The Evolution of Language

3rd International Conference
3 - 6 April 2000
Ecole Nationale Supérieure
des Télécommunications
Paris - France

TELECOM PARIS
CNRS - URA 820

Proceedings

Edited by Jean-Louis Dessalles & Laleh Ghadakpour

Sponsored by:

Sony CSL
Sony Computer Science Laboratory Paris
TELECOM PARIS
UeL University of East London
F.LLPR
france telecom
CNRS
GET

ENST 2000 S 002

The Evolution of Language

3rd International Conference - 3 - 6 April 2000

Ecole Nationale Supérieure des Télécommunications
Paris - France

Organised by:

Jean AITCHISON (Oxford University),
Jean-Louis DESSALLES (ENST Paris),
Jim HURFORD (Department of Linguistics, University of Edinburgh),
Chris KNIGHT (Department of Sociology and Anthropology, University of East London),
Luc STEELS (Sony CSL and Vrije Universiteit Brussel).

Progamme Committee:

Jean AITCHISON (Oxford University),
Robert C. BERWICK (M.I.T.),
Derek BICKERTON (Univ. Hawaï),
Ted BRISCOE (University of Cambridge Computer Laboratory),
René CARRÉ (ENST Paris),
Bernard COMRIE (Max Planck Institute for Evolutionary Anthropology, Leipzig),
Jean-Louis DESSALLES (ENST Paris),
Jean-Marie HOMBERT (MSH Rhône-Alpes),
James R. HURFORD (University of Edinburgh),
Michel IMBERT (Université de Toulouse),
Judy KEGL (University of Southern Maine),
Simon KIRBY (University of Edinburgh),
Chris KNIGHT (University of East London),
André LANGANEY (Musée de l'Homme),
Frederick J. NEWMEYER (University of Washington),
Michael STUDDERT-KENNEDY (Haskins Laboratories),
Luc STEELS (Sony CSL & Vrije Universiteit Brussel),
Bernard VICTORRI (ENS Paris).

Local organisation:

Jean-Louis DESSALLES (ENST Paris), Laleh GHADAKPOUR (CREA Paris), Frédéric KAPLAN (Sony CSL), Luc STEELS (Sony CSL and Vrije Universiteit Brussel), François YVON (ENST Paris).

The Evolution of Language

3rd International Conference

> **"Human language**
> **is an embarassement**
> **for evolutionary theory"**
> **(D. Premack)**

This is the third conference in a series concerned with the evolutionary emergence of speech. From a wide range of disciplines, we seek to attract researchers willing to integrate their perspectives with those of modern Darwinism.

The aim is to bring together linguists, computer scientists, anthropologists, palaeontologists, ethologists, geneticists, neuroscientists and other scientists who are concerned with the question of the origin and evolution of language.

Who's afraid of the co-evolution of medial and lateral cortices for speech?

Christian ABRY & Rafael LABOISSIÈRE

Institut de la Communication Parlée, CNRS UMR 5009,
ICP-Stendhal, BP 25, F-38040, Grenoble Cedex 9.
abry@icp.inpg.fr

ABSTRACT

Speech ontogeny and pathology: the medial carrier component of babbling and CVCV aphasia

MacNeilage gave in BBS 98 a thorough presentation of his “Frame/Content Theory of Evolution of Speech Production”. In our comment, we supported globally his view that 7-months babbling cannot originate in Broca’s area. In addition to his arguments in favour of SMA (or a part of it), coming from cortical mapping and irritative lesions, we added a new piece of evidence, still unnoticed by its discoverers themselves: CVCV aphasia. We can now totalize an amount of about 60 patients, “brothers/sisters” of famous Broca's patient Tan, who actually repeated “tan, tan”. Moreover such fluent “Tantan” or “Titi-Titi” global aphasics, some of them with lateral lesions, cannot support the proposal opposed to MacNeilage by three eminent students in motor control (Abbs, Jürgens and Lund), namely to lump into the lateral system both frame and content.

Concerning MacNeilage's view of babbling as a simple jaw based cyclicity, we are now able to demonstrate, by using articulatory models built from different subjects, that his “pure” labial frames are in fact idiosyncratic: labial, coronal or corono-labial, depending mainly on the anatomy of the baby (i.e. “why papa, mama…and tata” explained). Pushing a step farther, and given lips and tongue compliance, babbling variegation, could simply originate from biological noise in the motor signal generating jaw cyclicity.

A central question remains: what is the role of Broca? and when? Contrary to MacNeilage, we argue that a lot of speech behaviour in the first two years of life does not need to make a “loan to the intelligence” of Broca, as proponed by his BBS target paper. We will evaluate the benefit of delaying the appearance on the scene of such a morphosyntactic device as Broca’s, for the emergence of phonological universals, say the syllable related to first word prosody, and segment types related to articulator independence. This for ontogenetic developement.

Speech phylogeny: The vocal self and the alien monitoring medial systems

Backwards, on the phylogenetic side, Rizzolatti made the same premature loan, when becoming acquainted of MacNeilage’s proposal. He argued on the base of the perception/action system, discovered by his team, called “mirror neurons”, that neurons in the homologue of Broca’s area could match the “observation” and the “execution” of visual lipsmack communicative cyclicities. In fact, for the 3-day-old rhesus monkey, which demonstrates these affiliative lipsmacks, a “Broca” control is truly impossible, and Rizzolatti’s proposal is clearly a visuo-motorically biased misreading of MacNeilage’s target article, which insists both on the visual and audio nature of such a “precursor”. Hence we need to go back to the non-lateral system (if not to SMA).

In fact Rizzolatti, like many other neuroscientists, seems to admit that non human primate vocalisations are not under cortical volitional control. This is obviously true when homologues of Broca’s and Wernicke’s areas (hence B-W) are lesioned. But not for the anterior cingulate gyrus (ACG, the head of the animal vocalisation system) and its neocortical extension SMA. We need to emphasize the endogenous, volitional control of this pathway, clearly defined as instrumental vocalization. For us this volitional aspect is the stem (and logically the necessary condition) of the vocal self monitoring system. This animal model of the vocal self has inspired a human model of auditory hallucinations in schizophrenia by Frith. One of its crucial feature is the use of B-W circuitry for the implementation of the corollary discharge hypothesis: this enables the recognition of one's own expected vocalizations.

We will claim that, for the moment, it seems impossible to get B-W circuitry involved farther in the evolution of language, without calling for an evolutionary Theory of Mind. This is precisely what is comprised into Frith’s and colleagues’ recent theoretical and experimental advances under the heading of a vocal alien monitoring system, found in the medial area BM 8. Whereas the elaboration of a Theory of Mind seems to us problematic with the action understanding system of “mirror neurons”. It remains that, in our opinion, such a system is still in need of an adequate biocybernetic model, Jordan-Wolpert’s forward models, as the one used by Frith and colleagues in collaboration with Wolpert, being just able to implement the corollary discharge hypothesis, i.e. the initial vocal self monitoring system.

Lateral and medial cortical speech in evolutionary cognition

Our conclusion will be that, due to the focalization of research on the perisylvian system of language, the fractionation of the basic subsystems of speech has been waiting for too long. It's now time to look (i) at the basic system allowing the first steps of speech – babbling and its remnant in a very specific aphasia; and also (ii) at the two even more basic systems which allow people to talk to each other – the vocal self and alien monitoring systems. The co-evolution of these medial systems with the lateral ones will consequently appear less miraculous: (i) the emergence of the degrees of freedom for the carried articulators, i.e. the birth of their relative independence from the jaw, allowing the learning of language-specific segments on the basis of universal phonetic “places” and “manners” ; (ii) the connection with the medial cortex of the rich and specific systems for understanding the semantics of action, which is crucial for language learning.

Two conceptions of the emergence of phonemic structure

Irene APPELBAUM

Department of Philosophy, University of Montana
Missoula, Montana 59812
appel@selway.umt.edu

Introduction[1]

In a series of papers Björn Lindblom argues that phonemic structure is emergent, in the sense that it arises in a self-organizing manner[2]. Central to this description is the claim that phonemic structure is not prespecified as a “mental blueprint”. More generally, on this view, phonemic structure is not the direct consequence of the causal forces that produce it; rather, it emerges as a by-product of the interaction among sound-pattern primitives, performance constraints, and the demand for increased vocabulary. The sound-pattern primitives, called “gestures”, are specified as sequences of articulatory closures and openings, corresponding to stop-consonant and vowel configurations, respectively. Thus, the gesture inputs are indistinguishable - in articulatory, acoustic, and perceptual terms - from CV syllables. This gives rise to the following objection: Phonemic structure arises from primitives which are indistinguishable from CV syllables. The presence of CV syllables seems to indicate the presence of phonemic structure. Thus, phonemic structure seems to be primitive and prespecified, contrary to the requirements of a self-organizing system.

Lindblom, anticipating this objection, denies that the gestural primitives of his model are in fact CV syllables[3]. He acknowledges that they “resemble” CV syllables, but emphasizes that they are not “analyzed” as such; instead, the gesture is defined as a “holistic transition” and as a “Gestalt trajectory”:

> Although these gestures resembled stop-vowel syllables, their specification did not presuppose an analysis in terms of segments. Rather a possible gesture was defined as a holistic transition running between an arbitrary point in the universal phonetic space of 'possible closures' and a similarly arbitrary point in the universal space of 'possible vowels'. (Lindblom et al, p. 181)

My aim in this essay is to analyze and evaluate Lindblom's above response. I will first try to show that his reply, as it stands, is inadequate. I will then present two different ways his reply can be elaborated and strengthened, each of which leads to a different conception of what precisely the self-organization of phonemic structure consists in. I will conclude by suggesting that neither conception is wholly satisfactory.

Lindblom's reply

The problem with Lindblom's reply is it suggests that what the emergence of phonemic

---

[1] I am grateful to the Center for the Study of Language and Information, Stanford University, for supporting this research.

[2] See, for example, Lindblom (1984, 1986, 1998) and Lindblom, MacNeilage, and Studdert-Kennedy (1984).

[3] Actually, in earlier papers, Lindblom identifies these sound-pattern primitives as "syllables" (e.g., "We make the assumption that the syllable is an axiomatically given primitive in our theory." (1984, p. 72); "We shall assume that the syllable is an axiomatically given primitive of our theory". (1986, p. 502). In more recent papers he is careful to avoid this formulation.

structure consists in, is simply a change in perspective on pre-existent structure, rather than a change in the structure itself. It suggests that the process he describes as emergent is not itself structure-generating, but simply structure-recognizing. The inputs to Lindblom's model are CV syllables which are not analyzed as such and the outputs are CV syllables which are analyzed as such. Hence, it appears that the only thing that changes is the analysis, not the structure itself.

Compare the above with Lindblom's favorite non-linguistic example of a self-organizing system: termite nest building. The output of the latter process is an intricate structure of pillars and arches. But the inputs are just randomly distributed piles of (glutinous) sand. Thus the emergence of termite nests consists in a radical change in the organization and structure of the sand, not a change in the analysis of the sand.

Nevertheless, I think there are two stronger defenses implicit in Lindblom's reply, which I will try to make explicit below. I will call the first reply “Recombination” and the second “Asymmetric Dependence”.

The Recombination View

The Recombination View denies that phonemic structure is present from the outset, but not by denying that the primitives are physically indistinguishable from CV syllables. Rather, the Recombination View denies that physical structure is the relevant level of structure at all. According to this view, the essential feature of phonemic structure is not that it is composed of a series of physically independent units (closures and openings)[1], but rather that the units are functionally independent. It is only once a single closure or opening by itself signifies a difference in meaning - that is, only once minimal pairs are present - that phonemic structure can be said to have arisen. By contrast, the initial input contains pairs of physically independent closures and openings, but they do not function independently.

Note that the Recombination View does not depend on how the original inputs are analyzed. In particular, it does not depend on refraining from interpreting the inputs as a sequence of two segments. Rather, the point is, it makes no difference whether we do or not, because however we analyze them, the initial inputs function as holistic units. On the other hand, once minimal pairs are present, the sound patterns are analyzed as two independent segments because they now function that way. Thus although a sequence corresponding to [ba] would not be analyzed as a CV syllable in the initial input and it would be analyzed as such in the output system, the change is not merely a change in the analysis. Rather, the change in the analysis itself reflects a change in the input - it now exhibits functional structure.

So it appears that phonemic structure is not prespecified as the initial objection charges. But is this account of how phonemic structure arises appropriately described as a self-organizing process? The local causal mechanism is recombining closures and openings. Whether this is a self-organizing process depends on whether it is best described as a case of indirect causation.

It seems safe to assume, as Lindblom does, that the process is not the result of an explicit and conscious attempt on the part of our ancestors to create phonemic structure. But this does not settle the question of how directly caused the process is. Lindblom is clear about the pressures that produce minimal pairs - demand for increased vocabulary and demand for sounds with low articulatory costs relative to perceptual benefit. The latter constraint limits the available pool of closures and openings and hence the number of patterns which do not share either a closure or

---

[1] In one sense, of course, the closure and opening are physically continuous, but the present point is simply that they are physically distinguishable and that Lindblom treats them as independent insofar as the initial inputs are specified as consisting of one closure and one opening.

an opening. When the former demand overloads the resources of the latter, reusing closures and openings is the only way to satisfy demand. So there need be no antecedent plan to generate phonemic structure; the external constraints are sufficient.

But there remains a question of whether phonemic structure arises as a direct or indirect consequence of these external pressures. It would seem to be as a direct consequence because these pressures lead directly (if blindly) to reusing closures and openings, and there is little, if any, explanatory distance between the process of recombining closures and openings and the process of generating phonemic structure. The former process just is the generation of phonemic structure.

By contrast, consider again, the termite nest building case. The local causal process in this case is termites depositing sand in response to the presence of pheromones. In so doing, they end up creating a structure of pillars and arches. But the process of depositing sand in response to pheromone presence and the process of generating an intricate pillar and arch structure are not ipso facto the same. One could easily imagine the sand structure being produced without a process of responding to pheromones, but it is more difficult to imagine phonemic structure being produced without a process of recombining primitives.

Thus, it seems that the process which generates phonemic structure is too direct to count as emergent. However, I am less concerned to press this conclusion, than to emphasize that in order to maintain that phonemic structure is a self-organizing process (on the Recombination View) one must deny this conclusion.

The asymmetric dependence view

But there is a second way to interpret Lindblom's claim that the gestural primitives are not analyzed as exhibiting phonemic structure, and hence that phonemic structure is not prespecified. On this second account what's important is not in the first instance whether discrete segments form minimal pairs or not, but whether discrete segments are present at all. Consider that the CV syllable patterns are, in one sense, composed of discrete units and in another sense, not. They are defined as consisting of two distinct components - closure and opening. They can be distinguished in articulatory and acoustic terms.

But, from another point of view, they are holistic. One hears a syllable as a single sound contour; one doesn't hear it as two physically independent segments. More specifically, one cannot separate the syllable into two perceptually independent units insofar as the stop-consonant onset cannot be pronounced without some (i.e., even tiny) bit of the following vowel. By contrast, vowels are not similarly dependent on onsets. Vowels can be pronounced independently; a syllable can consist of a vowel only, but not of a consonant only[1]. So one can say there is an asymmetric dependence between consonant and vowels (or between closures and openings) in the formation of CV syllables. Consonants are dependent on vowels in a way that vowels are not dependent on consonants..

This asymmetric dependence is present whether we are considering syllables in a system with or without minimal pairs. It's a physical, not functional, dependence. Yet in a system with minimal pairs - that is, in a system with phonemic structure, the consonants take on an independent functional existence as well: [pa] and [ba] is every bit as much a minimal pair as [pa] and [pi].

Once phonemic structure is present there is a mismatch between the autonomy of the segments at the physical and functional levels (at least for consonants). Prior to the emergence of phonemic structure, one has sound patterns that are holistic both physically and functionally;

---

[1] Ignoring syllabic [r]'s.

subsequently one has segments, but they are functional segments. According to the Asymmetric Dependence View, then, what the emergence of phonemic structure consists in is the creation of discrete structure from physically and perceptually holistic patterns.

The difference between this view and the Recombination View is that on the latter view, having holistic syllables which differ in only either a closing or an opening, but not both, just is phonemic structure, whereas on the present view, it is not phonemic structure until it is analyzed as such. On the Asymmetric Dependence view, it is only when holistic syllables are compared or juxtaposed, that segments emerge for the first time. What makes them emergent on this view is that they arise by abstraction from contrasting syllables; they are implied or inferred from the physical patterns.

According to the present view, Lindblom's claim that the initial inputs are not sequences of segments because they are not analyzed as such, means that you only get sequences of segments at the functional level; neither initially nor ultimately, do you have physical segments. And you only get sequences of segments at the functional level when you have holistic patterns that are only partially distinct. These partial distinctions are then analyzed as discrete segments.

Notice, though, that on this view, phonemic structure emerges only after the model's outputs are given a functional analysis. The outputs themselves do not yet constitute phonemic structure. Even though certain pairs of them differ only in the initial or latter part, they are nevertheless holistic syllables until discrete structure is imputed to them. And discrete structure is imputed to them only by an indirect process of comparing holistic syllables. Thus it does not seem problematic on the present view to say that phonemic structure arises by process of indirect causation. But then what the emergence of phonemic structure consists in is largely analyzing physically holistic syllables as discrete units but, significantly, the holistic syllables are already functioning as minimal pairs, insofar as pairs of them may differ in only an opening or only a closing.

Conclusion

On the Recombination View, phonemic structure arises as the consequence of recombining primitives. But it arises as a perhaps too-direct consequence to count as emergent. On the Asymmetric Dependence View, phonemic structure emerges by an indirect process of comparing of holistic syllables and abstracting discrete units from them. However, these holistic syllables are the outputs of Lindblom's model and already function as minimal pairs. Thus, we seem to be left with the choice between a view (Recombination) which takes phonemic structure to emerge from pre-existent discrete, physical structure, and a view (Asymmetric Dependence) which takes it to emerge from pre-existent continuous, functional structure.

Bibliography

Lindblom, B. (1984). Can the models of evolutionary biology be applied to phonetic problems?. In Van den Broecke, M.P.R. & Cohen, A. (eds.) Proceedings of the Tenth International Congress of Phonetic Sciences. Dordrecht: Foris: 67-81.

Lindblom, B. (1986). On the origin and purpose of discreteness and invariance in sound patterns. In Perkell, J.S. & Klatt, D.H. (eds.) Invariance and Variability in Speech Processes. Hillsdale, NJ: Earlbaum: 493-510.

Lindblom, B. (1998). Systematic constraints and adaptive change in the formation of sound structure. In Hurford, J.R., Studdert-Kennedy, M., & Knight, C. (eds.) Approaches to the Evolution of Language. Cambridge: Cambridge University Press: 242-264.

Lindblom, B., MacNeilage, P., & Studdert-Kennedy, M. (1984). Self-organizing processes and the explanation of phonological universals. In Butterworth, B, Comrie, B. & Dahl, O. (eds.) Explanations for Language Universals. New York: Mouton: 181-203.

The origin of symbols in the brain

Christian BALKENIUS, Peter GÄRDENFORS & Lars HALL

Lund University Cognitive Science
Kungshuset, Lundagård
S-222 22 Lund, Sweden
Christian.Balkenius@lucs.lu.se, Peter.Gardenfors@lucs.lu.se,
Lars.Hall@lucs.lu.se

Introduction

Deacon's (1997) book is an interesting attempt to explain the critical aspects of the evolution of language as the learning of symbolic relationships. Deacon blurs the traditional distinction between syntax and semantics by arguing that the meaning of symbols is primarily determined via the combinatorial relations between symbols, and only secondarily via an indexical relation between a symbol and a referent (Deacon 1997, Ch. 3). However, this account of how acquisition of symbols involves multiple hierarchies of associative learning has proved rather difficult to understand (Hurford 1997), and even more difficult to incorporate into an explicit representational model. In this article, we want to use Deacon's theory as a platform for a more elaborated and precise model of symbol learning.

Our model will be presented in rough phylogenetic order, and will contain only those cognitive elements that are minimally required for the learning of symbols. These mechanisms have evolved for other purposes than symbolic processing, but taken together they form a substrate for the emergence of a symbolic ability. The purpose of this article is therefore to show that given the coevolutionary dynamics of Baldwinian evolution (i.e. that behavioral adaptation tends to precede biological/cognitive change, Deacon 1997, p345, see also Laland, Odling-Smee & Feldman, to appear), and some rather conservative assumptions about the cognitive capacities preceeding the evolution of language, the end state of modern human language ability can be reached.

These capacities can be divided into three groups: (1) general learning abilities that are present in many animal species; (2) communicative capacities, some of which are present in apes and some of which coevolve with language; and (3) learning of symbols in the form of symmetrical signs, i.e. relations between vocal sounds (or gestures) and representations, and also relations between such signs.

General learning abilities

1. *Classical Conditioning* is the ability to form associations from sensory representations to events that are rewarding or punishing and it critically depends on the amygdala (Rolls, 1995). This is an ability that is present even in very primitive organisms.

2. *Internal Inhibition* reverses the influence of conditioning when expected rewards or punishments are omitted. This is called extinction and is controlled by the orbital prefrontal cortex (Rolls, 1995).

3. *Expectation Learning*. More advanced organisms can associate more freely between sensory representations to predict the occurrence of an event based on another event also in

cases when they are not connected with any innate needs (Mackintosh, 1983). However, these associations may still be limited and are not necessarily symmetrical.

4. *Context Learning*. An important step forward is the evolution of the hippocampus which makes contextual representations possible. Unlike simple associative learning that deals with the sensory representations in isolation, context learning is concerned with relations between representations. Initially, these representations are mainly spatial (Eichenbaum et al. 1999). The connections between objects and locations are mostly arbitrary which makes locations efficient cues for recall of objects (Balkenius & Morén 2000). In neural network parlance, they make the representations of specific objects more orthogonal.

