma
Italy

GARY LUPYAN **lupyan@wisc.edu**

Department of Psychology
University of Wisconsin, Madison
1202 West Johnson Street
Madison, WI 53706-1969
USA

BOB MCMURRAY **bob-mcmurray@uiowa.edu**

Department of Psychology
E11 SSH
University of Iowa
Iowa City, IA 52240
USA

D. KIMBROUGH OLLER **koller@memphis.edu**

School of Communication Sciences and Disorders
The University of Memphis
807 Jefferson Ave.
Memphis, TN 38105
USA
&
Konrad Lorenz Institute for Evolution and Cognition Research
Adolf Lorenz Gasse 2
A-3422 Altenberg
Austria

PIERRE-YVES OUDEYER **pierre-yves.oudeyer@inria.fr**

Inria
200, avenue de la Vieille Tour
33405 Talence
France

LINDA SMITH **smith4@indiana.edu**

Psychological and Brain Sciences
Indiana University
1101 E. 10th St.
Bloomington, IN 47405
USA

ANDREW WEDEL **wedel@u.arizona.edu**

Department of Linguistics
University of Arizona
1030 N. Olive Road
Tucson, Arizona 85721-0028
USA

GERT WESTERMANN **g.westermann@lancaster.ac.uk**

Fylde College
Lancaster University
Bailrigg
Lancaster
United Kingdom LA1 4YF

Origins of Complex Communication and Language: Epigenetic Modeling and Ethological Observation

Thursday 5 July	Evening	
6.00 pm		Welcome reception, introductions, and dinner at the KLI

Friday 6 July	Morning	Origins of Language, Comparative and Developmental Perspectives	Chair: Dale & Griebel
9.00 am – 9.45 am	Griebel & Oller	"My Dog Understands Every Word I Say": Cognitive Underpinnings of Language and Evidence from the Animal Kingdom	
9.45 am – 10.30 am	Oller & Griebel	Vocal Development as a Guide to Modeling the Evolution of Language	
10.30 am – 11.00 am	Coffee		
11.00 am – 11.45 am	Dale	Language Evolution by Multimodal Synergy	
11.45 am – 12:30 pm	Smith	Sensorimotor Origins of Reference	
12:30 pm – 2.00 pm	Lunch	at the KLI	

Friday 6 July	Afternoon	Modeling 1	Chair: Griebel
2.00 pm – 2.45 pm	Christiansen	The Importance of Chunking in Language Learning and Evolution	
2.45 pm – 3.30 pm	Westermann	Experience-dependent Brain Development as a Key to Understanding the Language Faculty	
3.30 pm – 4:00 pm	Coffee		
4.00 pm – 4.45 pm	Dediu	Language Acquisition, Change and Evolution: Genetic Influences and Phylogenetic Approaches	
4.45 pm – 5.30 pm		First General Discussion	
6.00 pm		Departure for Dinner and open evening for exploration of Vienna (or the environs)	

Saturday 7 July	Morning	Robotics	Chair: Oller
9.00 am – 9.45 am	Breazeal	Human-Robot Communication and Coordination through Non-verbal Behavior	
9.45 am – 10.30 am	Oudeyer	Bootstrapping Language Development out of Multimodal Curiosity and Socially Driven Development of Sensorimotor Skills in Robots	
10.30 am – 11.00 am	Coffee		

Saturday 7 July	Morning	A Perspective from Linguistics	Chair: Oller
11.00 am – 11.45 am	Gussenhoven	Functions of Intonation: Complex Language Structures versus Physiologically Based Paralinguistic Communication	
11.45 am – 12.30 pm	Wedel	Context-specific Pronunciation and the Creation of Regular Sound Change: Effects of Short versus Long-term Predictability on the Pronunciation of Individual Words	
12:30 pm – 2.00 pm	Lunch	at the KLI	

Saturday 7 July	Afternoon	Development and Cognitive Processes	Chair: Dale
2.00 pm – 2.45 pm	McMurray	EvoDevo in the Third Dimension: The Role of Real-time Processes in Language Development and Evolution	
2.45 pm – 3.30 pm	Bergen	How did Grammar Come to Modulate Embodied Simulation?	
3.30 pm – 4.00 pm	Coffee		
4.00 pm – 4.45 pm	Lupyan	Cognitive Functions of Language and their Implications for Language Evolution	
4.45 pm – 5.30 pm		Second General Discussion	
6.00 pm		Departure for Dinner at a Viennese Heurigen	

Sunday 8 July	Morning	Modeling 2	Chair: Griebel
9.00 am – 9.45 am	Loreto	A Cultural Route to the Emergence of Duality of Patterning	
9.45 am – 10.30 am	De Boer	Interaction of Biological and Cultural Evolution of Speech: The Case of Combinatorial Structure	
10.30 am – 11.00 am	Coffee		
11.00 am – 12.00 pm		Third General Discussion	
12.00 pm – 1.00 pm		Plans for Review and Publication	
1.00 pm – 2.00 pm	Lunch	at the KLI	
2.00 pm		Departure for a Boat Trip on the Danube with Dinner and Merry-Making in Dürnstein	
9.30 pm		Return to Vienna	

Abstracts

Ulrike GRIEBEL (with D. Kimbrough OLLER)
University of Memphis

"My Dog Understands Every Word I Say": Cognitive Underpinnings of Language and Evidence from the Animal Kingdom

The human species seems to be the only one on this planet that has evolved a highly complex communication system that can be called language. The reason or reasons for this are highly debated, as are the exact changes in features that actually caused human communication to diverge from primate communication systems during the evolution of modern language. Some argue that changes in a certain set of cognitive features was enough to make an "emergence" of language possible, but what exactly these features comprise is not well defined. Since our cognitive abilities have evolved from those of our ancestors, we suggest a comparative approach, examining animal language learning studies to determine what animals actually understand about human language. This approach might lead to better understanding of the cognitive underpinnings of language in general, and help delineate a general cognitive background for the evolution of a complex communication system such as language. But it might also help to pinpoint features that seem to be uniquely evolved in humans in adaptation to their complex communication needs.

In this paper we survey evidence from experimental studies with language learning animals evaluating their capabilities with regard to some mostly undisputed "ingredients" of language, which include but are not limited to: Symbolism, contextual and functional (illocutionary) flexibility, complex serial ordering, grammar (formal relationships between types of symbols), displacement in time and space, open-endedness of repertoire, creativity (creation of new symbols), recursion (embedding phrases within phrases of the same type in a hierarchical structure), and cultural transmission (teaching). We also discuss the puzzling discrepancies

between cognitive abilities of many non-humans to acquire certain basic language skills and their seemingly primitive natural communication systems.

So far most modeling in human language has been done with full-blown language including words and syntax etc. Our hope is that considering data from animal language learning studies about the most foundational properties of human language will at some point enable modelers to get to the root of language evolution.

D. Kimbrough OLLER (with Ulrike GRIEBEL and Anne S. WARLAUMONT)
University of Memphis

Vocal Development as a Guide to Modeling the Evolution of Language

Most modeling that we are aware of for both evolution and development of language has utilized mature units of spoken language as both targets and inputs. For example, “phonemes” or some other alphabetical units derivable from text have proven accessible and convenient in modeling. Similarly, lexical items, formed of both bound and free morphemes, are commonly targeted by modelers. But this approach is inherently incapable of addressing the earliest phases of language development (and presumably evolution) because the human infant is unable to produce any of these commonly targeted language features in the first 10 months of life.

Vocal development events occurring in the human infant prior to that age form essential infrastructure for all that vocal language comes to be. Here we do *not* refer to cry and laughter, which are species-specific sounds resembling the “calls” of non-human primates and other mammals (Owren et al. 2011), but to sounds that appear to be specific precursors to speech, the “proto-phones”. Thus modeling that aims to capture the essence of development and evolution will need to account for the infrastructure built in the first months of life by addressing (1) the protophones that emerge before mature speech categories are commanded, as well as (2) the nature of the prelinguistic exploratory and communicative utilization of the protophones prior to the existence of productive lexicon.