5. *Contextual Discrimination*. One important role played by contextual representations appears experimentally in omission training (Schmajuk & Holland 1998). While the initial learning is mainly context independent, extinction through omission of reward depends almost entirely on the context (Bouton 1991). This mechanism learns to suppress previously learnt associations when certain relations among representations are present. The prefrontal cortex has been implicated in this type of inhibitory learning (Fuster 1997, Rolls 1995) which has evolved from the more primitive extinction mechanism described above.

6. *Expectation Matching*. The above mechanism is driven by the omission of reward but many animals are also able to learn by expectation matching independent of any reward or punishment. When expectations are not met, a contextual inhibition develops for the incorrect expectations. This matching is one of the main functions of the septo-hippocampal system (Gray 1995).

Communication

The following capacities are connected with the evolution of communication. As Oliphant (to appear) makes clear, the difficulties inherent in moving from an innate to a learned system of communication lies not primarily in how to learn to associate a sound (or sign) with an object, but how to observe the correct sound-object meaning pairs.

1. *Joint attention*. By exploiting gaze contact, head-direction, body-posture, etc., individuals can establish joint attention to an object or event (Langton, Watt & Bruce 2000). Joint attention makes communication by ostention possible since one can then show that one wants to show something (Gomez 1998). Comparative evidence suggests that chimpanzees in the wild engage in such ostensive behavior (Leavens & Hopkins 1998).

2. *Vocalisation/miming*. If ostensive communication is combined with vocalisation or a miming gesture, signs for different kinds of objects can be established within a community. This sets up selection pressure for increased cortical control of the hands and facial muscles, and enables further enrichment of the system (Deacon 1997, ch. 8). The extent to which imitative and/or theory-of-mind-like abilities are required for, or a consequence of, the development and maintenance of such a conventional system, remains unclear (Dennett 1996, Burling 1999, Schaal 1999).

3. *Spontaneous Generation*. The next step on the road to full symbolization is spontaneous production, or babbling. Most importantly, this was not a matter of one clever hominid suddenly opening its mouth and letting the world know what it thought. As Dennett (1997) explains, the habit of adding a ‘soundtrack’ to one’s activities must be driven into place *before* the words themselves are understood. “For a word to serve as a useful, manipulable label in the refinement of the resources of a brain, it must be a ready enhancer of sought-for associations that are already to some extent laid down in the system” (Dennett 1997, p 347).

Symbolicity

1 *Symmetrical Signs*. In a phylogenetic perspective, babbling most probably coevolved with the advent of crude indexical language use, but as a simple example of automatic, instinctual resource enhancement it had far-reaching cognitive consequences. In our model it lays the foundation for the development of *symmetrical* associations between a representation, its corresponding sound pattern and the patterns of muscle activity that produces it. It seems unlikely that nonhuman primates can acquire such symmetrical associations (Lowe & Dugdale 2000). We want to propose that this paves the way for an emerging Saussurian sign (de Saussure 1972). As a consequence, this opens up the all important class of *implicit negative evidence* that modern language learners rely upon.

2. *Words as Orthogonal Code*. Since words are arbitrary, they are in a similar position as a place representation to enhance storage of a sensory representation (Clark & Thornton 1997). These orthogonal codes make it possible for signs to enter into indexical associative relations (Deacon 1997) as well as contextual representations which form the basis for symbolic relations.

3. *Symbolic Discrimination*. The final mechanism required for symbolic learning is a generalization of the earlier context system. Instead of working on spatial locations, it now operates on symmetrical signs, where sound patterns take on the role previously played by spatial locations.

It is well known that the prefrontal cortex is involved in the representation of sequential structures and it is well fitted for the inhibition of “ungrammatical” sequences or combinations. We suggest that by being explicitly unrewarded or having expectations unconfirmed, combinations of tokens can take on combinatorial properties. This implies that the learning of symbols proceeds through the acquisition of context-sensitive rules that code for incorrect combinations. The context, in this case, mainly consists of other symbols. As a consequence, explicit negative feedback is not necessary for learning symbolic relations, in contrast to what has been claimed by Chomsky and his followers (cf. Regier 1996, Rhode and Plaut 1999, Schoenemann 1999, Morris, Cottrell & Elman, in press).

Our model also fits well with recent empirical studies of language learning in children. The constructive grammar accounts by Carpenter, Nagell, Tomasello (1998), Gillette, et al. (1999) and Tomasello (2000) all show that children tend to build their communicative competence in an incremental fashion by rote learning of pragmatically relevant sentences, and then eventually *discover* the common structure behind them. This process is described by Tomasello (2000) as children “not just combining words or isolated linguistic categories, they are combining pre-compiled linguistic constructions of various shapes, size, and level of abstractness.” (p 243). As Tomasello admits, the *mechanisms* underlying this capacity (what he calls ‘structure mapping’ and ‘analogy making’) are still underspecified.

Our model attempts to address this problem and specify the mechanisms behind symbolic learning. Several parts of the model have already been tested in computer simulations (Balkenius & Morén 1999, 2000) and in the future we intend to perform simulations of the complete model.

References

Balkenius, C. & Morén, J. (1999) Dynamics of a classical conditioning model. *Autonomous Robots*, 1(7).

Balkenius, C. & Morén, J. (2000). A computational model of context processing, submitted to SAB 2000.

Bouton, M. E. (1991). Context and retrieval in extinction and in other examples of interference in simple associative learning. In L. Dachowski & C. F. Flaherty (Eds.) *Current Topics in Animal Learning: Brain, Emotion and Cognition*, Erlbaum: Hillsdale, N. J.

Burling, R. (1999). The Cognitive Prerequisites for Language. *Psycoloquy* 10(032).

Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development.* 255:63:4.

Clark, A., & Thornton, C. (1997). Trading spaces: computation, representation and the limits of uninformed learning. *Behavioral and Brain Sciences*. 20. 57-90.

De Saussure, F. (1972). *Course in General Linguistics*. London: Duckworth.

Deacon, T. (1997). *The Symbolic Species*. London: Penguin Books.

Dennett, D. C. (1996). *Kinds of Minds: Toward an Understanding of Consciousness*. Basic Books.

Dennett, D. C. (1997). How to do other things with words. *Philosophy: the journal of the Royal institute of philosophy.*, Iss. 42.

Eichenbaum, H., Dudchenko, P., Wood, E., Shapiro, M., & Tanila, H. (1999) The Hippocampus, Memory, and Place Cells: Is It Spatial Memory or a Memory Space? *Neuron*, Vol. 23, pp. 209–226.

Fuster, J. M. (1997). *The Prefrontal Cortex: Anatomy, Physiology, and Neuropsychology of the Frontal Lobe*. Philadelphia: Lippincott-Raven.

Gillete, J., et al. (1999). Human simulations of vocabulary learning. *Cognition*. 73. 135-176.

Gomez, J-C., (1998). Some thoughts about the evolution of LADS: With Special reference to TOM & SAM. In Carruthers, P., & Boucher, J. (Eds.). *Language & Thought: Interdisciplinary themes*. Cambridge: Cambridge University Press.

Gray, J. A. (1995). A model of the limbic system and basal ganglia: Application to anxiety and schizophrenia. In Gazzaniga, M. S. (Ed.) *The cognitive neurosciences* (pp. 1165-1176). Cambridge, MA: MIT Press.

Hurford, J. (1997). Review of Deacon: The Symbolic Species, Allen Lane, The Penguin Press, London. *The Times Literary Supplement*, October 23rd, 1998, p.34.

Laland, K. N., Odling-Smee, J., & Feldman. M. W. (to appear) Niche Construction, Biological Evolution and Cultural Change. *Behavioral and Brain Sciences*.

Leavens, D. A., & Hopkins, W. D. (1998). Intentional communication by chimpanzees: A cross-sectional study of the use of referential gestures. *Developmental Psychology.* 34:5. 813-822.

Lowe, F., Dugdale, N. (2000). Testing for symmetry in the conditional discrimination of language-trained chimpanzees. *Journal of the experimental analysis of behavior*. 73. 5-22

Mackintosh, N. J. (1983). *Conditioning and associative learning*. Oxford: Oxford University Press.

Morris W.C., Cottrell, G.W., & Elman, J.L. (in press). A connectionist simulation of the empirical acquisition of grammatical relations. In Stefan Wermter & Run Sun (Eds.) *Hybrid Neural Symbolic Integration*. Springer Verlag

Oliphant, M. (to appear). The learning barrier: Moving from innate to learned systems of communciation. *Adaptive Behavior*.

Regier, T. (1996). *The Human Semantic Potential: Spatial Language and Constrained Connectionism*. Cambridge, MA: The MIT Press.

Rhode, D. L. T. & Plaut, D. C. (1999). Language acquisition in the absence of explicit negative evidence: how important is starting small? *Cognition* 72. 67-109.

Rolls, E. T. (1995). A theory of emotion and consciousness, and its application to understanding the neural basis of emotion. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 1091-1106). Cambridge, MA: MIT Press.

Schaal, S. (1999). Is imitation learning the route to humanoid robots. *Trends in Cognitive Sciences*. 3:6. 233-242.

Schmajuk, N. A. & Holland, P. C. (Eds.). (1998). *Occasion Setting: Associative Learning and Cognition in Animals*. Washington, DC: American Psychological Association.

Schoenemann, P. T (1999). Syntax as an emergent characteristic of the evolution of semantic complexity. Minds and Machines. 9:3. 309-346.

Langton, S, R. H., Watt, R. J., & Bruce, V. (2000). Do the eyes have it? Cues to the direction of social attention. *Trends in Cognitive Sciences*. 4:2. 50-59

Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition*. 74. 209-253.

Negotiating syntax

John Batali

batali@cogsci.ucsd.edu
Department of Cognitive Science,
University of California at San Diego
9500 Gilman Drive; La Jolla, CA 92093-0515 - (619) 534-7308
http://cogsci.ucsd.edu/~batali/

ABSTRACT

If accurate communication provides benefit to both participants in a communicative event, at least some of the time, the sender will motivated to produce signals that the receiver is likely to interpret accurately. and the receiver will be motivated to interpret a signal as the meaning it is most often used to express. In a population where this is true, a conventional communication system can result from a process of of negotiation, as the agents alternate between contributing to, and conforming with, the emerging system.

A conventional communication system could be based on a simple table in which each meaning to be conveyed is associated with a unique signal. However as the number of meanings increases, the system becomes increasingly difficult to learn and to use.

If the set of meaning is large, the agents probably don't use a distinct internal representation for each of them. Their representations are most likely constructed from a relatively small number of component types. The interpretation of an internal representation depends on the specific components use to construct it, and how they are configured.

When an agent expresses a meaning, it will, in general, use aspects of the structure of its representation of the meaning in its derivation of the structure of the signal used to express that meaning. The receiver's interpretation of the signal will make use of the signal's structure to derive a representation of the meaning it conveys. The derivations performed by the sender and receiver therefore constitute implicit analyses of the relations between the structures of meanings and signals.

The structural derivations performed by the sender and the receiver might have no relation to one another, except that they both involve the same signal, and, if the agents are are lucky, equivalent meanings.

For them to achieve better than chance accuracy, I assume that the members of a population obey a set of negotiated conventions regarding possible derivational relationships between the structures of signals and meanings. These conventions can emerge if senders perform their derivations of signals while considering how receivers might interpret them, and if receivers derive their interpretations of signals while considering how they might have been constructed by senders. By including recursive characterizations of structural properties, the set of conventions can be used to coordinate communication of an unbounded set of meanings with a relatively small set of conventions. Analyses of structural mappings between signals and meanings are used by learners as they attempt to discover the conventions the users of a communication system obey, and as they attempt to perform derivations of signals and meanings in accord with the conventions.

In this account, syntactic structure is a representation of the conventional aspects of the structural derivations performed by senders and receivers. It is not reducible to the domains of meaning or phonology, although regularities or constraints in either domain may influence aspects of syntactic structure. Other syntactic regularities and constraints can emerge as a result of the negotiation process.

Results from a number of recent computational models consistent with this account suggest that agents with fairly general representational and learning abilities can negotiate communication systems capable of accurately conveying very large numbers of meanings. The systems that emerge from the negotiations incorporate structural regularities and constraints that resemble some aspects of the syntax of human languages.

Foraging versus social intelligence in the evolution of protolanguage

Derek BICKERTON

ABSTRACT

Currently there is wide agreement that the origins of language are to be sought in primate social intelligence, rather than in foraging, tool-making and the like. This is probably correct for the syntactic structure that gave us true language (Calvin and Bickerton 1999) but unlikely for the protolanguage that preceded it. Several independent lines of argument support this conclusion:

1. *Selective uniqueness*: Only humans among primates developed language. This suggests that the original impetus towards language came from a selective pressure unique to the hominid line, rather than from the highly developed social intelligence that characterizes all advanced primates.

2. *Absence of complexity*: In response to (1) it is sometimes claimed that growing complexity in hominid social life selected for language. But (a) primate social life is already complex (b) there is no independent evidence that early hominid social life was more complex than that of chimps or bonobos (c) any higher degree of complexity in contemporary social life results from the addition of language to typical primate relations.

3. *Hominid ecology*: With few if any serious predators and locally-concentrated, readily-available, year-round food supplies, primates and modern humans have abundant time for developing social life. Early hominids did not: they faced formidable predation and utilized sparse, diverse, scattered and seasonal food sources.

4. *Initial functionality*: For any trait to fix, it has to be functionally effective from the outset. Before language can be used socially, it must have reached a certain level of complexity.

5. *Cheapness of tokens*: Words are cheap, therefore intrinsically unreliable if first used in social contexts where deception and manipulation are the norm.

Initially at least they would require validation through (a) trust (use in contexts where co-operation rather than competition was called for) (b) validation (use in contexts where their veracity–or otherwise–could be readily determined),

If protolanguage emerged in co-ordinating group foraging over large day-ranges and/or coping with predation, no problems arise. The selective pressure would be unique to hominids. No growing social complexity would be required. Strategies of survival would have bulked larger in hominid life than among other primates. Single-word tokens of warning or indications of food-source nature/location would have paid off from the start. The co-operative contexts of finding food and escaping predators would have generated trust, and practical proofs (finding food, escaping predators) would have quickly distinguished genuine from fake messages.

On the role of bridge theories in accounts of the evolution of language

Rudolf P BOTHA

Department of General Linguistics, University of Stellenbosch,
Stellenbosch, South Africa
rpb@maties.sun.ac.za

Theories of the evolution of human language express by their very nature claims of a historical sort: claims about why, when, where or how language emerged and/or developed in some distant past. An essential feature of these claims is that they are made in the absence of sufficient historical evidence about the evolutionary events, biological processes, physical forces, environmental pressures, kinds of (pre)linguistic entities and so on involved in the evolution of language. The paucity of this historical evidence – i.e., evidence derived from data contained in natural or man-made records of these evolutionary events, etc. – is generally seen as one of the most formidable obstacles to serious work on language evolution.

In modern work on the evolution of language, various strategies have been adopted for remedying the problem of evidential paucity and, in Pinker and Bloom's (1990, p. 727) words, for countering the “[s]kepticism about the possibility of saying anything of scientific value about language evolution”. A number of these strategies involve the use of new sources of data that are believed to bear on the truth of claims about language evolution. A first source comprises data of a historical sort, including paleoneurological data. For instance, in work such as that by Wilkins and Wakefield (1995), data about impressions on the interior surface of fossil skulls – or corresponding bumps or ridges on the exterior surface of endocasts – are brought to bear on claims about the emergence of (certain “preconditions” for) the language faculty. A second source comprises data of a nonhistorical sort. For example, Pinker and Bloom (1990) attempt to bring, amongst others, data about the reliance on lexical association by right-handed people with a family history of left-handedness to bear on the truth of claims expressed in their selectionist account of the evolution of the language faculty.

Both historical data of the sort used by Wilkins and Wakefield and nonhistorical data of the kind drawn on by Pinker and Bloom are data about entities belonging to ontological domains that are distinct from the ontological domain to which the language faculty belongs. As a cognitive entity, an ancestral language faculty is separated by a distinct ontological gap from both fossil skulls and the processing behaviour of people living at present. Being the different kind of thing they are, fossilized (fragments of) skulls simply do not contain direct information about cognitive faculties. Similarly, the processing behaviour of presently living right-handers with a family history of left-handedness does not offer in any direct way information about the evolution of a mental faculty that is claimed to have come into existence in a distant past.

To be able to bring data about fossil skulls or data about the processing behaviour of right-handers with a family history of left-handedness to bear on the truth of claims about the evolution of the language faculty, the ontological gap in question has to be bridged in a proper way. If this can not be done, data about the former entities can not be relevant to the truth of the latter claims. What are in essence required, are warrants licensing the various inferential jumps from paleontological or behavioural data to the truth of claims about the evolution of the language faculty. This gives rise to the question of what the source of the required inference licences might be.

Developing ideas outlined in (Botha, 1998, 1999), the present paper identifies bridge theories as such a source of inference licences. A bridge theory is characterized as a structured set of empirical assumptions which systematically interrelate properties of entities of one ontological domain with properties of entities of another ontological domain. To elucidate and appraise the role of bridge theories in modern work on language evolution, the paper focuses on the way in which data about fossil skulls or endocasts have been brought to bear on the truth of claims about the evolutionary emergence of language as a mental faculty. The peer discussion in *Behavioral and Brain Sciences* (Vol. 18, 1995) of Wilkins and Wakefield's target article on the emergence of the neural preconditions for language is used for illustrating the following points:

(1) The constituent assumptions of the required bridge theories are often not stated clearly or justified explicitly.

(2) The various licensing assumptions are not always “parcelled out” into different bridge theories but are lumped together into a single holdall theory.

(3) Some of the required bridge theories – or their constituent assumptions – are often quite controversial.

If the problem of evidential paucity is to be solved, the paper concludes, it is imperative that flaws such as (1) - (3) are eliminated from bridge theories.

References

Botha, R.P. 1998. Neo-Darwinian accounts of the evolution of language: 3. Questions about their evidential bases, logic and rhetoric. *Language and Communication 18*, 17 - 46.

Botha, R.P. 1999. On the testability of theories of language evolution. *Stellenbosch Papers in Linguistics 32*, 1 - 52.

Pinker, S., Bloom, P. 1990. Natural language and natural selection. *Behavioral and Brain Sciences 13*, 707 - 727, 765 - 784.

Wilkins, W., Wakefield, J. 1995. Brain evolution and neurolinguistic conditions. *Behavioral and Brain Sciences 18*, 161 - 182, 205 - 226.

Attention and the evolution of intentional communication

Ingar BRINCK

Dept of Philosophy, Lund University, Sweden ingar.brinck@fil.lu.se

A fundamental problem in explaining the evolution of human language is to account for the transition between *informational* communication, aimed at directly influencing the behaviour of the recepient, and *intentional* communication, aimed at indirectly influencing the behaviour as a consequence of changing the mental state of the recepient. Intentional communication is often held to depend on language and conceptual representations.

Informational communication is *goal-directed*. In goal-directed agency the goal is represented in, or during, acting, i.e., while the action is performed (Searle 1983). Informational communication aims at influencing the behaviour of the recipient to the advantage of the sender, usually by conveying information about the sender. It is successful when the recipient responds to the act (say, by fleeing when a warning call is emitted), or copies, or reproduces, the act (as sometimes is done with warning calls). Examples are the butterfly signalling that it is not edible by not being visible against the background of the flower it is sitting on, and the dog showing its teeth in anger.

Intentional communication is purposeful or deliberate, *goal-intended*, and about something else than the sender herself: the intended object. That it is goal-intended means that the sender represents the goal prior to acting (Searle 1983). The representation of the goal is independent of whether the action actually is performed or not. The aim of intentional communication is to change the mental state of the recipient and as a consequence the behaviour. A typical case is linguistic acts. Successful communication depends on that the recipient understands and recognizes the intention, or point, of the communicative act (Grice 1989; Sperber and Wilson 1986). One requirement for doing so is that sender and recepient succeed in *sharing the intentional object*.

The leap from informational to intentional communication can be explained by the emergence of certain areas of the brain. The question remains what made these areas evolve. I suggest that ways of communicating intentionally existed before language as we know it now developed. Pressures from existing quasi-intentional communicative strategies made the brain develop as it did, in order to handle increasingly complex situations, such as the increasing social co-ordination and planning pressed forward by environmental changes (the fourth glaciation); keeping up and cementing social bonds in groups (Dunbar 1993); or cultural adaptation as a result of intraspecies competition (Donald 1991).

To explain how communication can be intentional without language, I will address two questions: first, the nature of non-conceptual intentionality, and, second, how an intentional object may be determined and shared without language.

An important characteristic of intentionality is *directedness* (Brentano 1973): intentional states are *about* something (they have a content), and they are *aimed at* an object. Furthermore, the intentional object may not exist, and the representation of it is independent of any particular contextual constraints on the content. *Full-blown* intentional communication can make use of both referent- and context-independent representations (Tomasello and Call 1997).

I maintain that the elements necessary for intentional communication emerge with the capacity for *attention-focusing*. Attention-focusing has directedness in common with intentionality. To explain the notion of attention-focusing and how it is intentional, I will start from a general account of attention.

Attention is important for any information-gathering system, since it helps the organism in sifting among input. It has been debated wether there is an early selection or a late selection among input. Recent research suggests that there is selection at different stages of information-processing. Stimuli rejected at an early stage is less completely analysed than at a later stage. Semantic processing occurs late, while perceptual processing occurs early. Attention thus proves to be complex, making use of different kinds of processing (sensational, perceptual, conceptual) and working against different kinds of memory (iconic, procedural, perceptual, semantic, etc).

I distinguish between three attention-mechanisms: scanning, attention attraction, and attention-focusing (Brinck 1997). Under normal conditions animals (and humans) are immersed in a constant flow of information that provides a basic state of arousal. It functions as a background against which the subject herself and the part of her environment that she attends to stand out (Luria 1973; Gibson 1986). Attention consists in an increased awareness of something either external or internal to the subject. It can be directed at behaviours, sensations, perceptions, or conceptions, and may be involuntary as well as deliberate.