On the protophones: Among the most remarkable motoric events of the first year in humans is the appearance of canonical babbling at around 7 months and not later in typically developing infants than 10 months (Koopmans-van Beinum & van der Stelt 1986; Stark 1981), typically recognized when infants produce vocal/articulatory sequences that adult listeners often interpret as “baba” or “mama,” and that are taken by adult listeners to be candidates for lexical

formation. In these protophones, infants display the culmination of a series of vocal developments that begin in the first days of life and progress through at least three discernible prior stages (Oller 1980), in which even more primitive precursors to speech sounds are produced. Loosely the many vocal categories that have been recognized in longitudinal research as precanonical protophones, encompass sounds associated with control of normal phonation (quasivowels), vocal pitch and voice quality (squeals, vocants, growls), and precanonical articulation (gooing, marginal babbling, raspberries). It is sensible to postulate that all of these vocal development stages were replicated, at least loosely, by our hominin ancestors as they followed the evolutionary path that led to the production of the well-formed or canonical syllables required in language (Oller 2000).

On the utilization of vocalization in early infancy: Of similar importance for realistic modeling of language development and evolution is the recognition of ways the protophones are used. From the first month of life they appear to be produced spontaneously, with no obvious intent. They appear to be used by parents as signals of infant state, and they appear to constitute endogenously motivated exploration of the vocal capacity. By the third or fourth month the sounds of the infant are used in systematic face-to-face turn-taking exchanges with caregivers, proto-conversations (Trevarthen 1974). About the same time, the protophones have been diversified such that they fall into a small set of discernible though fuzzy categories, recognized by parents as such, and forming the basis for systematic interaction *about* the categories, as indicated for example by parental imitation of them and attempts to elicit them. By not later than the fourth month, it can be demonstrated that infants show functional flexibility of the protophones (Griebel & Oller 2008), the ability to produce each type with multiple valences ranging from positive to neutral to negative. All of these infrastructural properties of human vocalization in the first half year of life (spontaneous production, proto-conversation, category formation, and functional flexibility) as well as several others to be discussed in the paper, are either absent altogether or scarcely present in the vocal activities of our primate relatives.

Since non-human primates neither possess the ability to produce sounds outside their “call” repertoire (they possess no protophones) nor the ability to utilize the calls they do possess in the flexible ways human infants do with the protophones, we have proposed that our hominin ancestors must have evolved similar capabilities as infrastructure for the subsequent evolution of language. We have been developing this EvoDevo (Hall 1992; West-Eberhard 2003) style argument for many years, and now propose that modeling of the evolution and development of language should directly characterize the sequential emergence of the typical infant vocalization categories as well as their apparently human-specific patterns of utilization as critical infrastructure for language. This will represent a significant change in the nature of modeling of language evolution, since real infant sounds begin very distant from speech (though they are more like speech than sounds of any other primate at any age) and are transformed in a series of naturally logical steps, becoming at each stage more speech-like.

This paper will outline an agenda for modeling where the protophones along with their typical patterns of utilization are seen as the initial targets of acquisition/development, and where successively more speech-like sounds and utilization characteristics emerge across time. The proposal will also suggest social/perceptual influences that may help guide the emergence of protophones and their utilization patterns.

Rick DALE (with Christopher T. KELLO)
University of California, Merced

Language Evolution by Multimodal Synergy

We gesture. We smile. We speak and laugh. We exploit phonetic regularities while structuring whole narratives. In fact, we sometimes do all of this simultaneously, in matters of just seconds. We argue multimodality is not just a feature of language, but was a crucial ingredient in language evolution, and remains a relatively underexplored ingredient in models of language evolution. In many respects this is not a new proposal. For example, Peter Carruthers has referred to language as a kind of "integrative system" (Carruthers 1998, 2002). Jeff Elman has pointed to language emergence as a kind of "conspiracy theory" of many interacting constraints (Elman 1999; see also, among many others, MacWhinney, 1999; Christiansen in press; Christiansen et al. 2001; Louwerse et al. in press; Seidenberg & MacDonald 2001; van Rooij 2012). Schoenemann (1999, 2009) also argued for growing integrative brain systems as a crucial basis for complex symbols. For quite some time, the duality-patterning conception of language (from Hockett and beyond) has focused on language's systematically mapped layers of complexity.

This talk will consider the manner in which these layers of language served to constrain each other mutually, and through that synergy, bring about structure (both in the semantic, and syntactic, sense). To understand how multimodality can engender such structure, a number of global informational principles for human language will be introduced from graph theory, and we review existing proposals (e.g., Ferrer-i-Cancho et al. 2005).

Articulating these principles cannot speak directly to their generating mechanisms, which must be the subject of modeling. We discuss the implications of these principles in modeling, and how this synergy across levels can be pursued. We draw from a wide array of evidence that this is what may have taken place,

including: comparative brain structure studies, human behavioral experiments, development, corpus analysis and other computational models.

In the end, we argue that without a rich array of multiple modalities underlying language, conditions would not be sufficient for it to thrive ("robustness" in the sense of Winter & Christiansen 2012). Many proposed conditions of the pre-hominid line encourage tales of selection pressure for multimodality, such as complex social organization and memory.

The upshot of the discussion is that, quite in opposition to Hauser, Chomsky, and Fitch (2002), recursion is an epiphenomenon of this process. The degrees of freedom in human cognition and action offer our large and integrative multimodal nervous system the chance to weave them into systematic linguistic behavior under particular forms of ecological constraint.

Linda B. SMITH
Indiana University

Sensorimotor Origins of Reference

Theorists who study early word learning typically do so from the top-town, and in terms of theoretical constructs at the macro- or cognitive level: inference, intention, meaning, concepts, joint attention. This talk will consider early word learning from the bottom-up, in terms of the sensory-motor processes that support the visual isolation of objects, stabilized attention, and the binding of names to things. The data derives from a series of studies in which mothers and toddlers interact with toys and mothers sometimes names those toys. We measure at high temporal resolution the head, hand and eye movements of the two participants, the first-person views of the interaction, and the toddlers learning of any of the toy names after the play session. The bottom-up view of word learning both grounds and challenges traditional cognitive accounts. The unsolved problem of how to unify phenomena across levels of analysis—across macro and micro scales of behavior—will be considered.

Morten H. CHRISTIANSEN
Cornell University & Santa Fe Institute

The Importance of Chunking in Language Learning and Evolution

Language happens in the here-and-now. If the linguistic input is not processed immediately, nothing can be learned from it. To successfully deal with the continual deluge of linguistic information, the brain must compress and recode the input as rapidly as possible. As a consequence, incoming language incrementally gets recoded into chunks of increasing granularity, from sounds to constructions and beyond. Thus, units at different levels of linguistic analysis come for free as a consequence of the transient nature of language. The specific units vary cross-linguistically due to historical differences in the trajectories taken by cultural evolution for each language. And these units may also change during development as the child learns to use language. To illustrate, I discuss a recent chunk-based computational model of early syntactic acquisition. I conclude by arguing that the immediacy of language processing provides a fundamental constraint on accounts of language learning and evolution.

Gert WESTERMANN
Lancaster University

Experience-dependent Brain Development as a Key to Understanding the Language Faculty

An influential view of the nature of the language faculty is that a system of rules is combined with a lexicon that contains the words of the language together with, depending on the particular viewpoint taken, a more or less rich representation of their context. Alternative views, usually based on connectionist modeling, attempt to explain the structure of language on the basis of complex associative processes. The English past tense has emerged as a focus on which different language theories are explored, because here we find a model system for language as a whole: regular forms appear to be generated by a rule process and irregulars are memorized, thereby creating a system of rules and exceptions.