Scanning of the environment is continuous. Information is registered in a search directed at discovering possibilities to act (Gibson 1986). It guides movements and triggers actions in particular contexts. Action is supported by pragmatic representations that represent object attributes as affording specific motor patterns, not as cues for a given perceptual category (Jeannerod 1994). This does not call for binding attributes to a single entity. Crick and Koch (1990) hold that a fleeting awareness detects innate or overlearned features, while a serial attentional mechanism is in use when detected features are novel and unbound. Bodily movements, perception, and environmental changes are continuously attuned in an on-going, on-line co-ordination, forming an equilibrium (Shanon 1993).

Attention is *attracted* by events that are at odds with what is expected on the basis of previous experience. The perceptual system is geared to perceive changes in real time (Freyd 1987). Such changes shift the direction of the scanning and the behaviour. Discrimination in scanning and attention attraction depends on the detection of dissimilarity or similarity between items. Whether two items are discriminated or, on the contrary, perceived as similar, will depend on the context and on with what else they are juxtaposed.

In some cases, the subject will *focus* on that which happened to attract her attention and categorise it perceptually (Barsalou and Prinz 1997; Mandler 1992, 1997). She freezes a piece of the transient reality by suppressing most of the incoming information and simultaneously adding information in making perceptual inferences based on previous encounters with similar situations.

I maintain that attention-focusing makes it possible to focus on context-independent types of *attentional objects*. It is central for the capacity to re-identify events and objects across contexts and thus for grasping identities through time and space. Attention-focusing is necessary for context-independent content as used in intentional communication.

By producing an intentional object that can be shared among communicators, attention-focusing makes possible *goal-intended communication*. Goal-direction is a direct consequence of attention attraction, triggered by the environment and tuned to a particular action. Goal-intention, to the contrary, is generalised and the goal may be formed independently of the context. Since the sender's attention is towards categorised items, she can communicate about something else than herself.

The recepient can know what the sender intends and share the attentional object if sender and recepient engage in *joint attention*. Joint attention is, as noted by Gomez (1994, 1998) and Tomasello (1998, to appear), the clue to intentional communication. In what follows I develop the general notion of joint attention. The concepts of shared object and goal-intention are central to the account.

Joint attention is based on the ability of two or more subjects to focus their perception simultaneously on a *shared* attentional *object*: to engage in *object-focused* attention. An item is shared when two or more subjects can interact with or use it in a similar way. Sharing does not imply that subjects share all possible ways to interact with it. The notion is instrumentalist. Shared objects do not only exist beforehand, like trees to climb in, but can be created in interaction with the environment and other subjects, like branches as (chimpanzee) nutcrackers.

In *subject-focused* attention, attention is directed at the behaviour of other subjects. By, e.g., looking at each other, two subjects can find out their respective attentional objects on the basis of the direction of their respective movements in combination with a salient object that functions as a target. Co-ordination of attention based on saliency and behavioural co-ordination result in *mutual attention-focusing*. Note that perceptual saliency of features is connected to species-specific as well as individual values and affordances.

Mutual attention-focusing can spread automatically as a function of attention-attraction. Consider the behaviour of a group of antelopes when a predator is approaching. The predator attracts the attention of an antelope, and the other antelopes have their attention attracted by the behaviour of the first antelope. Attentive behaviour may spread automatically without resulting in mutual attention-focusing, if the object that attracted the attention of the first animal is no longer present or salient when the attentive behaviour of the group as a whole has been co-ordinated.

Joint attention consists in *subject-subject attention*. The subjects attend to each other as subjects capable of attending, and, moreover, of attending in a goal-intended way, i.e., in a way that is not controlled by the object of attention (as in attention attraction). Goal-intention provides for the capacity to either direct or follow the attention of the other subject in the absence of salient objects. This means that attention-focusing can be guided by the subjects' mutual attention to each other instead of by the environment.

Animals capable of subject-subject attention attend to each others attentional states, not to the behaviour that is a consequence of attention. They distinguish non-attentional head and body-orientation from attentional gaze and can engage in attention contact, during which they simultaneously check each others state of attention, e.g., by eye contact (Gomez 1994).

Directing somebody by gaze and pointing are ways of making one's focus of attention accessible to others. It is the target of the attentional state that is shared and made available for others, not the mental states themselves or their internal content. To achieve joint attention higher-order intentional states are not necessary. Since attentional states are manifested behaviourally, one does not need to represent the mental contents of other subjects in order to make them behave as one intends.

It is the perceivable directedness and implicit imperative force of gaze and gestures that make them communicative. Joint attention is referential, and reference serves the aim of inviting action. Thus communication can by exploiting attention be intentional without being metarepresentational or conceptual.

References

Barsalou, L. and Prinz, J. 1997: Mundane Creativity in Perceptual Symbol Systems. Ward, T. B., Smith, S. M., and Vaid, J. (eds) *Creative Thought: An Investigation of Conceptual structures and Processes*. Washington, DC: American Psychological Association

Brentano, F. 1973: *Psychology from an Empirical Standpoint*. London: Routledge & Kegan Paul

Brinck, I. 1997: *The Indexical 'I'*. Dordrecht: Kluwer Academic Publishers

Crick, F. and C. Koch 1990: Towards a Neurobiological Theory of Consciousness. *Seminars in Neurosciences* 2, pp 263-275

Donald, M. 1991: *The Origins of the Modern Mind*. Cambridge, MA: Harvard University Press

Dunbar, R. 1993: Co-evolution of neocortex size, group size and language in humans. *Behavioral and Brain Sciences* 16, pp. 681-735

Freyd, J. J. 1987: Dynamic Mental Representation. *Psychological Review*, vol. 94, 427-438

Gibson, J. J. 1986: *The Ecological Approach To Visual Perception*. Hillsdale, NJ: Lawrence Erlbaum Ass.

Gómez, J. C. 1994: Mutual Awareness in Primate Communication: A Gricean Approach. *Self-Awareness in Animals and Humans*. Parker, S. T., Mitchell, R. and Boccia, M. (eds) Cambridge: Cambridge University Press

Gómez, J. C. 1998: Some Thoughts about the Evolution of LADS, with Special Reference to TOM and SAM. *Language and Thought*. Carruthers, P. and Boucher, J. (eds) Cambridge: Cambridge University Press

Grice, P. 1989: *Studies in the Way of Words*. Cambridge, MA: Harvard University Press

Jeannerod, M. 1994: The Representing Brain: Neural Correlates of Motor Intentions and Imagery. *Behavioral and Brain Sciences* 17, pp. 187-202

Luria, A. R. 1973: *The Working Brain*. New York: Basic Books

Mandler, J. 1992: How to Build a Baby: II. Conceptual Primitives. *Psychological Review* Vol. 99 No. 4, 587-604

Mandler, J. 1997: Preverbal Representation and Language. *Language and Space*. Bloom, P., Peterson, M., Nadel, L. and Garrett, M. (eds) Cambridge, Mass.: MIT Press

Searle, J. 1983: *Intentionality*. Cambridge: Cambridge University Press

Shanon, B. 1993: *The Representational and the Presentational*. New York: Harvester Wheatsheaf

Sperber, D. and D. Wilson 1986: *Relevance: Communication and Cognition*. Cambridge, MA: Harvard University Press

Tomasello, M. 1998: Reference: Intending that Others Jointly Attend. *Pragmatics & Cognition* Vol. 6 (1/2), 229-243

Tomasello, M. to appear: Perceiving Intention and Learning Words in the Second Year of Life. *Language Acquisition and Conceptual Development*. Bowerman, M. and Levinson, S. (eds) Cambridge: Cambridge University Press

Tomasello, M. and J. Call 1997: *Primate Cognition*. New York: Oxford University Press

Macro and Micro Models of Linguistic Evolution

Ted BRISCOE

Computer Laboratory, University of Cambridge
ejb@cl.cam.ac.uk

Introduction

Mathematical modelling and computational simulation of dynamical systems are useful tools for refining and checking hypotheses about linguistic evolution. I compare macro-evolutionary models, which typically assume non-overlapping generations and in infinite populations to achieve analytic tractability, to stochastic micro-evolutionary models, which typically trade analytic tractability for more realistic demographic assumptions (e.g. Renshaw, 1991). The predictions of a macro and micro model for a simple linguistic example are dramatically different.

The NB model

Niyogi and Berwick, 1997; Niyogi, 2000; hereafter NB) have developed a model of linguistic evolution based on a macro-evolutionary model in which languages are treated as dynamical systems, the aggregate output of a population of grammars, and evolution of the system corresponds to changes in the distribution of (variant) grammars. This distribution changes as each new generation of language learners each acquire a grammar from the data provided by their speech community (i.e. the previous generation of learners).

The NB model has three main components: a class of grammars, G, from which a learner selects on the basis of data; a learning algorithm, A, used by the learner to choose a grammar, $g \in G$; and a probability distribution, P, with which sentences are presented to the learner. P is defined in terms of the distribution on sentences for each $g \in G$ and the proportions of each g in the population. A dynamical system can now be defined in which each state of the system is represented by a P for state, S_i, and P' for state S_{i+1} can be calculated by an update rule which depends only on P, A and G. Crucially, this deterministic update rule relies on the assumption of non-overlapping generations of learners and speakers, and the abstraction to infinite populations. The former assumption makes the analytic calculation of P for each state of the system tractable and the latter abstraction amounts to the assumption that random sampling effects are irrelevant in the calculation of the proportions of learners who converge to specific grammars given P.

A critical result which follows from one instantiation of the NB model is that the evolution of linguistic systems will be S-shaped or logistic. NB argue that it is a strength of their model that logistic behaviour can be derived analytically from the properties of the update rule, given certain assumptions about G, A and P. Diachronic linguistic work has shown that language change often follows a broadly S-shaped pattern but has not been able to derive this behaviour from more fundamental assumptions (e.g. Lightfoot, 1999:101f). To derive the logistic map, NB assume a two grammar / language system in which A selects between g^1 and g^2 on the basis of 2 example sentences from P. If the last sentence is unambiguously from one grammar, then this grammar is selected. If the first sentence is unambiguously from one grammar and the last is ambiguous, then the learner selects the grammar on the basis of the first example. Otherwise, a random unbiased selection is made. The update rule is defined in terms of the consequent probabilities of A selecting g^1 or g^2 given P. If these probabilities are not equal then the population will converge logistically to the better represented grammar over time. If they are equal then the system is stable and does not evolve.

The critical assumption for the analytic derivation of logistic behaviour lies not in the specific assumptions about *G*, *A* or *P*, but rather in *D*, the model of a dynamical system that NB adopt. (This is not to say that *G*, *P* and particularly *A* are not important - Robert Clark (1996) demonstrates via simulation that logistic change is the exception rather than rule in the NB model, and NB only derive this behaviour analytically for the specific case of selecting between g^1 and g^2 from 2 examples.) NB characterise the states of the system in terms of the proportion of average or arbitrary learners exposed to *P* who converge to g^1 (equivalently g^2). This is a macro-evolutionary model in which what is modelled is the gross statistical behaviour of learners and thus of the linguistic systems, rather than the behaviour of individual learners within the population.

A simple stochastic model

If we replace *D* with a stochastic micro-evolutionary model, *D'* in which there is a finite population of non-overlapping generations, and we model the behaviour of each individual learner while keeping assumptions about *P*, *A* and *G* identical, we find very different behaviour - at least until population sizes become very large. The differences are most obvious when we consider the case where each learner has an equal chance of being exposed to an unambiguous sentence from g^1 or g^2. In the NB model this leads to stasis, but in a micro model stasis is extremely improbable.

For simplicity assume a starting point in which there are equal numbers of g^1 and g^2 speakers in the population, ½ of sentences from g^1 and g^2 can distinguish the two grammars (i.e. are unambiguous with respect to the source grammar which generated them), and *P* is a uniform distribution. The probability that a learner selecting a grammar based on 2 sentences will select on the basis of an initial random unbiased setting, because the two sentences are ambiguous, is ¼, because for each independently drawn observation from *P* the chance of seeing an ambiguous sentence is ½. Therefore, the learner will select g^1 on the basis of data with probability $^3\!/_8$ ($P = 0.375$). (NB give equations for calculating such probabilities for *A*.)

For stasis we require exactly half of the learners to acquire g^1. Suppose there are 100 learners; what is the probability that exactly half will select g^1 in the first generation? The data provided to each learner is stochastically independent so this is equivalent to asking how probable is it that in 100 tosses of an unbiased coin exactly 50 will come up heads, and is given by the binomial theorem: $P = 0.0795$ (e.g. McColl, 1995). Therefore, it is very improbable that the distribution *P* will remain unaltered, and unbiased between g^1 and g^2, for the next generation of learners. This result is in marked contrast from that of NB and follows directly from modelling the fact that each individual learner will be exposed to a different (random) sample of sentences.

To see how likely it is that, given a biased distribution, *P*, on g^1 and g^2, the dominant grammar will spread logistically through the population given *D'*, we need to consider the shape of the skewed binomial distribution arising from the bias. For example, if we minimally modify the example above by assuming that ¾ of the adult population speak g^1, the probability that a learner will acquire g^1 given 2 sentences is now $^{11}\!/_{16}$ ($P = 0.687$). (Note that it is not $^{12}\!/_{16}$ ($P = 0.75$) because of the possibility of selection according to the initial unbiased setting when the data seen is ambiguous.) Consequently, the probability that more than 75 learners will acquire g^1 is only $P = 0.070$, though the probability that more than 50 will acquire g^1 is $P > 0.999$. In fact, the distribution peaks at 69 learners predicting not logistic growth but rather a probable slight decline in the number of g^1 speakers in the next generation. In the limit, if the whole population speak g^1, the probability that a learner will select g^1 is 7/8 ($P = 0.875$) because there remains a 1/8 chance that a learner will see 2 ambiguous sentences and select g^2 on the basis of a random initial setting. Therefore, even in this case the resulting binomial distribution peaks with 88 learners acquiring g^1 in the next generation. Therefore, given these assumptions for *P*, the micro model predicts endless random drift in the proportions of g^1 and g^2 speakers.

It might be objected that this result follows primarily from choosing *G* and *P* with a high proportion of ambiguous sentences, so that learners frequently select grammars on the basis of random (unbiased) initial settings (though *G* and *P* here are in this respect similar to several of the more realistic examples NB consider, derived from Gibson and Wexler, 1994). If we assume, that g^1 and g^2 are as highly differentiated as possible and share no sentences, then a learner will select between them with probability directly correlated with the proportions of g^1 and g^2 speakers in the adult population. In the case of equal proportions, the probability that exactly half the population of learners will acquire g^1 (equivalently g^2) is still given by the unbiased binomial distribution in D', and thus remains low. The binomial distributions for each generation of learners will now peak at exactly the point predicted by the proportion of adult speakers, but this still only allows us to predict that ± 13 learners around this peak will acquire g^1 with $P > 0.99$ for a population of 100 learners. Therefore, we can still expect to see an oscillating pattern of random drift prior to eventual on one variant.

Realistic stochastic models

For some types of language change the idealisation of *D* to infinite populations may not be harmful; for example, diffusion through American English within the last 50 years might be such a case. However, even then we would need to be clear that there is an analytic and thus predictive advantage to macro modelling for *realistic* versions of *G*, *A* and *P*, and this has not been demonstrated as yet. In all cases where evolution of a linguistic system is likely to have taken place in small relatively isolated speech communities - for example, modelling of prehistoric development or of a process like creolisation, where the relevant populations are likely to have been at most in the low hundreds - abstracting away from sampling issues is dangerous.

Furthermore, the specific behaviour which we want to derive, such as logistic change in the system, may simply follow directly from more realistic demographic assumptions than are possible with macro models. For example, population movement, birthrate, the proportion of language learners in the population and the resultant linguistic mix of the population are critical factors in understanding creolisation. Briscoe (2000) discusses several models enriching *D*' with more realistic demography and realistic accounts of *A* and *G*.

References

Briscoe, Ted (2000) `Evolutionary perspectives on diachronic syntax' in Susan Pintzuk, George Tsoulas and Anthony Warner (ed.), *Diachronic Syntax: Models and Mechanisms*, Oxford: Oxford University Press.

Clark, R.A.J. (1996) *Internal and External Factors Affecting language Change*: *A Computational Model*, MSc Dissertation, University of Edinburgh.

Gibson, E. and Wexler, K. (1994) `Triggers', *Linguistic Inquiry, vol.25.3*, 407-454.

Lightfoot, D. (1999) *The Development of Language: Acquisition, Change, and Evolution*, Blackwell, Oxford.

McColl, J.H. (1995) *Probability*, Edward Arnold, London.

Niyogi, P. (2000, in press) `Theories of Cultural Change and their Application to Language Evolution' in Briscoe, E.J. (ed.), *Language Acquisition and Linguistic Evolution: Formal and Computational Approaches*, Cambridge University Press, Cambridge.

Niyogi, P. and Berwick, R. (697-719) `Evolutionary consequences of language learning', *Linguistics and Philosophy, vol.20*. 1997

Renshaw, E. (1991) *Modelling Biological Populations in Space and Time*, Cambridge University Press, Cambridge.

Something to talk about: Conflict and coincidence of interest in the evolution of shared meaning

Seth BULLOCK

School of Computer Studies, University of Leeds
seth@scs.leeds.ac.uk

Introduction

If we are seriously to consider the possibility that human language arose from animal communication through a process of evolutionary change, even if such consideration is merely in order to discount this possibility, we must address the conceptual problems at the heart of our current understanding of animal signalling. In doing so we may throw light upon not only the origins of human language, but also its character. In this paper I will identify a problem in understanding the semantics of evolved communication. This problem stems from the fact that whilst the meaning of human language is typically considered to be "shared" across a community of language users, the interests of evolving creatures are often not.

Biologists have been aware of the fact that conflicts of interest may prohibit the establishment and maintenance of honest signalling for some time, and have proposed various attempts to deal with this problem (see Johnstone, 1997, for a recent review of this literature). They have also pointed out that the same conflicts of interest prohibit a straightforward reading of meaning in such signalling systems (Dawkins & Krebs, 1978; Krebs & Dawkins, 1984; Maynard Smith & Harper, 1995). I will argue that even if the problems concerning how honesty may be established and maintained in an evolving system are solved or avoided, the latter problems concerning the semantics of evolved signalling systems remain. I will suggest that a reanalysis of the issue of conflicting interests points to a solution to these problems.

Stability and meaning

As group-selectionist accounts lost currency in evolutionary biology, those interested in the evolution of animal signalling confronted the fact that since the interests of different animals often conflicted, it was not clear why their signals should be honest, and as a result not clear why their signals should be attended to. Despite this, it appeared that signalling was near ubiquitous across the natural world.

Commentators have presented this issue of conflicting interests as essentially dichotomous (e.g., Maynard Smith & Harper, 1995). Either one is like a bee, living in a community enjoying entirely homogeneous interests due to its peculiar genetics, or one lives in a world of conflict stemming from the struggle to out-reproduce one's competitors. Whilst the adaptiveness of communication is trivially apparent in the former case, how on earth could stable communication survive the tumult of competition implied by the latter?

Faced with the task of building honest signalling systems from deceitful beasts, biologists have taken two approaches. Some have claimed that the notion of stable honest signalling between agents with conflicting interests is a myth. Under this reading, most natural signalling systems are in a constant state of flux – signallers attempting to *manipulate* receivers to their own ends, whilst receivers attempt to *mind-read* the secret thoughts and intentions of signallers

(Dawkins & Krebs, 1978; Krebs & Dawkins, 1984). Others have sought to shoe-horn competitive animals into co-operative harmony by identifying restraints on signallers (or receivers) of such gravity that their self-interest can only be served by kow-towing to the yoke of receiver (or signaller) interests. The costs of signalling, for instance, may preclude deceitfulness that would otherwise have been adaptive (Zahavi & Zahavi, 1997).

Both of these approaches have problematic consequences for the semantics of signalling systems. The gravest of these consequences is the possibility that the behaviours picked out by the theories do not count as attending to, or producing, signals at all. If a vulture spots some carrion which is as yet undetected by its conspecifics (with whom it competes for food), and descends from the sky to devour it, can this act of descending, as seen by other vultures, be taken as a signal from the first vulture to its competitors to the effect that “there is some food to be eaten over here”? Similarly, consider a species that likes to swim towards magnetic north. If an aquatic predator could generate an electro-magnetic field about itself that attracted its prey by mimicking magnetic north, in what sense is this magnetic field acting as a signal?

Millikan (1984, 1993) has claimed that neither of the above examples qualify as instances of signalling. The behaviour of one party is merely adapted to exploit the character of the other. No more meaning can be attached to these manipulations or mind-readings than to the pushing or pulling of animals attempting to manually manipulate each other with specially adapted pincers. In contrast, for a system to be a signalling system, Millikan demands that both agents involved have been adapted by natural selection to play their role in a co-operative signalling episode. Millikan uses the bee dance as a paradigmatic example of such a signalling system. The interests of the dancing bees and their audience are coincident. This coincidence is one way to guarantee that the system is evolutionarily stable, and is the only way to ensure that it counts as a *signalling* system and hence involves *meaningful signals*.

Millikan’s account implies that systems may be evolutionarily stable and may even be understood to involve the traffic of information, yet still not count as signalling systems. Biologists’ efforts to stabilise natural signalling systems using restraints on the agents involved would, under this reading, not result in stable *signalling* systems, as the conflict of interest at their heart would prevent them meeting Millikan’s criteria. For instance, sensory exploitation, in which a “signaller” exploits a tendency in an observer to respond to a particular stimulus (Ryan, 1990), may be evolutionarily stable if the observer’s tendency is so valuable in some other context that it is not worthwhile extinguishing, despite the cost of being exploited. For example, the magnetic-north-seeking organism may continue to seek magnetic north despite the risk of predation if this is a successful foraging strategy. Although evolutionarily stable, this system still suffers the problems identified by Millikan. What is being signalled by the exploiter? What is the meaning of the exploitational behaviour? Similarly, a handicap signalling scenario (Zahavi & Zahavi, 1997) in which, for instance, only some males can afford to exhibit a sexual ornament, will not count straightforwardly as a signalling system, since the signallers and receivers involved have not coevolved to achieve a common aim. Rather, they are fundamentally at odds – signallers being selected to obtain as many mates as possible, while receivers are selected to get the best mates possible. The meaning of the mating display is not shared by signaller and receiver since the job of the sexual ornament is not the same for both parties.