Here I extend work on the English past tense by considering the shortcomings of both the words-and-rules and the connectionist approaches. I argue that the adult language system cannot be understood properly unless we consider how it has arisen through development. The recently emerging field of cultural neurolinguistics acknowledges that a specific language environment shapes the brain structures responsible for processing this language in an experience dependent way. Taking on board the importance of experience dependent brain development I describe a 'neuroconstructivist' connectionist network model of past tense acquisition, adult processing and impaired processing after brain damage, and I show how it can account for the empirical data better that static, non-developmental models of either theoretical persuasion. This model puts experience dependent brain development in response to a specific language environment at the heart of understanding not only language development but adult language processing as well.

Dan DEDIU
Max Planck Institute for Psycholinguistics

Language Acquisition, Change and Evolution: Genetic Influences and Phylogenetic Approaches

The patterning of linguistic diversity and the universal tendencies shared by languages are potentially important input data for modelers. In turn, models will help us understand better the interplay between historical accidents and constraints (articulatory, cognitive, linguistic, cultural, etc.) in shaping both the diversity and shared property of languages. I will briefly introduce some examples of genetic effects on language, and of the inverse feedback from language to genes through participation in the construction of the complex cultural niche characterizing modern humans. I will also overview some recent methods and findings resulting from the application of modern phylogenetic methods to studying language history and their effects on our understanding of how language evolves and the timescales involved. I will conclude by discussing possible ways of integrating evolutionary (both genetic and cultural) phenomena into a broader understanding of language.

Cynthia BREAZEAL
Massachusetts Institute of Technology

Human-Robot Communication and Coordination through Non-verbal Behavior

As personal robots enter our workplaces and homes, it will be important for them to interact naturally and cooperate effectively with people. Social robots, in particular, are designed to exhibit and support anthropomorphic cues and encourage people to engage them as social agents. Hence, this new "breed" robot is designed to interact with people more as a partner rather than as a tool and opens new applications for socially intelligent machines in the future.

In this talk, I survey a body of research in the field of Social Robots that highlights the importance and impact of robots' non-verbal behaviors in a wide range of collaborative and communicative scenarios with people, from the formation of social judgments toward robots, to teamwork, and aspects of social learning.

I begin by briefly presenting a study to illustrate how social robots can successfully "tap into people's social brain." Specifically, I illustrate how the non-verbal gestures of a robot can influence people's social judgments of it, for instance, how trustworthy the robot is perceived to be to is interlocutor. In fact, the robot uses the same cues hypothesized to influence the judgment of trustworthiness between people, and the robot has been successfully used as a new kind of scientific instrument for social psychologists to validate those cues. The take away message is that people behave like people, even when interacting with a robot. Subconsciously, our brains intuit social robots as social actors.

Next, I turn to the design problem. How can we design robots that "think" as social actors? I present a high-level overview of the design and development of a socio-cognitive architecture for our robots. Specifically, our approach is guided by the hypothesis that imitative interactions between infant and caregiver, starting with facial mimicry, are a significant stepping-stone to developing appropriate

social behavior—to predicting others' actions and ultimately to understanding people as social beings. The architecture begins with the simplest mechanisms of imitation, contingency and attention to enable the robot to engage in an imitation-based social learning process with a human teacher. The design of our architecture is inspired by embodied theories of cognition to enable the robot to leverage its physical and cognitive embodiment through simulation-theoretic mechanisms to first learn how its body and internal states map onto those of its human interlocutor (the "Like Me" hypothesis). Hence, our approach is loosely inspired by theories for how human infants learn to communicate with caregivers and come to understand the actions and expressive behavior of others in intentional and motivational terms.

I shall describe how we can then build on this foundation, layering on more sophisticated capabilities, such as shared attention and perspective taking, to enable a robot to infer selected mental states of others by observing their behavior. This enables the robot to perform simple false belief tasks, and reason about beliefs in order to appropriately assist a person with intended goals (even if their enacted plan is invalid).

Building further, we show how people structure social behavior to mediate the interaction of attention with learning, in effect to serve as "social filters." Social filters can be social-cognitive capabilities such as perspective taking that focuses the robot's attention on the subset of the problem space that is important to the teacher. This constrained attention allows the robot to overcome ambiguity and incompleteness that can often be present in human demonstrations and thus learn what the teacher intends to teach. Other social filters can be external, dynamic, embodied cues through which the teacher uses his or her body to spatially structure the learning environment to direct the attention of the learner.

None of these interactions involve language. Clearly language is critical for how people communicate. However, this work highlights how important non-verbal signals and behaviors are in understanding one another and coordinating complex joint action. In the future, robots should be able to communicate,

cooperate, and learn in partnership with people through sophisticated forms of verbal and non-verbal expression. That remains a grand challenge of robotics and AI.

Pierre-Yves OUDEYER
Inria

Bootstrapping Language Development Out of Multimodal Curiosity and Socially Driven Development of Sensorimotor Skills in Robots

Robots are extraordinary experimental tools to explore and generate novel hypothesis for understanding human sensorimotor, cognitive and social development (Oudeyer 2010). In this presentation, I will present several robot experiments that address the question of how elementary linguistic skills, and how primitives within language, can be formed out of multimodal and general mechanisms for sensorimotor and social development, pushing further the limits of theories conceptualizing language as fundamentally grounded in action. Indeed, I will suggest that fundamental aspects of language could be understood as *formed out of* sensorimotor development, rather than simply *grounded in* sensorimotor development.

*Discovering social interaction out of curiosity-driven body babbling***.** First, I will explain how intrinsic motivation systems, also called curiosity-driven exploration, can explain in a unified manner autonomous exploration and learning of one's own body (e.g. hand-eye or vocal tract-ear relationships through body babbling) and its relation with both the physical world (e.g. hand-object) and the social world (e.g. sound/gaze-sound/gaze). In short, I will show that a robot can be lead to interact vocally, in imitation-like interactions, with "others" as a side effect of itself exploring and learning through curiosity what his own body can effect in its surrounding (Oudeyer and Kaplan 2006). I will also show that such a mechanism does not need to have a prior concept of "other," but rather that the distinction self/object/other can emerge out of such a developmental mechanism (Kaplan and Oudeyer 2007).

Bootstrapping linguistic primitives out of multimodal sub-symbolic sensorimotor flow. Can a child learn linguistic primitives such as "phonemes" and "words," as well as perceptual or motor primitives to which they are associated such as

"visual objects" or "primitive actions," without starting from innate mechanisms for segmenting the corresponding sub-symbolic percepts out of ambiguous high-dimensional multimodal sensorimotor flow? I will show that certain modern statistical learning approaches, based on what is called dictionary learning, can have such capabilities, thus making it non impossible in principle for children (ten Bosch et al. in preparation).

Bootstrapping meaning: from language Gavagai to motor Gavagai. In a third series of experiment, I will show that the problem of learning word meanings can be bootstrapped as a particular case of context-dependent learning of motor skills through imitation (Cederborg and Oudeyer 2011). In particular, I will argue that the fundamental characteristics of the so-called language Gavagai problem are already included in imitation learning of motor skills, and thus introduce the "motor Gavagai" problem.

Carlos GUSSENHOVEN
Radboud University Nijmegen

Functions of Intonation: Complex Language Structures versus Physiologically Based Paralinguistic Communication

Human vocal communication proceeds through two systems, animal communication ('paralinguistics') and language. Their simultaneous use is best seen in vocal fold vibration, and tone and intonation therefore provide a unique area of investigation. Paralinguistic form-meaning relations, which precede language, are metaphorical interpretations of the effects on vocal fold vibration of anatomical/ physiological conditions. Humans exploit these in partly community specific ways. Linguistic intonation has the usual linguistic structure. That is, it is discretized, linearized, hierarchical, metrical and has dual articulation (i.e. phonology and morphosyntax). Examples from different languages show that the discretization may range from modest to excessive and that the range of complexity is large. The paralinguistic form-meaning relations may well remain one force among others in shaping language change.