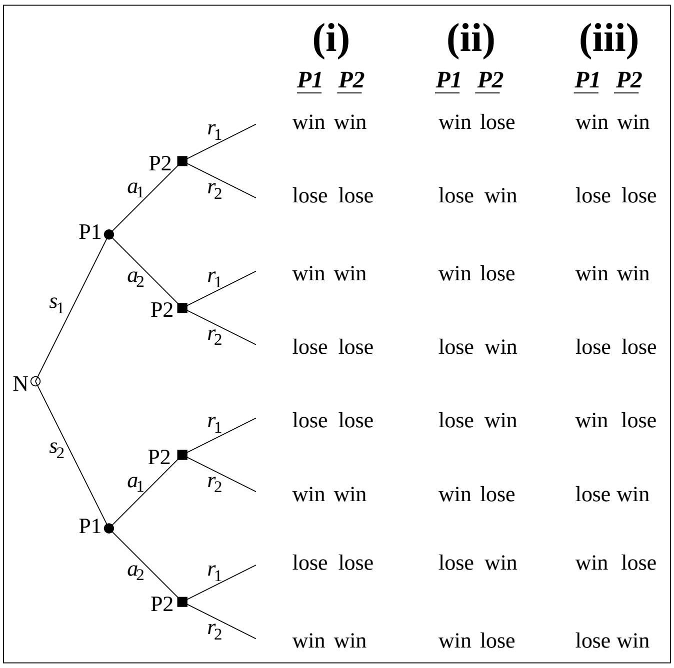
Human language

Within models of the evolution of human language (e.g., Kirby & Hurford, 1997; Kirby 1998, in press; Batali, 1998), there is often little attention to the possibility of deceit (but see Noble, in press). These accounts either implicitly or explicitly presuppose that the agents involved are taking part in an essentially co-operative enterprise. Language is used by these agents to transmit information. Disregarding some perspectival differences, the information conveyed in an utterance is shared between speaker and listener. Can we justify this conception of human language given that (i) the semantics of animal signalling are threatened by the conflicts of interest they suffer, (ii) this threat is not extinguished when such signalling systems are stabilised by, for instance, signaller costs, and (iii) human signalling evolved from animal signalling. If we cannot square these three assertions, must we abandon continuity between animal signalling and human language?

Conflict and coincidence

Figure 1 (overleaf) presents three simple action-response games (e.g., Bullock, 1997). In each game there are two possible states of the world, $S \in \{s_1 , s_2\}$. Player 1 (P1) is aware of the state of the world and must make one of two possible actions, $A \gamma \{a_1 , a_2\}$. Player 2 (P2) sees the action of P1, but is unaware of the state of the world. She must make one of two possible responses, $R \in \{r_1 , r_2\}$. In each game the state of the world makes a difference to at least one of the players in terms of which outcomes (S , A , R) they prefer.

In the first game, P1 and P2 experience a coincidence of interest. Whenever P1 wins, P2 does also. Whenever P1 loses, so does P2. In the second game the players suffer a complete conflict of interest. Whenever P1 wins, P2 loses, and vice versa. The potential for honest signalling (defined, *sensu* Enquist (1985), as *A* dependent on *S*, and *R* dependent on *A*, at equilibrium) in these two games is clear – in the first game the players have everything to talk about, in the second, nothing.



(i)		(ii)		(iii)	
P1	*P2*	*P1*	*P2*	*P1*	*P2*
win	win	win	lose	win	win
lose	lose	lose	win	lose	lose
win	win	win	lose	win	win
lose	lose	lose	win	lose	lose
lose	lose	lose	win	win	lose
win	win	win	lose	lose	win
lose	lose	lose	win	win	lose
win	win	win	lose	lose	win

The first game is capable of supporting a *co-operative signalling system* – one that satisfies Millikan's criteria as outlined above. Since the interests of the two parties coincide, P1 can, and should, use his action as a signal conveying the state of the world to P2. The meaning of this signal is shared by both players – it is both produced and consumed as a representation of *S*. In the second game P1's action cannot serve any purpose. In fact, P1 should not allow his action to in any way predict the state of the world, since this correlation could be exploited to his disadvantage by P2. The third game is more interesting precisely because the potential for honest signalling is not immediately clear – do the players have something to talk about? In this game there is a *partial* conflict of interest, or equivalently a partial coincidence of interest. Given that s_1 obtains, the interests of P1 and P2 coincide – they both prefer the outcome (s_1 , A , r_1) over (s_1 , A , r_2). Given that s_2 obtains, the interests of P1 and P2 conflict – P1 prefers (s_2 , A , r_1), while player 2 prefers (s_2 , A , r_2).

In fact, when biologists discuss signalling between agents suffering a conflict of interest, they are always referring to games of this form. Sexual signalling, aggressive displays, etc., all involve the potential for shared interests (when both parties would prefer the same outcome) but this is not guaranteed (the suitor may be poor, the aggressor weak, etc.). The former aspect ensures the possibility of a signalling system, the latter undermines this possibility. Thus, the interests of evolving agents are not best characterised as either coinciding (game 1) *or* conflicting (game 2). An important intermediate scenario (game 3) exists.

Given this partial coincidence of interest, can we argue that the behaviour of the players at equilibrium is sometimes best characterised by the notion of a co-operative signalling system being parasitised? For example, in the same way that an *eavesdropper* might gain knowledge of S by spying on P1, or a *fraud* might manipulate the behaviour of P2 by mimicking P1, the behaviour of the players, given that s_2 obtains, might be parasitic upon that of themselves given that s_1 obtains. Can we separate the two halves of the game, and claim that the co-operative half is a signalling system in its own right, despite the threat to it posed by the parasitic half?

Assume P2 plays the strategy ($r_1 \Rightarrow a_1$, $r_2 \Rightarrow a_2$). Given that s_2 obtains, P1 can either lie (a_1) or admit (a_2). Neither action can be awarded the status of signal. In the former case, P1 manipulates P2 to his own benefit (he lies that he is a high quality suitor, a strong aggressor, a poisonous prey item, etc.). In the latter case, P2 mind-reads P1 to her own benefit (P1 admits that he is a poor suitor, or a weak aggressor, or a palatable prey item). External forces (e.g., signalling costs) are required to make it reasonable for P1 to make this admission, just as external forces (e.g., a valuable sensory bias) are required to make it reasonable for P2 to believe a lie. These forces might stabilise honest signalling, but would fail to give the status of signalling episode to either the sequence $s_2 \Rightarrow a_2 \Rightarrow r_2$ or $s_2 \Rightarrow a_1 \Rightarrow r_1$.

This analysis suggests that the relationship of P1's actions to the interests of both parties confers upon them one of three possible statuses which may be glossed as signal, lie, or admission. Only the first of these is representational, but the meaning involved is meaning that is shared between signaller and receiver.

Conclusion

What impact does this reanalysis have on our understanding of the semantics of animal signalling, and the evolution of language? First, it allows that there may be a valid way to talk about signals between agents lacking entirely coincident interests, which preserves the idea of shared meaning. In the same way that the bee dance is both produced and understood as a representation of the location of nectar, signals between prey and predator, suitor and prospect, attacker and defender, offspring and parent, etc., may sometimes represent prey palatability, mate quality, etc. Second, this potential for shared meaning to exist allows that our conception of human language as involving meaning which is shared by a community is commensurable with the notion that there is evolutionary continuity between animal signalling and human language.

References

Batali, J. (1998). Computational simulations of the emergence of grammar. In Hurford et al. (1998).

Bullock, S. (1997). An exploration of signalling behaviour by both analytic and simulation means for both discrete and continuous models. In Husbands & Harvey (1997), pp. 454-463.

Dawkins, R., & Krebs, J. R. (1978). Animal signals: Information or manipulation. In J. R. Krebs, & N. B. Davies (Eds.), *Behavioural Ecology: An Evolutionary Approach*, pp. 282-309. Blackwell, Oxford.

Enquist, M. (1985). Communication during aggressive interactions with particular reference to variation in choice of behaviour. *Animal Behaviour*, *33*, 1152-1161.

Hurford, J. R., Knight, C., & Studdert-Kennedy, M. (Eds.) (1998). *Approaches to the Evolution of Language – Social and Cognitive Bases*. Cambridge Univeristy Press.

Husbands, P., & Harvey, I. (Eds.). (1997). *Proceedings of the Fourth European Conference on Artificial Life*, MIT Press, Cambridge, MA.

Johnstone, R. A. (1997). The evolution of animal signals. In Krebs, J. R., & Davies, N. B. (Eds.), *Behavioural Ecology: An Evolutionary Approach* (4th edition), pp. 155-178. Blackwell, Oxford.

Kirby, S. (1998). Fitness and the selective adaptation of language. In Hurford et al.(1998).

Kirby, S. (in press). Syntax without natural selection: How compositionality emerges from vocabulary in a population of learners. In Knight et al. (in press).

Kirby, S., & Hurford, J. (1997). Learning, culture and evolution in the origin of linguistic constraints. In Husbands & Harvey (1997), pp. 493-502.

Knight, C., Hurford, J. R., & Studdert-Kennedy, M. (Eds.). (in press*). The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form*. Cambridge Univeristy Press.

Krebs, J. R., & Dawkins, R. (1984). Animal signals: Mind reading and manipulation. In J. R. Krebs, & N. B. Davies (Eds.), *Behavioural Ecology: An Evolutionary Approach* (2nd edition), Blackwell, Oxford.

Maynard Smith, J., & Harper, D. G. C. (1995). Animal signals: Models and terminology. *Journal of Theoretical Biology*, *177*, 305-311.

Millikan, R. G. (1984). *Language, Thought and Other Biological Categories*. MIT Press, Cambridge, MA.

Millikan, R. G. (1993). *White Queen Psychology and Other Essays for Alice*. MIT Press, Cambridge, MA.

Noble, J. (in press). Co-operation, competition and the evolution of pre-linguistic communication. In Knight et al. (in press).

Ryan, M. J. (1990). Sexual selection, sensory systems, and sensory exploitation. *Oxford Survey of Evolutionary Biology*, *7*, 157-195.

Zahavi, A., & Zahavi, A. (1997). *The Handicap Principle: A Missing Piece of Darwin's Puzzle*. Oxford University Press.

The slow growth of language in children

Robbins BURLING

University of Michigan
rburling@umich.edu

ABSTRACT

A surprisingly broad consensus has grown among scholars interested in the origin of language that the ability to speak is a relatively recent and quite sudden development. One support for this presumption is the difficulty some linguists have imagining a partial syntax. No part of syntax, it is supposed, could exist without all the rest. Among other possible candidates for partial syntax is child language, so if one is to deny the possibility of partial syntax one must interpret child language as jumping abruptly from a state where it consists of little more than strings of poorly joined individual words (or "proto-language") to a stage of full syntax.

Unfortunately for those who hold this position, the facts of child language, as we can observe it each day in the behavior of our own children, do not fit this abrupt scenario. While it is true that children frequently appear to achieve complex syntax with enviable speed, they do not move from proto-language to full syntax in a few months, let alone in a single day. Two factors give us an illusion of much greater speed in language learning than the evidence can support.

First, an important part of language learning takes place silently, well before children actually produce the forms they have learned. It would be reckless to ignore the confidence of all parents that their children understand far more than they can say. Their level of understanding during the period before they use multiword sentences points to the priority of comprehension. Producing a syntactic construction should be looked upon as only the final stage in a long developmental process.

Second, syntactic learning continues well beyond the age of five. The illusion of the mature five-year-old is probably due to the fact that children of this age have generally overcome the imperfect phonology of their earlier years and no longer make the obvious morphological errors of younger children. The speak with few obvious mistakes, but they avoid syntactic mistakes by the simple expedient of avoiding complex syntactic constructions, not by knowing all there is to know about syntax.

If syntax starts to be learned months before it appears in production, and if it continues to be learned well past the age of five, its acquisition is not as magically fast as it sometimes seems. Partial syntax of many levels of complexity is there to be observed in our own children. If partial syntax is so evident in our own children, I see no reason to doubt the possibility of partial syntax, of many levels of complexity, among our early ancestors.

I do *not* mean to imply that ontogeny recapitulates phylogeny. The slow growth of language in children does *not* require us to believe in the slow development of language in prehistory. I mean only to insist that one cannot use the evidence of children to support the opposite claim, the odd notion that syntax *had* to develop abruptly in our ancestors because a partial syntax is impossible.

Corticocortical coherence as an enabling step for fluent nested embedding

William H. CALVIN

Department of Psychiatry and Behavioral Sciences
University of Washington
Seattle, Washington 98195-1800 USA
wcalvin@u.washington.edu
http://faculty.washington.edu/wcalvin

In our book[1], *Lingua ex Machina*, Derek Bickerton and I offer three possible preadaptations for the big step up from protolanguage to syntax:

1. The cognitive categories needed for keeping track of "Who owes what to whom" (evolved for minimizing the freeloader problem in reciprocal altruism) are a nice setup for the agent-theme-goal role tags so useful for quickly communicating "Who did what to whom."
2. The prefrontal-premotor circuitry needed for planning novel throwing and hammering movements looks a lot like the binary trees of phrase structure.

The repeated payoffs of even more altruism, or even more accurate throwing, seem capable of adapting brain circuitry that could be occasionally borrowed for structured language.

3. The cortical "plainchant choirs" needed for preprogramming accurate throws can, when extensive enough, make the long corticocortical paths (corpus callosum and such) temporarily coherent, overcoming the usual blur and jumble. Coherence would permit the various phrases and clauses to be located at diverse cortical sites, yet still "sing as a unified choir" as alternative sentence interpretations compete with one another.

The threshold for coherence provides a "capstone" candidate for fluent structured thought and talk – something that might have triggered the flowering of art and technology seen late in hominid evolution, after brain size itself had stopped growing.

Since corticocortical coherence is the one that builds atop the other two, and is the one that might have had a sudden emergence (the others look like gradual carryover so far, perhaps for millions of years), let me start by briefly explaining the first two pre-syntax "stones" of this arch so that the "capstone" role in enabling *fluent* syntax can be appreciated.

Antecedents of argument structure

Bickerton's preadaptation is from reciprocal altruism. What pays for the improvements in abstract cognitive categories for actor, recipient, and action are that they help keep track of who owes what to whom. This is not only handy for minimizing the cheater problem, telling the individuals possessing such cognitive traits when they ought to find a partner more likely to reciprocate occasionally, but it aids the furtherance of coalitions between nonrelatives.

---

[1] William H. Calvin and Derek Bickerton, Lingua ex Machina: Reconciling Darwin and Chomsky with the Human Brain (MIT Press, 2000). Available on the web at http://faculty.washington.edu/wcalvin/LEM.

Bickerton's idea is that, once you have the cognitive categories for keeping track of "Who owes what to whom," you can start communicating "Who did what to whom" – and have a syntax-ready recipient with the preformed mental categories needed for disambiguating the longer utterances. You get the essentials of argument structure (categories for agent, goal, and theme) from the payoffs of ever more cooperation. You may further improve the underlying neuroanatomy with new payoffs from language itself, but the foundations came "free," paid for via the prior cooperation/altruism use.

Multifunctional cortex

Conversions of function are a familiar story in evolutionary theory (Darwin spoke of them before introducing his example of the fish swim bladder converting into an amphibian lung). And any conversion goes through an intermediate phase of multifunctionality.

Despite our mapmaking tendencies that lead us into erroneously assuming one region, one function (it's another instance of the reification fallacy), cortical regions are notoriously multifunctional. While a region may have a specialization that is essential (you can't do the function without it, as seen in the stroke and cortical surface stimulation results), the region may participate in other functions (as seen via its increase in blood flow when performing nonspecialist tasks). Even the common strokes illustrate the multifunctionality: most aphasic patients also suffer from hand-arm apraxia, suggesting a core of neural machinery for novel sequencing that is shared by novel hand-arm and oral-facial movements.

Antecedents of phrase structure

The other common way of disambiguating a long string of words is what, in pre-minimalist days, was called phrase structure. The nested embedding of phrases and clauses, one within another, is strongly reminiscent of the structured planning needed for planning multi-joint ballistic movements during "get set." The hand movement is embedded in the elbow rotation which is embedded in the shoulder rotation which is embedded in the body's forward motion. Planning involves getting the whole thing right; there may be dozens of ways to hit the target, but they're hidden in a sea of millions of wrong solutions, ones that would cause dinner to run away.

What pays for planning improvements is aimed throwing for hunting and precision hammering for toolmaking; when novel sequences of hand-arm movement are not being planned, such neural machinery may be available for structured planning of other things, such as long utterances or agendas. Again, once structured language pays for further improvements, it may in passing improve aimed throwing or precision hammering, to the extent that they continue to share neural machinery.

The important feature of ballistic planning is its use of cortex for *novel* sequences (i.e., not standardized target distances as in darts and basketball free-throws, where subcortical circuits probably take over) with demanding requirements for timing precision (small targets have very brief "launch windows" and so timing jitter must be minimized).

Corticocortical coherence for fluent structured thought

I recently refined Hebb's 1949 notion of a cell-assembly[1] whose spatiotemporal firing pattern (think of a short song) represents a concept, relationship, action – or a phrase or a clause. Even the complete sentence should have one, if it is to compete with alternatives. This firing pattern would be a "code" for the concept, but one would expect the code for, say, *comb* to be different in visual association cortex than near auditory or motor areas.

A uniform-across-the-cortex code would be powerful but, for the same reasons as it took Europeans so long to invent the Euro, it isn't a default solution. The corticocortical connection from concept-laden temporal lobe to movement-schema-laden frontal lobe via the arcuate fasciculus (second only the corpus callosum in size) is surely jumbled (neighboring axons may not remain neighbors at the other end) and the fanout of connections guarantees blur. So a different code arrives, equally good in most respects (just like changing money when border crossing) but when it gets sent back to the originating cortex, it is doubly distorted. For frequently used concepts, of course, an identity relation can soon be learned.

But when novel messages are being sent around ("'A square green tomato' – anyone recognize this phrase?"), it runs into the same problem as trying to get moneychangers to recognize a wooden nickel. You can't get the virtues of a universal code for novel on-the-fly concepts without the equivalent of a coherent corticocortical connection.

What I demonstrated[2] in Chapter 7 of *The Cerebral Code* was that enough clones of the code in the sending cortex (think of a plainchant choir recruiting additional members from neighboring cortex, all singing the same song) would suffice to create a small choir singing the same song at the destination. (Such choirs are needed for reducing jitter and are also an outcome of a Darwinian copying competition for achieving quality on-the-fly.)

This allows novel codes to be passed back and forth without slowly learned identities. While handy for association tasks in general, the role of corticocortical coherence in nested embedding is where the common code really shines. It seems capable of making syntax an everyday, subconscious task that operates in seconds.

The "meaning of the sentence" is, in this model, an abstract cerebral code which competes for territory with codes for alternative interpretations, often in the manner of a Darwinian cloning competition. Phrases and clauses require coherent corticocortical links to contributing territories, having their own competitions and tendencies to die out if not reinforced by backprojecting codes.

It starts to look like a choral work of many voices, each singing a different tune but with the requirement that it mesh well with all the others. Indeed, the symphonic metaphor might be appropriate for the more complex sentences that we can generate and understand. Certainly the reverse-order analogy to Benjamin Britten's *Young Person's Guide to the Orchestra*, the all-together version being succeeded by the various voices playing separately, is the best metaphor I know for the read-out process that converts the parallel-structured plan into serial-ordered speech.

---

[1] D. O. Hebb, The Organization of Behavior (Wiley 1949), p.62.

[2] William H. Calvin, The Cerebral Code (MIT Press 1996).
Available on the web at http://faculty.washington.edu/wcalvin/bk9/.

Degrading syntax back to protolanguage

Consider the implications of efficiently linking the concept-filled temporal lobe with the prepare-for-action frontal lobe, with a common code replacing the degenerate codes – and then dropping back to the old system, with now-incoherent paths forcing a reliance on slowly established identities. Without coherence, you'd still have a vocabulary (the temporal lobe still works). You'd still be able to plan some nonlanguage actions (you'd pass many of the neuropsychological tests for frontal lobe functioning), but your ability to quickly invent new trial run associations would suffer.

Not only couldn't you form up a syntactic sentence to speak (except for stock phrases), but you couldn't judge sentences that you heard someone else speak because you could no longer judge the quality of your trial interpretations, whether they were nonsense, good guesses, or sure things. Your quality associations would be too slow for the windows of opportunity, and the results would be of poor quality because not shaped up very far by Darwinian copying competitions in the brain. And so your performance on language tasks would drop back to something like protolanguage, a wide choice of words but with novel sentences limited to just a few words to avoid ambiguity.

The big step up: Homo Sapiens and structured thought

Carryover from reciprocal altruism's cognitive categories and ballistic movement's planning circuits are both compatible with slow language improvement over a few million years. But corticocortical coherence should have a threshold (the size of the plainchant choir, achieved mostly by temporarily borrowing neighboring cortex, not via brain size increases).

Once borrowing abilities cross the coherence threshold, structured thought and talk would have become far more fluent – and thus a capstone candidate for what triggered the flowering of art and technology seen late in hominid evolution, after brain size itself had stopped growing.

A proper *lingua ex machina* would be a language machine capable of nesting phrases and clauses inside one another, complete with evolutionary pedigree. Such circuitry for structured thought might also facilitate creative shaping up of quality (figuring out what to do with the leftovers in the refrigerator), contingency planning, procedural games, logic, and even music. And enhancing structured thought might give intelligence a big boost. Solve the cerebral circuitry for syntax, and you might solve them all. None of us are there yet, but this "three-stone archway" provides an illustration of how our big questions – what, how, and why – might hang together.

I thank Derek Bickerton for persistently steering me in this direction, and the Rockefeller and Mathers Foundations for provocative venues.

Evolution of symbolisation in chimpanzees and neural nets

Angelo CANGELOSI

Centre for Neural and Adaptive Systems
University of Plymouth (UK)
a.cangelosi@plymouth.ac.uk

Introduction

Animal communication systems and human languages can be characterised by the type of cognitive abilities that are required. If we consider the main semiotic distinction between communication using icons, signals, or symbols (Peirce, 1955; Harnad, 1990; Deacon, 1997) we can identify different cognitive loads for each type of reference. The use and understanding of icons require instinctive behaviour (e.g. emotions) or simple perceptual processes (e.g. visual similarities between an icon and its meaning). Communication systems that use signals are characterised by referential associations between objects and visual or auditory signals. They require the cognitive ability to learn stimulus associations, such as in conditional learning. Symbols have double associations. Initially, symbolic systems require the establishment of associations between signals and objects. Secondly, other types of relationships are learned between the signals themselves. The use of rule for the logical combination of symbols is an example of symbolic relationship. Symbolisation is the ability to acquire and handle symbols and symbolic relationships.

Symbolisation in chimpanzees

A great deal of research exists regarding the study of symbolisation in humans. Language is considered to be a prototypical example of human ability to learn and use symbols. However, when we look at the evolutionary roots of symbolisation and language, e.g. with animal experiments, many studies have investigated the general ability of different animal species to acquire human-like languages, as opposed to focusing on symbolisation. Some experiments on language acquisition in chimpanzees have specifically investigated the evolution of symbolisation in apes (Savage-Rumbaugh, 1986). In these studies, researchers made a clear and operational distinction between non-symbolic and real-symbolic language learning strategies. Non-symbolic linguistic strategies use simple conditional associations to link signals and objects. Alternatively, real symbolic languages are based on the acquisition of symbolic relationships for communication, and the decontextualisation of language from the restricted learning stimulus set. In Savage-Rumbaugh & Rumbaugh (1978) chimpanzees are trained to learn a set of lexigrams (pictures in a keypad) to communicate about foods and drinks. Animals first learn the lexigram of the individual foods and drinks, such as banana and orange, milk and coke. Subsequently, they are taught the lexigrams for two actions (“pour” for the drinks only, “give” for the foods only), together with the individual food/drink lexigram (e.g. “pour-milk”, “give-banana”). Animals successfully learned these lexigrams after a systematic training cycle. Savage-Rumbaugh & Rumbaugh also devised a test for symbolisation. They wanted to ascertain if the linguistic stimuli learned by the animals were used in a real symbolic way (e.g. identifying the logical rule to associate the lexigram “pour” with all drinks, but not any of the solid foods) or if the animals were simply associating the whole pair “pour-milk” to the event of pouring milk. They taught the chimpanzees the lexigrams for the names of new foods and drinks and checked if the animal was able to generalise the rule and associate the correct action lexigrams with the new name lexigrams. The test results showed that only some of the chimpanzees were able to make a correct rule generalisation. Other chimpanzees had to be retrained to learn the new pairs of action-name lexigrams. Savage-Rumbaugh et. al. (1980) presented similar results for a test on the use of lexigrams for classifying tools and foods. Other studies (Greenfield & savage-Rumbaugh, 1990) have shown that during the spontaneous learning of lexigram use in baby chimpanzees, some animals invented symbolic structures, such as the one resembling the “action-object” syntactic rule.

This experimental data suggests that apes can successfully learn symbolic relationships. However, this learning is only obtained under certain experimental conditions that, for example, stress the pragmatic aspects of communication during language acquisition. These experiments are lengthy and complex, but they are useful in the acquisition of symbolisation abilities in apes. They also indicate that animals can use symbols in ways that emulate human language without comprehending their representational function (Savage-Rumbaugh et. al., 1980). For Deacon (1997) this evidence contributes to the explanation of the gap between animal communication systems and human language. Deacon also suggests that animals, even apes, have great difficulties in learning symbolic relationships because of differences in the structure and function of their brain, in particular in the prefrontal cortex areas.

Computational models for symbolisation: Evolving neural networks

Artificial neural nets are computational models that are inspired by the function and structure of biological neural systems. Currently, they are used for modelling cognition. However, using neural net models to study symbol acquisition is still a controversial subject. Some researchers are very sceptical (e.g., Fodor & Pylyshyn, 1988; Marcus, 1998), whilst others support the use of neural nets for cognitive tasks requiring symbolisation (e.g., Rumelhart & McClelland, 1986). Recently, the integration of genetic algorithms as a model of evolution and neural nets for cognitive modelling, has been proposed for the study of the evolution of communication in populations of artificial organisms (Cangelosi & Harnad, in press). This method is part of the synthetic approach to the modelling of language evolution (Steels, 1997; Kirby, 1999). This paper uses a model of the evolution of communication (Cangelosi, 1999) to study symbolisation and symbol acquisition in neural nets. Such nets represent the organisms' cognitive systems that control behaviour and communication. The simulated tasks resemble that of Savage-Rumbaugh & Rumbaugh's (1978) chimpanzee experiments. This work proposes a complementary approach for the study of the evolution of symbolisation through computational modelling. Computer models will allow us to simulate and expand the experimental settings used in lengthy animal experiments. For example, neural nets can be used to test some of Deacon's (1997) hypotheses on the co-evolution of the brain, language and symbolisation.

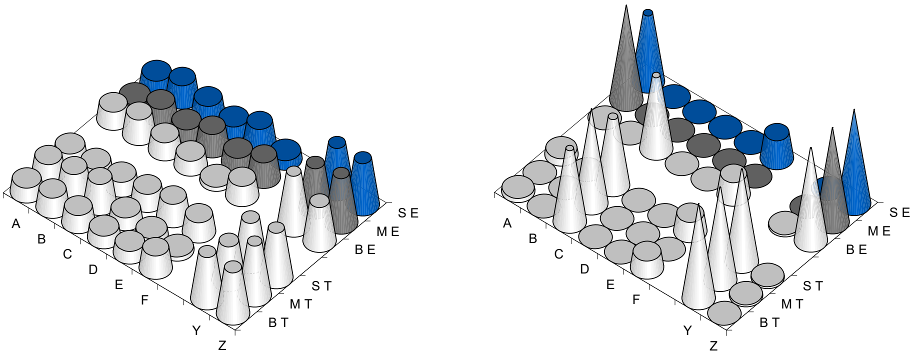
Model setup

The model setup is directly inspired by the ape language experiments. A population of 80 artificial organisms perform a foraging task by collecting edible mushrooms, whilst avoiding poisonous mushrooms (toadstools). The organisation of foraging task stimuli into a hierarchy of functional categories was derived from Savage-Rumbaugh & Rumbaugh's (1978) experiments. Our hierarchy consists of 2 high-level categories (edible and poisonous mushrooms) and 3 low-level categories (large, medium, and small mushrooms). Organisms will learn to name each of the three edible subcategories (“large edible”, “medium edible”, and “small edible”) and a common verb for the high-order edible category, i.e. “approach”. Each of the three toadstool subcategories (“large poisonous”, “medium poisonous”, and “small poisonous”) require the use of the same verb, i.e. “avoid”. The organisms' fitness and reproduction depend upon the number of edible foods correctly collected minus the number of toadstools collected. At each generation the 20 organisms with the highest fitness are selected and asexually reproduce 4 offspring each. The organism’s genotype is the connection weight matrix of its neural net. New offspring are subject to a 10% random mutation of their weights. During the first 300 generations, organisms evolve the ability to discriminate between the 6 types of mushrooms (3 edible and 3 poisonous). From generation 301 organisms are able to communicate using 8 linguistic input/output units to describe mushrooms. Organisms learn to label mushrooms using the backpropagation algorithm. The teaching input is provided from their parents.

A 3-layer feedforward neural net controls the behaviour of the organism. In the input layer 18 units encode the perceptual features of the closest mushroom, 3 units encode its location, and 8 units encode the 8 symbols available for communication. The hidden layer has 5 units. In the output layer 3 units are used to control the organism’s behaviour (movement and action depending on mushroom size), and 8 units are used to produce the communication symbols. Symbolic output units are organised in two winner-takes-all clusters of competitive units (one cluster of 6 units, one of 2).

Results

The simulation of the model was repeated 10 times, starting from different random populations. At generation 300, the fitness in 9 out of 10 replications increased to an optimal level. These 9 successful populations were used to evolve communication from generation 301 to 400. In approximately half of the replications, organisms evolved an optimal lexicon, i.e. the use of at least 4 symbols/symbols-combinations to distinguish 4 types of mushrooms (the toadstools + the three types of edible mushroom) (detailed description of the model's results can be found in Cangelosi, 1999). In the remaining populations, some mushrooms were incorrectly labelled and classified due to the lack of a specific symbol. Note that the majority of successful simulations evolved languages that used combinations of symbols, and in particular some evolved the “verb-noun” structure. Two different “verb” symbols were used respectively for toadstools and edible mushrooms. The other symbol is used to distinguish between the three subcategories of mushrooms. Figure 1 shows the charts of an evolved “verb-noun” language. Note that the two “verb” symbols (“Y” and “Z”) emerge in the early stages of language evolution and then stabilise. The names for the mushroom subcategories are subject to continuous change and only at the last generation they reach a stable and optimal point.



generation 300 generation 400

Figure 1: *Structure of evolved language at generation 300 and generation 400. (Letters A-Z for the 8 available symbols. SE, ME, BE, respectively for Small Edible, Medium Edible, and Big Edible mushrooms; ST, MT, BT, for Small Toadstool, Medium Toadstool, and Big Toadstool)*

The symbol acquisition test

To study the evolution of symbolisation it is important to establish if these apparent symbolic "verb-noun" structures are based on real symbolic relationships and if the organism is able to choose the correct verb with the name of each new edible or poisonous mushrooms. In order to analyse the type of referencing systems that organisms evolved, a symbol acquisition test was used, similar to that in Savage-Rumbaugh and Rumbaugh's (1978) chimpanzee experiments. The test was performed off-line, separated from the simulation on the evolution of language by auto-organisation. The goal was to teach organisms' neural nets a perfect "verb-noun" language. This language was imposed by providing the teaching input for the backpropagation cycle, as opposed to receiving it from the parent. The test consisted of three learning stages. In the first stage, organisms learned to label only four types of objects (large and medium toadstools, large and medium edible mushrooms). During this stage verbs were not used, and no names were taught for the remaining two categories (small edible and small poisonous mushrooms). In the second stage, organisms learned to associate the two verbs "approach" and "avoid" with the categories large/medium edible and large/medium poisonous mushrooms, respectively. At this point, it was expected that organisms would have learned the logical relationship between the names of the two edible mushrooms and the verb "approach", and the logical relationship between the verb "avoid" and the names of two toadstools. In the final stage the learning of the names of the small poisonous and small edible categories was finally introduced. The association of the two verbs with these new names was not taught. In fact, it was expected that only organisms that learned true symbolic relationships between verbs and names would be able to generalise this rule to new mushroom names.

The symbol acquisition test was repeated with ten different replications. After the three learning stages, seven populations produced the correct associations 'small_edible'-"approach" and 'small_toadstool'-"avoid". In three populations the learning of the names for small mushrooms did not produce the activation of the proper verb. It means that these organisms did not learn any symbolic association. In the seven successful populations, instead, the language is based on logical relationships between the mushrooms' names and the two verbs. The relationships between words and real objects, and between verbs and objects' name, allow neural nets to generalise the association of new names with the correct verb category. These results show that neural networks can learn simple languages that use symbolic associations.

Conclusion

The model simulation for the evolution of self-organising languages and the test of symbol acquisition show that neural nets, as chimpanzees, can be used as "models" for the study of evolution of language and of symbolisation. Some nets, such as some chimpanzees, were not able to learn a real symbolic language, even though they were apparently using languages with "verb-noun" rules. Further analyses of the nets' internal representations, and of the net's training history, will permit us to understand the conditions that lead to the acquisition of true symbolic languages. Moreover, computational models such as neural nets allow us to manipulate some of their features (e.g. the neural net architecture) to better understand the neural mechanisms for symbolisation and language acquisition.

References

Cangelosi A. (1999). Modeling the evolution of communication: From stimulus associations to grounded symbolic associations. In D. Floreano, J. Nicoud, F. Mondada (Eds.), *Advances in Artificial Life*, Berlin: Springer-Verlag, 654-663.

Cangelosi A. & Harnad S. (in press). The adaptive advantage of symbolic theft over sensorimotor toil: Grounding language in perceptual categories. *Evolution of Communication*. Presented at the 2nd International Conference on the Evolution of Language, London, April 1998.

Deacon T.W. (1997). *The symbolic species: The coevolution of language and human brain*, London: Penguin.

Fodor J. e Pylyshyn Z. (1988). Connectionism and cognitive architecture: A critical analysis. Cognition, 28, 3-71.

Greenfield P.M. & Savage-Rumbaugh S. (1990). Grammatical combination in Pan paniscus: Process of learning and invention in the evolution and development of language. In S.T. Parker & K.R. Gibson (eds), *Language and intelligence in monkeys and apes*, Cambridge University Press, 540-579.

Harnad S. (1990). The Symbol Grounding Problem. *Physica D* 42: 335-346

Kirby S. (1999). *Function, selection and innateness: The emergence of language universals*, Oxford University Press.

Marcus G.F. (1998). Rethinking eliminative connectionism. *Cognitive Psychology*, 37(3): 243-282.

Peirce C.S. (1955). Logic as semiotic: The theory of sign. In J. Buchler (Ed.), *The philosophical writing of Peirce*. New York: Dover Books.

Rumelhart D.E. e McClelland J.L. (Eds.) (1986). *Parallel Distributed Processing: Explorations in the microstructure of cognition*. Cambridge, MA: MIT Press.

Savage-Rumbaugh S. (1986). *Ape languages: From conditioned response to symbol*. New York: Columbia University Press.

Savage-Rumbaugh S. & Rumbaugh D.M. (1978). Symbolization, language, and Chimpanzees: A theoretical reevaluation on initial language acquisition processes in four Young Pan troglodytes. *Brain and Language*, 6: 265-300.

Savage-Rumbaugh S., Rumbaugh D.M., Smith S.T., & Lawson J. (1980). Reference: The linguistic essential. *Science*, 210, 922-925.

Steels, L. (1997) The synthetic modeling of language origins. *Evolution of Communication*, 1(1).

Using artificial language learning to study language evolution: Exploring the emergence of word order universals

Morten H. Christiansen

Southern Illinois University
morten@siu.edu

The study of the origin and evolution of language must *necessarily* be an interdisciplinary endeavor. Only by amassing evidence from many different disciplines can theorizing about the evolution of language be sufficiently constrained to remove it from the realm of pure speculation and allow it to become an area of legitimate scientific inquiry. Fueled by theoretical constraints derived from recent advances in the brain and cognitive science, the last decade of the twentieth century has seen a resurgence of scientific interest in the origin and evolution of language. Nonetheless, direct experimentation is needed in order to go beyond existing data. Computational modeling has become the paradigm of choice for such experimentation as evidenced by the many computational papers presented at the two previous Evolution of Language conferences. Computational models provide an important tool with which to investigate how various types of constraints may affect the evolution of language. One of the advantages of this approach is that specific constraints and/or interactions between constraints can be studied under controlled circumstances.

In this paper, I point to *artificial language learning* (ALL) as an additional, complementary paradigm for exploring and testing hypotheses about language evolution. ALL involves training human subjects on artificial languages with particular structural constraints, and then testing their knowledge of the language. Because ALL permits researchers to investigate the language learning abilities of infants and children in an highly controlled environment, the paradigm is becoming increasingly popular as a method for studying language acquisition (e.g. Saffran, Aslin & Newport,1996). I suggest that ALL can similarly be applied to the investigation of issues pertaining to the origin and evolution of language in much the same way as computational modeling is currently being used. In the remainder of this paper, I demonstrate the utility of ALL as a tool for studying the evolution of language by reporting on two ALL experiments that test predictions derived from previous computational work on the constraints governing the emergence of basic word order universals (Christiansen & Devlin, 1997).

Explaining the emergence of basic word order universals

There is a statistical tendency across the languages of the world to conform to a basic format in which the head of a phrase consistently is placed in the same position – either first or last – with respect to the remaining clause material. Within the Chomskyan approach to language, head direction consistency has been explained in terms of an innate module (X-bar theory) that specifies constraints on the phrase structure of languages. Pinker (1994) has further suggested that this module emerged as a product of natural selection.

This paper presents an alternative explanation for head-order consistency based on the suggestion by Christiansen (1994) that language has evolved to fit sequential learning and processing mechanisms existing prior to the appearance of language. These mechanisms presumably also underwent changes after the emergence of language, but the selective pressures are likely to have come not only from language but also from other kinds of complex hierarchical processing, such as the need for increasingly complex manual combination following tool sophistication. On this view, head direction consistency is a by-product of non-linguistic constraints on hierarchically organized temporal sequences. Christiansen & Devlin (1997) provided connectionist simulations in which simple recurrent networks were trained on corpora generated by 32 different grammars with differing amounts of head-order consistency. These networks did not have built-in linguistic biases; yet they were sensitive to the amount of head-order inconsistency found in the grammars. There was a strong correlation between the degree of head-order consistency of a given grammar and the degree to which the network had learned to master the grammatical regularities underlying that grammar: The higher the inconsistency, the more erroneous the network performance. This suggests that constraints on basic word order may derive from non-linguistic constraints on the learning and processing of complex sequential structure, thus obviating the need for an innate X-bar module for this purpose. Grammatical constructions incorporating a high degree of head-order inconsistency are difficult to learn and will therefore tend to disappear, whereas consistent constructions should proliferate in the evolution of language.

If this line of reasoning is correct, one would expect to be able to find evidence of sensitivity to head-order inconsistency in human sequential learning performance. Experiment 1 tests this prediction using an ALL task with normal adults. More generally, this account also predicts a strong association between language processing and the processing of sequential structure. Experiment 2 tests this prediction comparing the performance of agrammatic aphasics with matched controls in an ALL task.

Experiment 1: Testing for sensitivity to head-order consistency in sequential learning

Two artificial languages were created based on two grammars taken from the Christiansen and Devlin (1997) simulations (see Table 1). Note that the consistent grammar is all head-final to avoid possible contamination from the head-initial nature of English. Both grammars encoded subject-noun/verb agreement. Pairs of strings were generated–one from the consistent grammar and one from the inconsistent grammar – using a vocabulary consisting of six consonants (X = plur. N; Z = prep/post; Q = plur. N; V = sing. N; S = sing. V; M = plur. V). Each string in a pair has the same lexical items and the same grammatical structure as the other, but may differ in the sequential ordering of the lexical items depending on the grammar (e.g., the pair VVQXQXS and VQQVXXS). Thirty pairs in which the sequential ordering differed were selected for *training*. Thirty pairs of identical strings differing from the training items were selected to serve as grammatical test items. Thirty ungrammatical test items were generated by changing a single letter in each grammatical item (first and last letters excluded) to produce an item that was ungrammatical according to both grammars.

In the consistent condition (CON), 20 subjects were trained on the consistent items. In the inconsistent condition (INCON), 20 subjects were trained on the inconsistent items. During training each string was presented briefly on a computer screen, and the subject prompted to type it in using the keyboard. Subjects in both conditions were trained on three blocks of 30 training items, before being tested on two blocks of the 60 test items. Subjects were informed about the rule-based nature of the stimuli only prior to the test phase, and asked to classify the novel strings according to whether or not they followed the same rules as the training items. In a third control condition, 20 subjects went directly to the test phase.

Consistent Grammar	**Inconsistent Grammar**
S → NP VP	S → NP VP
NP → (PP) N	NP → (PP) N
PP → NP post	PP → pre NP
VP → (PP) (NP) V	VP → (PP) (NP) V
NP → (PossP) N	NP → (PossP) N
PossP → NP Poss	PossP → Poss NP

Table 1: ***The Two Grammars Used in Experiment 1***

With a classification performance of 63.0%, the CON group was significantly better at classifying the test items than the INCON group with only 58.3% ($t(38)=2.54, p<.02$). The CON group (67.8%) was numerically better than the INCON (65.8%) at classifying grammatical items, though this difference was not significant ($t < 1$). However, the INCON group was essentially at chance when it came to rejecting the ungrammatical items (51.7%), and performed significantly worse than the CON group (58.1%) on these items ($t(38)=2.47, p<.02$). With a correct classification of only 52.38%, the performance of the control subjects demonstrated that there was no bias in the test items. Thus, the results show that the subjects found the consistent grammar easier to learn than the inconsistent grammar as reflected in their ability to perform grammaticality judgments on the test items.

Experiment 2: Testing for an association between sequential learning and language

The theory of language evolution presented here suggests that language evolved to fit sequential learning and processing mechanisms. A straightforward prediction from this is that breakdown of language should be associated with breakdown of sequential learning. Christiansen, Kelly, Shillcock & Greenfield (2000) provide evidence in support for this prediction. They compared the ALL performance of seven agrammatic aphasic patients with seven normal controls matched for age, gender, socio-economic status and reasoning ability. A simple finite-state grammar was used for this experiment, with letter strings replaced by combinations of symbols (using the zapf dingbats font; e.g., βλνβν). For training, subjects carried out a match/mismatch pairing task in which they had to determine whether the second of two consecutively presented symbol strings matched the first. Testing was similar to Experiment 1, in that the subjects were informed about the rule-based nature of the training items and asked to classify 40 new test items (half grammatical/half ungrammatical).