Andrew WEDEL
University of Arizona

Context-specific Pronunciation and the Creation of Regular Sound Change: Effects of Short versus Long-term Predictability on the Pronunciation of Individual Words

Variationist/evolutionary models of phonology propose the existence of a causal chain that links biases at the utterance level to the development and consolidation of abstract phonological patterns over time within a speech community. Some of the properties of linguistic cognition that have been proposed to underlie this chain are (i) storage of experienced detail at multiple levels of description, (ii) feedback between perception and production, (iii) a similarity bias in the production and perception of variation, and (iv) some mechanism for the enhancement of cues to potentially ambiguous lexical items in usage.

Some of the individual links in this hypothesized chain have been supported by experiment, but others remain relatively understudied. Here, I'll focus on the theoretical prediction of the model that the pronunciation of a given lexical item must be influenced by three correlated factors: the phonetic and sentential predictabilities of the item in the local context, these predictabilities across previously experienced contexts, and the average contextual predictabilities of phonologically similar lexical items. Under this model, this nested set of influences creates a pathway that allows pronunciation variants that are frequently produced in individual utterances to propagate through the lexicon, creating regular sound change.

Because these factors are correlated with one another, it is not straightforward to establish their separate effects. In this paper I'll review results to date that are indirectly consistent with this pathway from computational simulation, laboratory experiment, corpus studies and phonological typology. Finally, I'll discuss possible experimental strategies to detect the separate influence of these effects in natural speech.

Bob MCMURRAY
University of Iowa

EvoDevo in the Third Dimension: The Role of Real-time Processes in Language Development and Evolution

Evolutionary developmental systems theory stresses two timescales—the development of an organism, and the evolution of a species. In this reframing of the modern synthesis, these timescales critically interact, as selection pressures operate on entire developmental systems (Lickliter & Honeycutt 2003) rather than genes (Dawkins 1976). Taking this approach to language points to the profound importance of understanding the mechanisms of development, both of language, and of other cognitive capacities more broadly, and it points to multiple avenues of heredity and evolution, including the environment and the language (Christiansen & Chater 2008). However, in putting the primacy on developmental mechanisms, this approach often neglects the real-time processes and interactions. In this talk, I will examine the role of this third timescale with two case-studies.

First, I examine how infants learn the phonetic categories of their language, and specifically how this is affected by infant directed speech (IDS). This is a classic EvoDevo story: infants learn speech categories (in part) via statistical learning mechanisms that are sensitive to the distributions of speech cues (Maye et al. 2003, 2008; McMurray et al. 2009), and at the same time, caregivers' use of IDS may selectively enhance these statistics (Kuhl et al. 1997; Liu et al. 2003). Thus, it appears that a specific cultural practice evolved to support development. I will present a much more detailed phonetic analysis of the acoustics of IDS that suggests that the differences between IDS and ADS (adult-directed speech) are not selectively beneficial for statistical learning, and may sometimes impede. These changes more likely derive from broader changes in speaking rate, prosody and affect, changes which are not "intended" for development, but rather are responses to the hear-and-now demands of getting an infant's attention, managing their arousal levels, and engaging them in language more broadly

(Smith & Trainor 2008). Here, a strictly developmental focus leads to a satisfying story, but may miss the underlying dynamics of the developmental system, and hence get both the development and the evolution wrong. Crucially it points to the role of real-time demands (that may be independent of developmental outcomes) in shaping development, and, in an EvoDevo framework, evolution.

The second case study examines the importance of real-time processes in characterizing development. I first show with a series of computational models of speech categorization (McMurray et al. 2009, 2011; Toscano & McMurray 2010) that statistical learning processes alone are insufficient to account for adult speech perception performance, and that they must be buttressed with real-time perceptual processes like competition and feedback from other levels of the system, moreover, it turns out that without such processes, unsupervised learning of the sort posited by statistical learning may not be possible. I next describe a computational model of word learning and referent selection that is built on the explicit interaction of real-time dynamic competition between interpretations of a word/scene, and longer-term associative learning that gradually build links between words and possible referents (McMurray et al. 2009 and in press). Crucially this model reframes many of the classic findings in word learning, suggesting that the apparent rapidity of early word learning may derive from slow gradual learning that is buttressed with in-the-moment processes that allow children to make inferences on the basis of only partial knowledge. Indeed, this flips the notion of performance/competence on its head, suggesting that in many cases, these real-time mechanisms allow children to perform in the moment much better than their underlying competence or knowledge. Again, this suggests that to understand the evolution of language skills like word learning, we must understand that pressures that operate at the level of real-time processes, and the way that the interact with slower learning processes over development.