Both aphasics and normals were able to perform the match/mismatch training task without any problems. However, the aphasic patients performed at chance (51.1%) on the classification task, significantly worse than the normal controls (62.1%; $t(12)=2.53, p<.03$). The controls were significantly better at classifying the grammatical items (65.0%) compared with the aphasics (38.6%; $t(12)=2.37, p<.04$). However, both groups were equally good at classifying the ungrammatical items (aphasics: 63.7%; controls: 59.3%; $t<1$). The results suggest that the aphasics had problems with sequential learning in this experiment, whereas amnesics have few problems on similar ALL tasks (Knowlton & Squire, 1994). Thus, as predicted, the language breakdown in aphasia appears to be associated with a breakdown of sequential learning.

Conclusion

The two ALL experiments and the computational simulations presented by Christiansen & Devlin (1997) support the claim that basic word order universals (head-ordering) can be explained in terms of non-linguistic constraints on sequential learning and processing, rather than as a product of innate linguistic knowledge. Thus, rather than a biological adaptation of learning mechanisms to fit linguistic structure, the evidence points to the adaptation of linguistic structure to fit pre-existing sequential learning mechanisms.

More generally, the results of the two experiments reported above and similar ALL results pertaining subjacency reported in Ellefson & Christiansen (2000) suggest that ALL is a useful tool for exploring issues relating to language evolution. It may be objected that the languages used in ALL experiments are simple and deviate significantly from natural language. However, the same objection can be raised against the computational models of language evolution, but this has not diminished their impact, nor their usefulness to the study of language evolution. Moreover, ALL also provides a new tool with which to study other aspects of language evolution, such as creolization (Hudson & Newport, 1998) and comparative aspects of language evolution (Hauser, Newport & Aslin, 2000). In this way, ALL promises to open up a whole new direction in the search for evidence to constrain scientific theories of language evolution.

References

Christiansen, M.H. (1994). *Infinite languages, finite minds: Connectionism, learning and linguistic structure*. Unpublished doctoral dissertation, Centre for Cognitive Science, University of Edinburgh, U.K.

Christiansen, M.H. & Devlin, J.T. (1997). Recursive inconsistencies are hard to learn: A connectionist perspective on universal word order correlations. In *Proceedings of the 19th Annual Cognitive Science Society Conference* (pp. 113-118). Mahwah, NJ: Lawrence Erlbaum Associates.

Christiansen, M.H., Kelly, L., Shillcock, R.C. & Greenfield, K. (2000). *Artificial grammar learning in agrammatism*. Manuscript in preparation.

Ellefson, M.R. & Christiansen, M.H. (2000). *The evolution of subjacency without Universal Grammar: Evidence from artificial language learning*. Paper to be presented at the 3rd Conference on the Evolution of Language, Paris, France.

Hauser, M., Newport, E. L. & Aslin, R. N. (2000). *Statistical learning of serial order in cotton-top tamarins*. Submitted manuscript.

Hudson, C.L. & Newport , E.L. (1998). Creolization: Could adults really have done it all? In A Greenhill, H. Littlefiled & C. Tano (Eds.), *Proceedings of the 23rd annual Boston University Conference on Language Development* (pp. 265-276). Boston: Cascadilla Press.

Knowlton, B.J. & Squire, L.R. (1994). The information acquired during artificial grammar learning. *JEP: Learning, Memory and Cognition, 20,* 79-91.

Pinker, S. (1994). *The language instinct: How the mind creates language*. New York: William Morrow.

Saffran, J.R., Aslin, R.N. & Newport, E.L. (1996). Statistical learning by 8-month olds. *Science, 274*, 1926-1928.

Did language evolve from manual gestures?

Michael C. CORBALLIS

University of Auckland

ABSTRACT

True language appears to be unique to humans. This has created severe problems in understanding how it evolved, since there is little that can be learned about its evolution from the communications of other species. Attempts at explanation have therefore depended more on speculation than on fact, prompting the Linguistic Society of Paris, as early as 1866, to ban all discussion of the language origins. Moreover, given that spoken language is based on abstract symbols (words), it is not clear how it began: How was it decided which symbols would stand for which concepts, and how was this propagated? As Rousseau remarked, “Words would seem to have been necessary to establish the use of words.” I shall argue that this apparent paradox can be at least partially resolved if it is supposed that language evolved from manual gestures, since gestures have at least the potential to represent concepts iconically rather than in abstract form. Once a set of iconic representations is established, increasing usage can then lead to more stylized and ultimately abstract representation, as has occurred in the evolution of writing systems.

Manual and bodily gestures play a prominent part in contemporary language. The work of David McNeill and his colleagues at the University of Chicago have documented the role of gestures in present-day spoken language, and shown that gestures take on syntactic structure if people are prevented from speaking. A number of societies, notably certain groups of aboriginal people in Australia and the United States, have developed sign languages that can function in the absence of speech. The most accessible sign languages, however, are those invented by deaf communities, and these are clear examples of purely gestural language that have all of the important hallmarks of true language, including fully developed syntax. Children learning sign language from infancy go through essentially the same developmental stages as those learning to speak, and sign language also appears to be predominantly left-hemispheric.

Primates, including ourselves, are predominantly visual creatures, with excellent voluntary control over the muscles of the limbs, while in nonhuman primates control over vocalization is relatively poor and is largely emotional rather than voluntary. The vocal and auditory systems system in primates are better adapted to an arousing or alerting function than to a descriptive or narrative one. Where the intent is to convey information about a four-dimensional world of space and time, as is the case in human language, the early hominids were surely better preadapted to use gestures, which permit four-dimensional representation, rather than vocalization, which is essentially restricted to the single dimension of time. As evidence for this, there has been at least some measure of success in teaching a form of sign language to chimpanzees and gorillas, whereas attempts to teach them to speak have been fruitless. Moreover, it has recently been discovered that there are neurons in the prefrontal cortex of macaques that respond both when the animal makes a specific grasping movement, and when it observes the same grasping movement made by others. These so-called “mirror neurons”, it has been suggested, may represent a precursor to a gestural language. They may also relate to an

ability to take the mental perspective of others, which can be regarded as a necessary precursor to language.

The hominids split from the precursor of the modern chimpanzee about five million years ago. The main characteristic distinguishing the hominids was bipedalism, whereas the common ancestor was presumably a quadrupedal knuckle walker, as are present-day chimps and gorillas. Bipedalism freed the hands and arms from any major role in locomotion, and created a more frontal stance, both of which would have boosted a gestural, visual form of communication. It is conceivable that gestural communication was a factor in the selection of the bipedal stance, and not simply a fortuitous consequence of it. These early hominids lived on the savanna-like territory, mostly to the east of the Great Rift Valley in Africa, and an effective, silent form of communication may have been crucial to survival in a habitat populated by dangerous killers, such as the precursors to modern tigers, lions, and hyenas.

About two million years ago, at least one branch of hominids, now called *hominins*, began to show new characteristics. These included increased brain size, the emergence of manufactured stone tools, and the beginnings of migrations from Africa into Asia and Central Europe. I shall argue that these developments may have heralded the emergence of a more sophisticated language, still probably mainly gestural, but including a recursive syntax that enabled communication to be generative. It is likely that vocalizations increasingly accompanied gestures, which might explain why cerebral asymmetry links handedness with left-hemispheric control of vocalization.

Recent evidence suggests that modern-day humans evolved from an African branch of hominins some 100-150,000 years ago. This new species, *Homo sapiens*, also migrated out of Africa, beginning perhaps 70,000 years ago, and eventually replaced all other hominids, including the Neanderthals in Europe and *Homo erectus* in Java and Southern Asia. What explains the dominance of *H. sapiens* over equally large-brained hominids like the Neanderthals? I suggest that it was the conversion from a form that was dependent on gestures to one that could function entirely vocally. It presumably took place before *H. sapiens* migrated from Africa, since present-day humans speak, and it is unlikely that autonomous speech arose independently in different geographic locations. On the other hand, reconstructions of the vocal tract have been interpreted to mean that the Neanderthals would have been incapable of fluent speech. This suggests that the conversion to autonomous speech took place, perhaps gradually, somewhere between about 150,000 and 70,000 years ago. The conversion may have been quite a small step, since vocalizations probably played an increasing role throughout hominid evolution, but it was a crucial step because it freed the hands from communication. This would have enhanced tool manufacture, allowing people to explain techniques verbally while demonstrating them. This may have heralded the beginning of pedagogy. It would also have allowed communication at night, and when obstacles prevent communicating parties from viewing each other. It also places fewer demands on focal attention.

I suggest that the emergence of sophisticated tools, art, ornamentation, and human culture can be attributed to the emergence of a fully autonomous vocal language. There was a particularly dramatic increase in the sophistication of human artifacts around 30-40,000 years ago, which is perhaps too late to be attributed directly to the emergence of autonomous vocal language some 60-100,000 years earlier. However, recent evidence suggests that this so-called “evolutionary explosion” may have occurred much earlier in Africa, and earlier migrations, such as that to Australia, must have required sophisticated technology. I suggest that, with the switch to vocalization, technology and social complexity progressed in exponential fashion, and shows no signs of abating. One of the consequence of the freeing of the hands was the development of new visual forms of communication, including pictorial art, writing, and ultimately photography, film, and computer graphics. So we come back to ways of exploiting the visual sophistication that we inherited from our primate forbears.

The idea that language evolved from manual gestures is not new. It was suggested in 1746 by the philosopher Condillac, and was revived in the 1970s by the anthropologist Gordon W. Hewes. It received something of a boost, notably from the linguist William Stokoe, when it was fully understood that the sign languages of the deaf have all of the essential hallmarks of true language, including generative syntax. A number of linguists and anthropologists have supported the idea, but on the whole it has not been widely accepted. For example, one of the most popular and influential of recent books on language, Steven Pinker's *The Language Instinct*, essentially dismisses the idea, despite the fact that it includes examples from sign language to support ideas about the development and nature of language. I believe the idea is now sufficiently compelling that it should be taken seriously.

Putative location and identity of the gene for cerebral asymmetry and language

Tim J. CROW

POWIC
University Department of Psychiatry, Warneford Hospital
Oxford OX3 7JX.

"Human language is an embarrassment for evolutionary theory" (D. Premack) because as Chomsky (1972) pointed out in language we have a faculty without clear precursors in primate evolution, and as Bickerton (1995) has argued, that appeared suddenly and recently. Such an innovation is incompatible with the prevailing "biological species concept" that species transitions occur along genetic gradients that are generated in geographically separated environments. The origin of language appears to be an instance of "saltational" evolution consistent with the concepts of Goldschmidt (1940) or of punctuated equilibria of Eldredge & Gould (1972) that have lacked followers or a specified mechanism. One suggested mechanism – chromosomal re-arrangement (White, 1973; King, 1993) – has been generally disregarded on the grounds that such changes may occur without phenotypic effects, and have inconsistent relationships with species boundaries.

A possible neural correlate of language – asymmetry of the cerebral hemispheres – has been generally discounted because it is widely believed that such asymmetries are present in other vertebrates. This view is now strongly challenged by the observations of Marchant & McGrew (1996) that population-based directional hand preferences such as are present in all human populations are absent in chimpanzees, and of Buxhoeveden & Casanova (in press) that asymmetries of pyramidal cell columns in the superior temporal gyrus present in the human brain, are absent in that of the chimpanzee. The anatomical torque therefore appears to be a unique feature of the human brain, and a correlate of language.

The genetic origin of cerebral asymmetry has been unravelled through the neuropsychological correlates of sex chromosome aneuploidies – anomalies in the number of X chromosomes. Individuals who lack an X chromosome (Turner's syndrome) have non-dominant hemisphere deficits, whilst individuals with an extra X chromosome (XXY or Klinefelter's and XXX syndromes) have dominant or verbal deficits. This strong evidence that a gene for relative hemispheric development is present on the X chromosome must be contrasted with the absence of non-dominant hemisphere deficits in normal males who, like females with Turner's syndrome, have only one X. A gene on the Y chromosome therefore must balance the gene on the X chromosome. This logic identifies the human cerebral asymmetry factor as in the select class of X-Y homologous genes (Crow, 1993; Netley, 1998).

The history of the Y chromosome in primate evolution suggests a specific location. X-Y homologies have been generated by translocations of blocks from the X to the Y, and these can be dated in relation to species separations (Lambson et al, 1992; Affara et al, 1996). One such block – the homology between the Xq21.3 region of the X long arm and two blocks of sequences in the Y short arm (Yp) – was created by a translocation that occurred approximately 3 million years ago – after the separation of the chimpanzee and hominid lineages (Sargent et al, 1996). The two blocks on Yp were created by a subsequent paracentric inversion which occurred more recently – probably within the last 500,000 years (Schwartz et al, 1998). These events established gene sequences on the Y chromosome that are Homo sapiens specific. Any change (including the paracentric inversion) that is present within this region that might have an influence on gene expression and can be shown to be present in all human populations is a candidate for the speciation event.

Transmission of handedness (a correlate of cerebral asymmetry) within families is associated with sex in a manner that is consistent with X-Y linkage (Corballis et al, 1996). Evidence for an association of cerebral asymmetry with a specific region of the X chromosome was sought in a linkage study in 180 pairs of left-handed brothers (Laval et al, 1998). Increased allele sharing (consistent with the presence of a gene) was found directly over the Xq21.3 region.

Because of the particular interest of the Xq21.3/Yp region of homology in relation to human evolution the block has been given priority in the sequencing strategy of the Sanger Human Genome Centre. Two genes within the region have so far been identified – a phospho-diesterase and a proto-cadherin sequence. The latter is of particular interest because such sequences form a large family of cell adhesion molecules that are associated with specific neuronal pathways that differ in distribution between species(Crow, 2000).

A gene for asymmetry in homologous form within the non-recombining portions of the sex chromosomes will be subject to sexual selection (mate selection by criteria that may differ between the sexes). This is of interest because it has been suggested (eg by Kaneshiro, 1980) that sexual selection and speciation are related in that the primary event in speciation occurs in a characteristic that differentiates the sexes and that this then becomes subject to sexual selection. Such selection might account for an “escalation” in the development of that characteristic in the two sexes. It might explain how a change in a single individual spreads within the population, and how an initial change on the Y chromosome through selection acting on the homologous sequences on the X chromosome came to have an influence on both sexes (Crow, 1998a,b, 2000). Such sequential interactions between the chromosomes may be mediated by “epigenetic” modification of the DNA sequence eg through protection from the process of X inactivation (Jegalian & Page, 1999).

According to this theory the primary event in speciation that enabled the evolution of language was a discrete change on the Y chromosome that occurred in a single male. The change in gene dosage that this brought about became the subject of progressive modification in both males and females through a process of sexual selection. The effect of the change was to allow the cerebral hemispheres to develop with a degree of independence that may be assumed to have been associated with a delay in maturation. Evidence that the process of sexual selection acts through hemispheric differentiation is apparent in the relationship between degrees of lateralization and verbal ability that is similar in form but quantitatively displaced in females relative to males (Crow et al, 1998).

This explanation of the origins of language thus requires adherence to a saltational theory of speciation events mediated by a rare but selected change occurring on the sex chromosomes. The case of the evolution of language in Homo sapiens thus provides a paradigm for the genetic mechanism by which novel mate recognition systems are initiated and consolidated in the origin of species.

References

Affara, N., Bishop, C., Brown, W., Cooke, H., Davey, P., Ellis, N., Graves, J.M., Jones, M., Mitchell, M., Rappold, G., Tyler-Smith, C., Yen, P., & Lau, Y.C. (1996). Report of the second international workshop on Y chromosome mapping 1995. Cytogenetics and Cell Genetics, 73, 33-76.

Bickerton, D. (1995). Language and Human Behavior. Seattle: University of Washington.

Buxhoeveden, D., & Casanova, M. (1999) Comparative lateralization patterns in the language area of normal human, chimpanzee, and rhesus monkey brain. Laterality (in press).

Chomsky, N. (1972). Language and Mind. San Diego: Harcourt, Brace & Jovanovich.

Corballis, M.C., Lee, K., McManus I.C., & Crow T.J. (1996). Location of the handedness gene on the X and Y chromosomes. American Journal of Human Genetics (Neuropsychiatric Genetics) 67, 50-52.

Crow, T.J. (1993). Sexual selection, Machiavellian intelligence and the origins of psychosis. Lancet, 342, 594-598.

Crow, T.J. (1998a) Sexual selection, timing and the descent of man: A theory of the genetic origins of language. Current Psychology of Cognition, 17, 1079-1114.

Crow, T.J. (1998b) Why cerebral asymmetry is the key to the origin of Homo sapiens: How to find the gene or eliminate the theory. Current Psychology of Cognition, 17, 1237-1277.

Crow, T.J. (2000) Did Homo sapiens speciate on the Y chromosome ? Psycoloquy 11(01).

Crow, T.J., Crow, L.R., Done, D.J., & Leask, S.J. (1998). Relative hand skill predicts academic ability: global deficits of the point of hemispheric indecision. Neuropsychologia, 36, 1275-1282.

Eldredge, N., & Gould, S.J. (1972). Punctuated equilibria: an alternative to phyletic gradualism. In T.M. Schopf (Ed.), Models in Palaeobiology. (pp. p82-115). San Francisco: Freeman Cooper.

Goldschmidt, R. (1940) The Material Basis of Evolution. New Haven: Yale University Press.

Jegalian, K. & Page, D.C. (1998) A proposed mechanism by which genes common to mammalian X and Y chromosomes evolve to become X inactivated. Nature, 394, 776-780.

Kaneshiro, K.Y. (1980). Sexual isolation, speciation and the direction of evolution. Evolution, 34, 437-444.

King, M. (1993). Species Evolution: the role of chromosome change. Cambridge: Cambridge University Press.

Lambson, B., Affara, N.A., Mitchell, M., & Ferguson-Smith, M.A. (1992). Evolution of DNA sequence homologies between the sex chromosomes in primate species. Genomics, 14, 1032-1040.

S.H. Laval, J. Dann, R.J. Butler, J. Loftus, J. Rue, S.J. Leask, N. Bass, M.Comazzi, A. Vita, S. Nanko, S. Shaw, P. Peterson, G. Shields, A.B. Smith, J. Stewart, L.E. DeLisi and T.J. Crow, Evidence for linkage to psychosis and cerebral asymmetry (relative hand skill) on the X chromosome. Am. J. Med. Genet. (Neuropsychiatric Genet.), 81 (1998) 420-427.

Marchant, L.F., & McGrew, W.C. (1996) Laterality of limb function in wild chimpanzees of Gombe National park: comprehensive study of spontaneous activities. Journal of Human Evolution, 30, 427-443.

Netley, C. (1998) Sex chromosome aneuploidy and cognitive development. Current Psychology of Cognition, 17, 1190-1197.

Sargent, C.A., Briggs, H., Chalmers, I.J., Lambson, B., Walker, E., & Affara, N.A. (1996). The sequence organization of Yp/proximal Xq homologous regions of the human sex chromosomes is highly conserved. Genomics, 32, 200-209.

Schwartz, A., Chan, D.C., Brown, L.G., Alagappan, R., Pettay, D., Disteche, C., McGillivray, B., de la Chapelle, A., Page, D.C. (1998) Reconstructing hominid evolution: X-homologous block, created by X-Y transposition, was disrupted by Yp inversion through LINE-LINE recombination. Human Molecular Genetics 7: 1-11.

White, M.J.D. (1973) Modes of Speciation. San Francisco: WH Freeman.

Ontogenetic versus phylogenetic learning of the emergence of phonetic categories

Robert I. Damper

Image, Speech and Intelligent Systems (ISIS) Research Group
Department of Electronics and Computer Science
University of Southampton, Southampton S017 1BJ, UK.

Introduction

The scientific study of spoken language confronts two central problems: (1) How is the *continuous* sound pressure wave of speech decoded by a listener into the *discrete* percept of a linguistic message? (2) How does each individual listener acquire the particular form of sound system, morphology, grammar, etc. appropriate to the language of his/her social community? The first of these is the classical problem of categorical perception (Harnad 1987) and the second is that of language acquisition (Wexler & Culicover 1980; Pinker 1984). Both have been intensively studied, and given rise to intense debate and controversy. They are generally held to be rather different problems: plainly, however, each makes little or no sense by itself.

Language is acquired rapidly and robustly by speaker/listeners in spite of obvious theoretical difficulties such as "poverty of the stimulus" (Chomsky 1965), non-uniqueness of the object to be learned (Gold 1967), and the vast preponderance of positive examples (Angluin 1980). These issues have influenced ideas of language acquisition profoundly and need no further rehearsal here. They are generally taken as evidence for a nativist stance, holding that much of language acquisition (the 'universal' component) must be genetically constrained. The developing child's task is then to infer by exposure in early life the *particular*, local variant of spoken language.

So what are the prospects of explaining either or both of the above faculties in terms of evolutionary emergence – the topic of this workshop? Our starting point is the suggestion of Hurford (1992, p.292): "… that promising candidate design features for evolutionary explanation include … [I]n phonetics, the phenomena of categorical perception … and the tendency of speech to lump itself into segments." Yet this suggestion appears far too nativist. Since learners in different speech/language communities acquire different variants of sound system, it seems incontrovertible that there must be an ontogenetic aspect. This latter view is supported by experimental evidence from cross-language and infant perception studies (Wilson 1977; Simon & Fourcin 1978; Werker & Tees 1984) suggesting that human neonates are capable of distinguishing all phonetic contrasts in the world's languages, but these generalised abilities are restricted by learning through exposure to become specific, cf. recent ideas of 'perceptual magnets' (Kuhl 1991; Kuhl et al. 1992; Guenter & Gjaja 1996). Hence, the thesis explored in this paper is that an evolutionary explanation is only warranted for *part* of the process of acquiring phonetic categories. The question then is: which part? Specifically, it is argued – initially at least, before refining the argument – that phylogenetic (evolutionary) learning provides a pre-processor which eases the problem of ontogenetic (during life) language acquisition by the individual. This preprocessor is the peripheral auditory system.