These case studies lead to broader questions about feedback and selection in the system: at what time-scale do selection pressures for real-time behavior, development and evolution operate. In our theories, should pressure for good developmental outcomes trump the pressures faced by children and their care-

givers in the here and now? And if not, how do we reconcile this with the commitment that evolutionary selection pressures work at the level of developmental systems. Similarly, if specific real-time processes like competition are required for development and a fundamental part of the developmental system, how do we account for their development and evolution?

Benjamin K. BERGEN
University of California at San Diego

How did Grammar Come to Modulate Embodied Simulation?

Over the past decade, dozens of behavioral and neuroimaging studies have shown that people activate their perceptual and motor systems while processing language about perceptible scenes or motor actions. This has been interpreted as suggesting that language comprehension involves performing embodied simulations of described events. Moreover, the grammar of human languages serves to modulate this simulation in a variety of ways. For instance, grammatical person ("you" versus "he") modulates the perspective a simulation is constructed from; grammatical aspect ("he is opening the drawer" versus "he has opened the drawer") modulates whether the simulation focuses on the middle or the end of the event. In this talk, I will outline some of the pressures and mechanisms that might have given rise to this capacity, which appears to be uniquely human.

Gary LUPYAN
University of Wisconsin

Cognitive Functions of Language and their Implications for Language Evolution

Language is a defining trait of our species. A standard assumption shared by many in the cognitive sciences is that language simply allows for public expression of ideas that are themselves represented in a language-independent 'mentalese.' Similarly, capacities on which humans appear to differ markedly from other animals—relational reasoning, theory of mind, categorization, and executive function—are often viewed as developments largely unrelated to language. In contrast, recent empirical evidence suggests that normal human cognition is actually language-augmented cognition. Exploring the role that language learning and language use exerts on human cognition leads to a better understanding of the evolutionary trajectory of language and offers a partial solution to the puzzle of how humans have come to possess intellectual capacities that could not have evolved through natural selection ("Wallace's problem"). A computational framework for exploring the role language exerts on cognition is also discussed.

Vittorio LORETO
Sapienza University of Rome

A Cultural Route to the Emergence of Duality of Patterning

Duality of patterning refers to the organization of the meaningful elements in a language at two distinct levels: a combinatorial level where meaningless forms are combined into meaningful form, and a compositional level where meaningful forms are composed into larger lexical units. The question remains wide open regarding how such a structure did emerge. Here we address this question in the framework of a multi-agents modeling scheme, the Blending Game, where a population of individuals plays language games aiming at success in communication. We show that the two sides of duality of patterning can emerge simultaneously as a consequence of a pure cultural dynamics in a simulated environment which contains meaningful relations, when a simple constraint on message transmission fidelity is also considered. In addition we show how the theoretical predictions are in surprisingly good agreement with available empirical data, and we highlight new directions for experiments in this area.

Bart DE BOER
Vrije Universiteit Brussel

Interaction of Biological and Cultural Evolution of Speech: The Case of Combinatorial Structure

When investigating biologically evolved cognitive adaptations for language and speech, one has to confront the problem that languages are not static over time, but that they also evolve culturally. The interaction between fast cultural evolution and slow biological evolution causes the mapping between observable properties of language and underlying cognitive biases to become very indirect. This makes it difficult to establish what these biases are, and some researchers have even proposed recently that there are no language-specific cognitive biases.

This presentation focuses on one aspect of speech that humans most likely have cognitive adaptations for: combinatorial speech. It presents experimental ways to disentangle the effects of culture and cognitive biases, but also discusses from a cross-species comparative perspective how to measure combinatorial structure, and what combinatorial structure is exactly.