The notion of a preprocessor which evolves in Darwinian fashion and effects a recoding of the stimulus to ease the problem of learnability is consonant with the arguments of Clark and Thornton (1997). Distinguishing between tractable type-1 and difficult/intractable type-2 learning problems, they posit (but do not pursue the point) "that evolution gifts us with exactly the right set of recoding biases so as to reduce specific type-2 problems to … type-1 problems” (p.57). How might this work in the case of phonetic categorisation of speech sounds?

Modelling categorisation of speech sounds

According to MacWhinney (1998): “It now appears that the ability to discriminate the sounds of language is grounded on raw perceptual abilities of the mammalian auditory system” (p.202) yet: “Even in the areas to which they have been applied, emergentist models are limited in many ways … the development of the auditory and articulatory systems is not yet sufficiently grounded in physiological and neurological facts” (p.222). In previous work over many years, however, we have studied extensively the emergence of phonetic categories in a variety of computational models, but all having the general form shown in Figure 1 (Damper, Pont & Elenius 1990; Damper, Gunn & Gore, forthcoming; Damper & Harnad, forthcoming). By inclusion of a front-end simulation of the mammalian auditory system, we explicitly ground our model in physiological and neurological facts. Space precludes any detailed treatment: suffice to say that the computational models convincingly replicate the results of human and animal experimentation, in both category labelling and discrimination.

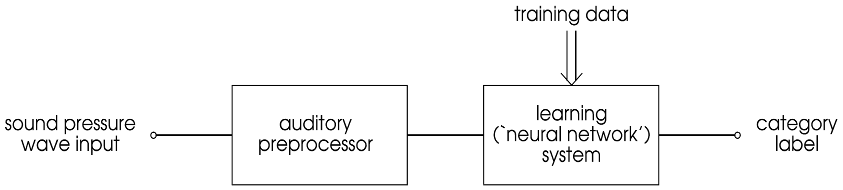
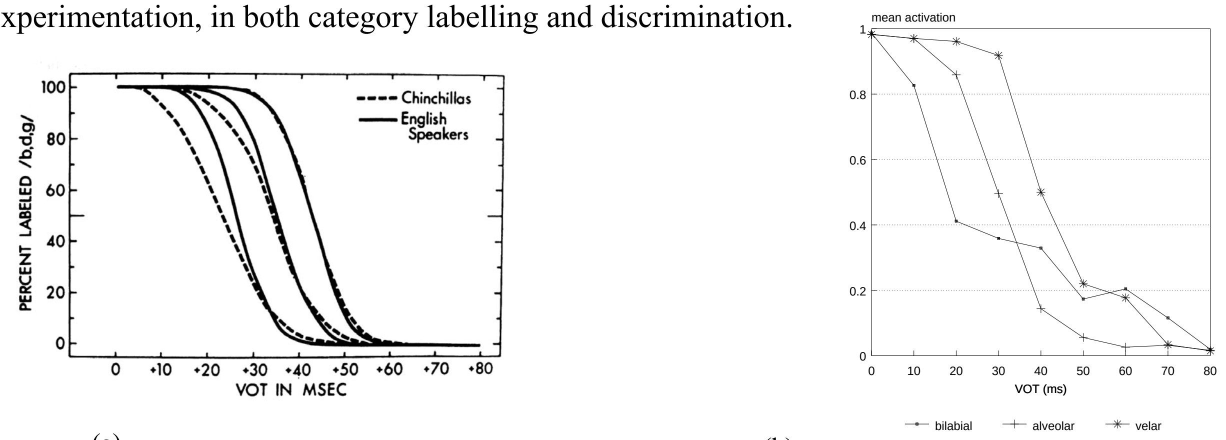


Figure 1: Two-stage computational model for the categorisation of speech sounds. An auditory model converts the sound pressure wave input into a time-frequency representation of auditory nerve firings, followed by a ‘neural network’ trained to convert the auditory time-frequency patterns of firing into a category label.



(a) (b)

Figure 2: *Labelling results for the categorisation of synthetic CV stimuli varying in voice onset time. (a) Human and chinchilla results from Kuhl and Miller (1978) and (b) typical results using a single-layer perceptron as the learning component of the model in Figure 1.*

To illustrate this, Figure 2 compares the results of labelling experiments with (a) human and chinchilla listeners and (b) one version of the computational model depicted in Figure 1. In both cases, input stimuli were tokens from synthetic series of initial-stop-consonant/vowel stimuli varying in voice onset time (VOT). There were three such series (bilabial, alveolar and velar), varying in place of articulation of the stop consonant. To mimic the operant training of the animals, the neural networks (one for each place of articulation) were trained on the extreme VOT values of 0 and 80 ms and then tested on the full range of values. (Of course, adult human listeners need no such training.) In (a), but not in (b), smooth curves have been fitted to the raw data which otherwise are essentially identical. Just like human and animal listeners, the computational model produces plateaux, corresponding to the two (voiced/unvoiced) categories, at the extremes of the labelling curve, with a sharp transition between them. Learning is robust (largely independent of details of the learning system) and rapid, requiring only a handful of training epochs. (No 'poverty of the stimulus' here!) Nor is supervised learning essential – a potentially important fact because feedback on correct categorisation is not obviously present in the real learner's environment. Also, the movement of category boundary with place of articulation is correct. The model does not merely place the boundary at the midpoint of the VOT range as we would expect a 'dumb' pattern classifier to do.

However, the learning component of our model *is* nothing other than a dumb (linear, at that) pattern classifier. This suggests that the separation according to place of articulation must be effected by the front-end auditory component. This is indeed the case: replacing the front end by a simpler Fourier spectrum analyser abolishes the boundary-movement effect. That is, the peripheral auditory system – a product of evolution – preprocesses the speech sounds into a form which promotes the learning of phonetic categories.

Discussion and implications

The work described here falls into the 'empiricist' approach to the study of speech perception exemplified by Kluender (1991), Nearey (1997), Sussman et al. (1998) and others. According to this increasingly influential tradition, "speech is decodable via simple pattern recognition because talkers adhere to 'orderly output constraints' " (Kluender & Lotto 1999, p.504). The latter authors warn, however: "With the suggestion that a complete model include both auditory and learning processes … falsifiability is at risk … part of the explanation of speech perception falls out of general auditory processes, and the remaining variance can be 'mopped up' by learning processes". But, of course, the complete model must all the time respect current knowledge of auditory system anatomy and physiology, computational learning theory, the facts of speech perception and so on, as we have tried to do here.

Kluender and Lotto (1999) also warn: "It is not very useful to hypothesize that learning plays some role. Of course it does" (p.508). So how successful have we been in teasing apart phylogenetic (evolutionary) and ontogenetic (learning) factors? Probably the main aspect in which we appear to be in conflict with available *human* data (reviewed above) is that these data indicate that neonates are born with pre-wired categories. These seem not to have to be learned, although they may have to be *un*learned. In principle, this could be incorporated in the model by some phylogenetic setting of the initial weights/parameters of the learning system. At this stage, it is uncertain whether this is reasonable or not. The issue should become clearer as the modelling work is extended and deepened to cover more phonetic categories and to deal with real speech.

References

Angluin, D. (1980). Inductive inferences of formal languages from positive data. *Information and Control*, *45*, 117-145.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Clark, A. and C. Thornton (1997). Trading spaces: Computation, representation and the limits of uninformed learning. *Behavioral and Brain Sciences*, *20*, 69-70.

Damper, R.I., S.R. Gunn, and M.O. Gore (forthcoming). Extracting phonetic knowledge from learning systems: Perceptrons, support vector machines and linear discriminants. *Applied Intelligence*, in press.

Damper, R.I. and S.R. Harnad (forthcoming). Neural network models of categorical perception. *Perception and Psychophysics*, in press.

Damper, R.I., M.J. Pont, and K. Elenius (1990). Representation of initial stop consonants in a computational model of the dorsal cochlear nucleus. Technical Report STL-QPSR,4/90, Speech Transmission Laboratory Quarterly Progress and Status Report, Royal Institute of Technology (KTH), Stockholm.

Gold, E. M. (1967). Language identification in the limit. *Information and Control*, *16*, 447-474.

Guenter, F. H. and M. N. Gjaja (1996). The perceptual magnet effect as an emergent property of neural map formation. *Journal of the Acoustical Society of America*, *100*, 1111-1121.

Harnad, S. (Ed.) (1987). *Categorical Perception: the Groundwork of Cognition*. Cambridge, UK: Cambridge University Press.

Hurford, J. R. (1992). An approach to the phylogeny of the language faculty. In J. A. Hawkins and M. Gell-Mann (Eds.), *The Evolution of Human Languages*, pp.273-303. Reading, MA: Addison-Wesley.

Kluender, K.R. (1991). Effects of first formant onset properties on voicing judgements resulting from processes not specific to humans. *Journal of the Acoustical Society of America*, *90*, 83-96.

Kluender, K.R. and A.J. Lotto (1999). Virtues and perils of an empiricist approach to speech perception. *Journal of the Acoustical Society of America*, *105*,503-511.

Kuhl, P.K. (1991). Human adults and human infants show a 'perceptual magnet effect' for the prototypes of speech categories, monkeys do not. *Perception and Psychophysics*, *50*, 93-107.

Kuhl, P.K. and J.D. Miller (1978). Speech perception by the chinchilla: Identification functions for synthetic VOT stimuli. *Journal of the Acoustical Society of America*, *63*, 905-917.

Kuhl, P.K., K.A. Williams, F. Lacerda, K.N. Stevens, and B. Lindblom (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, *225*, 606-608.

MacWhinney, B. (1998). Models of the emergence of language. *Annual Reviews of Psychology*, *49*, 199-227.

Nearey, T.M. (1997). Speech perception as pattern recognition. *Journal of the Acoustical Society of America*, *101*, 3241-3254.

Pinker, S. (1984). *Language Learnability and Language Development*. Cambridge, MA: Harvard University Press.

Simon, C. and A.J. Fourcin (1978). A cross-language study of speech-pattern learning. *Journal of the Acoustical Society of America*, *63*, 925-935.

Sussman, H.M., D. Fruchter, J. Hilbert, and J. Sirosh (1998). Linear correlates in the speech signal: The orderly output constraint. *Behavioral and Brain Sciences*, *21*, 241-299.

Werker, J. F. and R. C. Tees (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, *7*, 49-63.

Wexler, K. and P. Culicover (1980). *Formal Principles of Language Acquisition*. Cambridge, MA: MIT Press.

Williams, L. (1977). The perception of stop consonant voicing by Spanish-English bilinguals. *Perception and Psychophysics*, *21*, 289-297.

Tools, language and the origins of culture

Iain DAVIDSON

School of Human and Environmental Studies
University of New England, Armidale, NSW, AUSTRALIA
Iain.Davidson@une.edu.au

The three strands of evidence

Approaches to the study of the origins of language have three major sources of evidence: 1) From modern humans, including linguistics; 2) From modern non-human primates, including their communicative abilities; 3) From archaeology. This paper is about the third of these – the archaeological evidence of what actually happened. I will review the evidence from stone tools that has provided insight into the emergence of language over the last ten years since the Wenner-Gren conference in Cascais, Portugal. My paper will begin by outlining the argument I advanced in the conference volume *Tools, language and cognition in human evolution*.

The principal point was that the apparent existence of different tool types is not only a product of deliberate intentions to produce the forms recognised by archaeologists. It is this understanding which gave rise to our definition of a phenomenon we called "the finished artefact fallacy". This phrase has been misunderstood by Gowlett and by Mithen. I will clarify the concept in my presentation. Without plausible argument about the intentions of tool-makers, we cannot approach understanding of what tools tell us about the language-based symbolic representations in the mind.

"Imposed form" (Mellars)

One of the significant contributions to the discussion of the intentions involved in stone tool making is Mellars' introduction of the phrase "imposed form". If archaeologists can find a way to identify that hominins or humans imposed a form on an artefact, then it would be possible to comment on the process of conceptualisation among those hominins. For Mellars, imposed form was absent from Mousterian scrapers where the modification of the flake is related to its original form. But it was present in the Upper Palaeolithic of Europe where there were modifications of flakes which are visually distinctive, repetitive and standardised. Such modifications may or may not have been applied to flakes which were themselves of standardised form (often "blades").

At Cascais, I argued that the first appearance of this "imposed form" was with the distinctive geometric microliths of the Middle Stone Age in southern Africa. I would now reinforce this judgement because of the excellent documentation both of the processes of production of these tools and of bone tools in contemporary industries in different parts of Africa. I have been at pains to point out the constraints on artefact form imposed by the mechanics of stone knapping. No such constraints apply to the production of bone tools. In these African artefacts of 90 000 years ago, modification of form was *not* determined by the requirements of function or the contingencies of use. I believe an appropriate argument can be constructed that these products of the African MSA were made by people who used language. That they are accompanied by quantities of worked ochre is a comfort to some other arguments about language origins.

In Mellars' terms, the MSA tools were visually distinctive, repetitive, and standardised. The crucial issue is still whether there were earlier artefacts which fulfilled those criteria, and, if so, whether they indicate language use. The most obvious case which still needs discussion is the handaxes called Acheulean which seem to dominate the production of stone tools over 1.5 million years and over most of the world occupied by hominins over that period. One of the novelties of the last ten years has been the increasing recognition that handaxe forms are found in east Asia. I will outline some of Wynn's arguments about the cognitive significance of handaxes.

Handaxes/bifaces

At a conference in 1987, Dibble demonstrated that there were two common problems with the interpretation of Acheulean handaxes. First, there have been claims that there were many different types of handaxes. Dibble showed that the divisions were more apparent than real, resulting from the partitioning of continuous variation. There have been other demonstrations in the last decade that archaeologists are particularly inclined to partition continuous variation in stone artefact assemblages, with misleading results. Second, there was some element of the apparent standardisation of form of handaxes which resulted from simple procedures by archaeologists. I will present some more data in relation to this argument, and reinforce the suggestion that a relatively small number of constraints could determine the apparent standardisation of Acheulean handaxes.

In light of this claim, I recently began to look at cores from the oldest layers at the site of Tabun, from the meticulous excavations by Art Jelinek. If I am right in the contention that the form of handaxes is an outcome of the application of a small number of constraints in knapping, then some of the features of handaxes should be present on the non-handaxe cores in the related assemblage. I will show how my analysis of material from Tabun confirms this. In addition, Tabun provides good evidence confirming the claim by Bradley and Sampson that handaxe forms might be a stage in reduction sequence.

Arguments about the uniformity of handaxes over such vast times and regions, bring into question the issue of standardisation which is such an important part of Mellars's definition. I will discuss the arguments about standardisation by Chase, by Chazan and by Kuhn. The fundamental point here is that there are two major reasons why artefacts are standardised: either there is a technical, procedural or cognitive limit on the possible outcomes, or a single option of many available was chosen.

One of the significant changes over the last ten years has been the clarification of the chronology of Middle Pleistocene archaeological sites. In particular, it is now quite clear that there is no simple sequence from Oldowan to Acheulean to a late Mousterian and then the Upper Palaeolithic characterised by stone tools made on parallel-sided blades. I believe that the new chronology requires the abandonment of the notion that there is a continuous 1.5 million year long Acheulean tradition. Instead we should recognise that there are relatively small numbers of possible outcomes from stone-knapping. There is some confirmation of this alternative view from the discovery of bifacially flaked handaxe-like cores in Australia, a region colonised after the end of the Acheulean from a region where handaxes were said never to have been made.

Levallois technique

There remains the issue of the Levallois technique, if that had not been abolished by my argument in 1990. The appearance of "Levallois cores" in the early stone industries of Olduvai Gorge would be a product of the limited number of outcomes from knapping. Van Peer's work provided new evidence in support of my position. This work showed that, when conjoining flakes from single knapping events, non-Levallois flakes were very often missing (presumed used). This was confirmation of the interpretation of use-wear studies by Beyries that unretouched non-Levallois flakes were more often used than unretouched Levallois flakes. The argument that Levallois technique implies a long time-depth of intentionality is looking very shaky.

Nevertheless, the Levallois technique has come back into prominence because of argument by Foley and Lahr that it provides insight into the evolutionary history of hominins. I will outline the difficulty of accepting the position adopted by Foley and Lahr.

Later stone industries

There is new evidence and argument which could be presented about the earliest industries which I do regard as showing "imposed form" – those from the southern African Middle Stone Age (MSA) – and about the contemporary industries in Europe and adjacent areas. Although some of these industries may represent the emergence of modern human behaviour, I will confine my remaining comments to the question of the importance of long, narrow flakes known as blades.

By commenting on the various industries in Africa, Europe and the east Mediterranean which appear to have long parallel-sided flakes (much) earlier than the Upper Palaeolithic, I will argue that the appearance of "blades" alone cannot be used as an indicator of technical or cognitive abilities of hominins.

Discussion

The evidence I present draws attention to two issues: the problems arising from our language-based practice of naming categories; and the need to think more fundamentally about the factors contributing to variation and to lack of variation in stone tool assemblage.

The problem of naming is well illustrated by the various examples of the partitioning of continuous variation in identifying tool "types", but it applies equally to the naming of the assemblages of tools as "industries". Importantly, it is highly likely that issues of continuity or discontinuity in patterns of producing stone artefacts result partly from different approaches to giving names to categories in analysis. This problem is best illustrated by the continuing importance given to the issues surrounding the larger category "Upper Palaeolithic". In particular, there are issues about whether one industry or another belongs to the Middle or the Upper Palaeolithic, and then whether the Upper Palaeolithic necessarily indicates modern human behaviour. Arising from this is a widespread perception that somehow or other the Upper Palaeolithic indicates the emergence of language (despite the overwhelming evidence to the contrary). Finally, the issue of the reality of the names invented by archaeologists and palaeoanthropologists is starkly revealed by the problems of naming fossil hominins. One important issue here is that the naming also involves the partitioning of continuous variation – producing a conflict with an evolutionary approach to changing hominin morphological variation. In addition, the naming practices of archaeologists and palaeoanthropologists raise the question of whether a single stone industry could be made by more than one hominin species, as appears to be the case in the east Mediterranean where Mousterian industries were made by both Neandertals and anatomically modern humans.

Explanations of stone artefact variation need to take into account all of the various factors which create variation between stone artefacts, as well as the many different factors which cause there to be little variation. These factors may be mechanical, mental or cultural. Mechanical constraints include the nature of the raw material, the requirements of knapping, knapping technique, and the requirements and contingencies of use. Mental factors include the absence of conceptualisation as well as the ability to conceive of new forms and how to make them. Constraints due to the culture of the knappers involve the fundamental issue of the context of learning to knap, as well as the cultural significance attached to style and any cultural constraints on producing new forms. It seems highly unlikely that, if a single stone tool industry was made by two different species, we should regard the lack of variation in that industry as the product of culturally determined intentions.

Conclusion

The study of stone artefacts clearly has huge implications for understanding the origins of culture, particularly if we can tie down the appropriate criteria for recognising that the forms of artefacts were determined primarily by the self-conscious choices of the knappers. In this context, we will also need to understand not only how there can appear to be imposed form in non-language using industries (as in the Acheulean), but how there can appear not to be imposed form in industries made by people who certainly used language (as inAustralia). The story is undoubtedly more complicated than the text-books tell us, but we are beginning to get answers despite the difficulties of interpretation. To do this, we must understand the extent to which some of the story was determined by assumptions about the way hominin behaviour changed in the Pleistocene.

Studying language evolution with self-organisation and phonetics

Bart DE BOER

AI-Lab
Vrije Universiteit Brussel
bartb@arti.vub.ac.be

The work described in this extended abstract is a continuation of the work that was presented at the previous evolution of language conference. I will start with a short recapitulation of the philosophy and hypotheses behind the "self organisation in vowel systems" simulation. In order to support the claims made, new results of the experiments on self-organisation and their relation to vowel systems in real human languages will be presented. But the goal of this abstract is slightly more ambitious (and speculative) than that. I will argue that a good understanding of the role of self-organisation is crucial when trying to explain language evolution. I will also argue (not without self-interest) that phonetics and phonology are ideal aspects of language for testing theories on the role of functional adequacy and self-organisation in explaining the evolution of language.

Explaining the evolution of language is a daunting task. Not only is language a very complex and multifaceted system, it is also a cultural phenomenon. This has several implications. Because of its many-sidedness, it will not be possible to find a single explanation or mechanism for the evolution of language. Because language is a cultural phenomenon, it is very hard (although not altogether impossible) to find evidence of its evolution. Furthermore, there will be cultural evolution and cultural variation of language co-occurring with its actual physical evolution.

However, this cultural aspect of language actually makes it possible to simplify the task of finding an evolutionary account for the origins of language. If we are able to separate the phenomena that can be explained through cultural mechanisms from the phenomena that must be explained by biological evolution, the complexity of the evolutionary account can be reduced. Cultural mechanisms will be subject to functional pressure. Forms of language that are difficult to learn, produce or process will be disfavoured with respect to forms of language that are easier. The importance of the link between functional factors and evolutionary processes has already been stressed by other researchers (e.g. Kirby 1999; Hurford, *to appear*). The importance of purely cultural mechanisms has been stressed by e.g. Steels (1998).

In reducing the amount of linguistic phenomena to be explained by evolution, we therefore have to look for possible functional constraints that are relevant to these phenomena. These phenomena might turn out to be emergent properties in the population due to the functional constraints. Emergence of phenomena in a population is often called self-organisation. The outcome of self-organisation is very hard to predict directly from the known constraints on the individuals in the population. The best way to test the influence of these constraints is therefore either to observe the population over a very long period of time (as in historical linguistics) or to build a computer simulation of the phenomena. Computer simulations are of course faster and easier to change. An extra advantage of computer simulations is that one can abstract away from unnecessary detail. If a certain phenomenon is observed even in a simplified computer simulation as the result of self-organisation in the population, it is likely that it will also be emergent in human language. Of course, computer simulations can only be used to test a theory. The real acid test for the theory is of course its ability to predict data from human languages.

The example that will be elaborated in this extended abstract is that of the properties of human vowel systems. Certain of these properties were caused by biological evolution. The shape of the vocal tract for example has clearly been determined by biological evolution. However, other aspects of vowel systems might not be determined biologically. The preferences for certain vowels over others, for symmetries and for certain inventory sizes are, according to some researchers (starting with Chomsky & Halle, 1968) due to biologically determined properties of the human brain. Other researchers (e.g. Liljencrants & Lindblom 1972) have proposed functional criteria for explaining these phenomena. In this extended abstract, and in previous papers (de Boer, 1999, *to appear*) I show that the combination of functional constraints and self-organisation is sufficient to predict the shape of vowel systems.

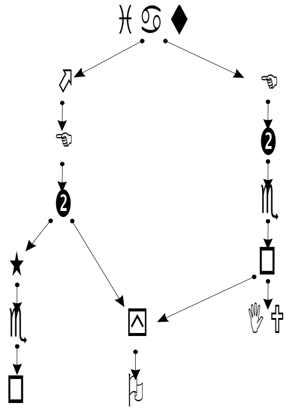


Figure 2: ***Vowel system hierarchy according to Crothers (1978).***

The computer simulation that was used to test the theory is based on a population of agents that are able to produce, perceive and store vowels in a human-like way. Production is based on an articulatory vowel synthesiser and perception is based on a distance function that uses the first formant and the effective second formant of the signal. The vowels are stored as prototypes. The interactions between the agents, so-called imitation games, consist of an agent picking a random vowel from its repertoire and producing it (imperfectly) with its synthesiser. Another random agent then imitates this sound, using its perception function, its list of vowel prototypes and its synthesiser. The first agent then checks whether the signal it perceives is analysed as the same vowel as the one it originally produced, and communicates the success or failure of the game using non-verbal feedback. The success of the game thus depends not only on the absolute qualities of the vowels used, but also on the repertoires of the agents that participate in the game. Depending on the outcome of the game, the agents update their repertoires by removing, adding, shifting or merging vowel prototypes, using only local information and a rather general learning mechanism (selectionism combined with hill-climbing). The agents and games are described in much more detail in (de Boer, 1999).

Assessing the realism of the vowel systems that emerge is best done by classifying them according to the same principles with which human vowel systems are classified. One can then check whether the emerged vowel systems show the same universal tendencies as human vowel systems. Crothers (1978) has made a study of human vowel systems and derived a list of universals from that. His universals can be represented as implicational universals, as shown in figure 2. An arrow in the figure indicates that if a human vowel system has a given vowel, it will

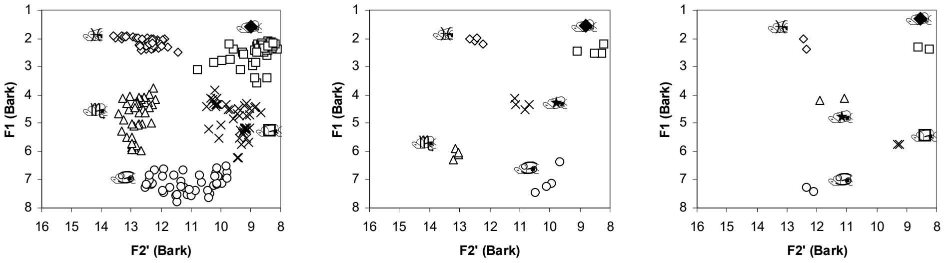


Figure 3: ***Classification of five vowel systems.***

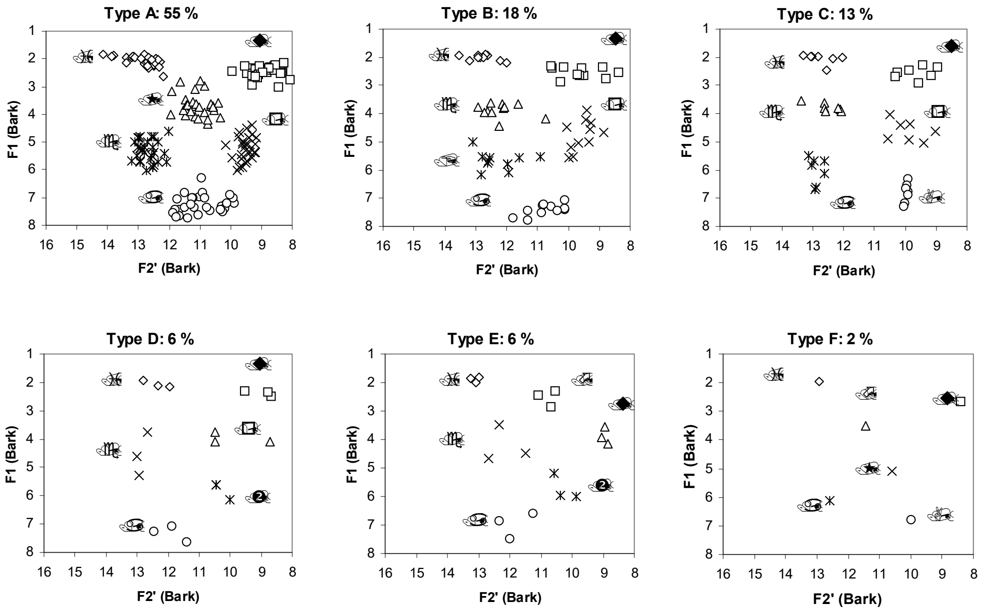


Figure 4: ***Classification of six vowel systems.***

also contain the vowels from which there are arrows pointing towards it.

An example of emerged vowel systems with five vowels is given in figure 3. These figures were made by creating 100 random populations with 20 agents. Then 25 000 imitation games were run, for a given parameter setting (15% acoustic noise). Then the populations with systems with five vowels that emerged were selected (there were 49 of such systems). From each population, a random agent was selected, and its vowel prototypes were plotted in the acoustic space consisting of the first formant and the effective second formant. The scales of the figures are logarithmic. Three different types of five-vowel systems emerge. The symmetric system appears in 88% of the cases, the system with more front vowels in 8% of the cases and the system with more back vowels in 4% of the cases. Except the rightmost systems, all emerged systems followed Crothers' universals. Furthermore, the percentages of occurrence match very well with those of human vowel systems (as measured by Schwartz *et al.* 1997) 88%, 8% and 4% versus 89%, 5% and 5% (for 109 languages). Similar good fits were found for systems of 4–8 vowels. For systems of 3 vowels and for systems with 9 vowels, the fit was less good. In order to further illustrate this, the 54 emerged systems for six vowels are shown in figure 4. Percentages from (Schwartz *et al.* 1997) were type A 55% type B: 20%, type C: 5%, type D: 7% and type E 13% (of 60 6-vowel systems in their data).

Systems of different sizes were obtained by running the simulation with different settings for acoustic noise. The more noise, the fewer vowels a system would contain. A very strong universal for human vowel systems is a tendency towards systems with five vowels. It was checked whether self-organisation in a population also showed a tendency towards particular sizes. For this reason, the system was run for noise levels between 9% and 28%. Below 9% no more vowel systems with less than eight vowels emerged. Above 28% no vowel systems with more than two vowels would emerge. The histogram for the different sizes is given in figure 5. These numbers have been calculated with a perception function that was slightly different from the perception function that was used in the previous experiments. The original perception function contained a small bug. This bug resulted in size four occurring the most frequently. As can be seen in the figure, with the bug repaired, size five occurs most frequently. The peak is not so pronounced as for human vowel systems, but it does occur.

These results show that systems of basic vowels can be predicted very well as the result of self-organisation under constraints of perception and production, without recourse to specialised innate structures determined by biological evolution. Of course, for larger vowel repertoires and for consonants and syllables, other factors start playing a role. Nevertheless the research presented here is a good demonstration of the power of self-organisation in predicting real linguistic phenomena.

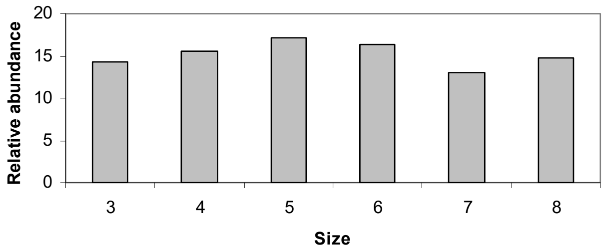


Figure 5: Relative abundance of vowel systems

Unfortunately, implementing similar simulations for more complex sounds turned out to be more difficult than expected. Still, investigating the role of functional constraints and self-organisation in phonetics and phonology is a promising way for gaining more understanding of the emergence of language. More complex speech units, such as syllables and words show compositionality, just as sentences do, but in a more limited way. Furthermore, functional constraints in phonetics are physical, and have relatively little to do with what Kirby (1999) calls processing. We know relatively little about the way the human brain processes language. Building computer simulations and verifying their results in the case of syntax, although very interesting from a linguistic point of view, involves a relatively large number of uncertainties and arbitrary decisions. On the other hand, many aspects of phonetics and phonology are much easier to measure and to model, as they involve physical signals. Therefore realistic computer simulations are easier to build and verify in this case, as has been shown by the results in this paper. Self-organisation and phonetics are therefore promising means for investigating the origins of human language.

References

Chomsky, N. & M. Halle (1968) *The sound pattern of English*, MIT Press, Cambridge, Mass.

Crothers, J. (1978) Typology and Universals of Vowel systems. In Joseph H. Greenberg, C. A. Ferguson & E. A. Moravcsik (eds.) *Universals of Human Language*, Volume 2 Phonology, Stanford: Stanford University Press pp. 93–152.

de Boer, B. G. (1999) *Self-organisation in vowel systems*, PhD. thesis Vrije Universiteit Brussel.

de Boer, B. G. (*to appear*) Emergence of sound systems through self organisation. In C. Knight & J. Hurford (eds.) *The evolution of language*, Cambridge: Cambridge University Press.

Hurford, J. R. (*to appear*) Social transmission favours linguistic generalization. In C. Knight & J. Hurford (eds.) *The evolution of language*, Cambridge: Cambridge Univ. Press.

Liljencrants, L. & B. Lindblom (1972) Numerical simulations of vowel quality systems: The role of perceptual contrast. *Language 48* pp. 839–862.

Kirby, S. (1999) *Function, Selection, and Innateness: The emergence of Language Universals*, Oxford: Oxford University Press.

Schwartz, J., L. Boë, N. Vallée & C. Abry (1997), Major trends in vowel system inventories. *Journal of Phonetics 25*, pp. 233–253

Steels, L. (1998) Synthesising the origins of language and meaning using co-evolution, self-organisation and level formation. In J. R. Hurford, M. Studdert-Kennedy & C. Knight (eds.) *Approaches to the Evolution of Language*, Cambridge: Cambridge University Press pp. 384–404.

Inheritance of language deficits in schizophrenia: Relationship to the evolution of human linguistic ability

Lynn E DeLisi

Department of Psychiatry, SUNY Stony Brook
Stony Brook, New York 11794
LDeLisi@notes.cc.sunysb.edu

The language abilities of modern humans are thought to be clearly distinct from non-human primates (Bickerton, 1990). Crow (1997, 1998a) has proposed that language disturbance is central to the genetic etiology of schizophrenia and may be the critical characteristic in the evolution of Homo sapiens. Oral soliloquys have been analyzed from schizophrenic patients at the onset of their illness, controls and families with a high density of illness. Anomalies in Bickerton's 5 distinctly human language devices were examined (sequential order, null elements, arguments of verbs, mechanisms for expansion and grammatical correctness) using a structured scoring format on the transcribed passages. Analyses show reduced use of clausal embedding in chronic patients that is not present at the time of the first-episode. These data are preliminary and need to be replicated in a larger sample of subjects. However, they are consistent with previous literature on language in patients with schizophrenia and further show that there is a familial component to these measures. Thus, these findings suggest that deficits in specifically human aspects of language may be related to the genetics of schizophrenia.

Background

Schizophrenia is a lifetime brain disorder expressed by thought, speech perception and organizational problems. It occurs in every human population (world-wide) and family, twin and adoption studies show a genetic predisposition for schizophrenia. Explanations for the thought and speech disorder of schizophrenia include that of (1) "dysfunctional executive control" (frontal lobe; e.g. Chaika, 1990; Morice,1995). (2) working/semantic memory deficits (e.g. Grove and Andreasen, 1985, Mortimer et al., 1995), and the extreme end of the variation in bilateral neural organization of the uniquely human components of language (Crow, 1998).

Bickerton (1990) proposes 5 unique characteristics of human language: (1) Differences in the superficial order of constituents, (2) use of null elements, (3) subcategorized argument structure of verbs, (4) mechanism for the expansion of utterances, and (5) grammatical items. It is hypothesized that the language disturbances of schizophrenia are related to defects in one or more of the uniquely human features of language and that the inheritance of the variation in these abilities is related to the inheritance of schizophrenia.

A relationship also exists between brain structural and functional asymmetries and both language abilities and schizophrenia (Corballis, 1991; Crow, 1998a; Geschwind and Galaburda, 1987). Furthermore, reduced functional asymmetry (hand skill) relate to reduced verbal ability (Crow et al., 1998). It has been proposed that the core clinical symptoms of schizophrenia relate to an underlying disorganization of language and its asymmetric perception and production (Crow, 1998b). Finally, some reduced posterior cerebral asymmetries present in schizophrenia are shown to be familial (DeLisi et al., 1997).

In previous studies (DeLisi et al., 1997; Shedlack et al., 1997) we found significantly more morphological errors and less clausal embedding in 1st episode patients compared with 52 controls. No associations were found with brain structural asymmetries. When ill and well members from 19 families with 2 or more siblings with schizophrenia were compared with 17 unrelated controls, soliloquys revealed that both ill and well siblings from these families had significantly less of one type of clausal embedding than controls, although the degree of embedding was not correlated within families and ill siblings did not have less embedding than their well siblings. The present study expands on our previous measures of language performance.

Methods

First-episode patients with schizophrenia, families with 2 or more siblings with schizophrenia and unrelated controls had their speech recorded and later transcribed (without punctuation) while an examiner asked each to use his/her imagination to describe what is happening in 7 pictures taken from the Thematic Aptitude Test (TAT). These pictures ranged from disturbing emotional scenes to tranquil scenery.

The following are examples of responses to two disturbing scenes, ones which tended to stipmulate more disorganized responses from patients.

> *"...I'm afraid to say something lude but it could connotate that be a grandfather coming in to touch his precious grand pre-post pubsescent granddaughter and yet he has sorta of a halo around him so it could be an angel or maybe I am saying that because there was a chance I was sexually abused which I have been exploring in the last year or two the innocent part of me says its an angel visiting the virgin Mary and the other part of me says its incest the beginning of an incestral relationship that has been going on for a while and not nervous so I'm alittle upset looking at that one..."*

> *"...it looks like he is tired he wants to go to bed he's tired I guess he has his hand over his face like he is really tired and he is aggravated he could be aggravated he is just getting up out of bed putting his clothes on getting out of bed getting something to drink or something he could be getting up to go to work in the morning you know his wife or his girlfriend is sleeping or maybe the wife dies or something he probably is going to go to work..."*

In addition , each subject was asked to describe the steps taken to get up in the morning and get ready for an appointment, from waking-up to arrival at the destination (sequential ability).

Laterality was assessed by (1) a hand use questionaire (Annett 1967) and (2) by a relative hand-skill timed "dot" test (Tapley and Bryden, 1985). The following items were scored: total # of words, total # of sentences, SYNTAX: Order (# of awkward sentences, # with incorrect word order), # of null elements, # sentences missing arguments to verbs, mechanisms for expansion (# of conjoined clauses and # embedded, # grammatical mistakes, SEMANTICS: Misuse of Words (#), neologisms (#), MISCELLANEOUS: Appropriateness of content and logical order to sequential task and amount of detail.

Two separate 2-way ANOVA's were performed (1) For the 1st epsode cases compared to chronic patients and controls covarying sex and age (2) For the 12 families only: with sex, diagnosis and family membership as factors, covarying for age.

	1st Episode Patients Mean +/-SD (N=9)	Chronic Pati (N=29)	Controls (N=12)	F(Dx)=	p<
Dependent Variable					
Age	23.4+/-5	33.8+/-8	32.6+/-7		
#Words	594+/-406	530+/-267	862+/-514	3.21	0.05
# Awkward Sentences	2.7+/-3	3.3+/-4	1.8+/-1	0.61	0.55
# Incorrect Order	0.3+/-.7	0.3+/-.9	0	1.23	0.30
# Null Elements	6.8+/-5	6.2+/-4	11.0+/-8	0.46	0.64
# Grammatical Mistakes	1.8+/-2	1.3+/-2	0.9+/-2	1.50	0.24
# Missing Arguments	2.0+/-2	1.1+/-2	0.9+/-1	0.31	0.73
# Conjoined Clauses	29.9+/-24	20.2+/-16	45.8+/-27	5.92	0.005
# Embedded Clauses	6.9+/-7	5.6+/-5	12.2+/-12	2.68	0.08
Misuse of Words	0.44+/-.5	0.38+/-.6	0.17+/-.4	0.54	0.59
Neologisms	0	0.14+/-.6	0	0.30	0.75
Poor Over All Content	11%	41%	0	3.32	0.05
Vagueness	33%	28%	0	2.11	0.13
Sentences Lack Connectivity	22%	14%	0	0.95	0.39
Incomplete Sentences	44%	38%	8%	1.09	0.35
Numerous False Starts	22%	35%	25%	0.80	0.46
Sequential Task					
Poor Content	20%	46%	9%	1.44	0.25
Poor Order	20%	50%	18%	1.43	0.25
Handedness (#R)	80%	96%	100%	0.60	0.55
Handskill (R/L)	1.3+/-.6	1.3+/-.4	1.3+/-.3	0.30	0.74

Note: There were no significant sex or sex by diagnosis effects.

Table 1. *Unrelated Patients vs. Controls*

	Family		Diagnosis	
	F=	(p<)	F=	(p<)
Dependent Variable				
# Words	1.44	0.31	0.74	0.42
# Awkward Sentences	0.89	0.59	0.46	0.52
# Incorrect Order	0.58	0.80	0.94	0.36
# Null Elements	7.87	0.006	0.04	0.84
# Grammatical Mistakes	1.08	0.48	0.22	0.66
# Missing Arguments	0.19	0.99	0.06	0.82
# Conjoined Clauses	3.42	0.05	1.03	0.34
# Embedded Clauses	5.19	0.01	1.28	0.30**
Misuse of Words	0.84	0.62	3.17	0.12
Poor Over All Content	0.88	0.59	0.20	0.66
Vagueness	0.68	0.73	1.22	0.30
Sentence Connectivity	1.48	0.30	0.23	0.65
Incomplete Sentences	5.13	0.01	0.13	0.73
Sequential Task				
Poor Content	5.86	0.01	0.12	0.74
Logical	0.31	0.95	0.62	0.46
Handedness	1.20	0.43	0.002	0.97
Handskill	0.51	0.83	0.01	0.91

**A significant family by diagnosis effect covarying for total # of words was present for embedded clauses (F=5.87, p=0.02).

Table 2. *12 families with 2 siblings, both of whom are diagnosed with schizophrenia (ill and well relatives compared and a within versus among family analysis performed.*

Results

Table 1 summarizes the data on the 1st episode patients compared with controls, while table 2 summarizes the family data. Reduced percentages of conjoined and a trend for reduced embedded clauses were present in the schizophrenic patient groups compared with controls.

Summary

The Characteristic thought disorder present in patients with schizophrenia can be shown to be linguistically deviant. Particularly, chronic patients (ill for at least a few years) had reduced use of conjoined and embedded clauses in their speech, used fewer words overall, more inappropriate content and less connectivity between sentences, than controls. A lack of detectable language abnormalities at the time of the 1st episode suggests progressive development after the onset of illness. In addition, a familial basis for clausal mechanisms of expansion is suggested from analyses of the above families.

Thus, the uniquely human aspects of language, particularly mechanisms for expansion are shown to be reduced in patients with schizophrenia and also shown to be familial and related to illness within families.

References

Annett M (1967) The binomial distribution of right, mixed, and left handedness. Quarterly Journal of Experimental Psychology 19: 327-333.

Bear DM, Schiff D, Saver J, Greenberg M, Freeman R (1986) Quantitative analysis of cerebral asymmetry; fronto-occipital correlation, sexual dimorphism and association with handedness. Archives of Neurology 43: 598-603.

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

Chaika EO (1990) Understanding Psychotic Speech. Charles C. Thomas Publisher, Springfield, Illinois

Corballis MC (1991) The Lopsided Ape: Evolution of the Generative Mind, Oxford University Press.

Crow TJ (1997) Is schizophrenia the price that Homo sapiens pays for language? Schizophrenia Research 28: 127-141.

Crow TJ (1998a) Sexual selection, timing and the descent of man: A theory of the genetic origins of language. Current Psychology of Cognition 17: 1079-1114.

Crow TJ (1998b) Nuclear schizophrenia symptoms as a window on the relationship between thought and speech. British J of Psychiatry 173: 303-309.

Crow TJ, Crow LR, Done DJ, Leask SJ (1998) Relative hand skill predicts academic ability: global deficits at the point of hemispheric indecision. Neuropsychologia, in press.

DeLisi LE (1997) Is schizophrenia a lifetime disorder of brain plasticity, growth and aging? Schizophrenia Research 23: 119-129.

DeLisi LE, Sakuma M, Kushner M, Finer DL, Hoff AL, Crow TJ (1997) Anomalous cerebral asymmetry and language processing in schizophrenia. Schizophrenia Bulletin 23: 255-271.

Geschwind N, Galaburda AM (1987) Cerebral Lateralization, MIT Press, Cambridge, Mass.

Grove WM, Andreasen NC (1985) Language and thinking in psychosis. Is there an input abnormality? Archives of General Psychiatry 42: 26-32.

Morice R (1995) Language impairments and executive dysfunction in schizophrenia. In: Speech and Language Disorders in Psychiatry (A Sims, editor), Gaskell Press, pp. 57-69.

Mortimer A, Corridan B, Rudge S, Kho K, Kelly F, Bristow M, Hodges J (1995) Thought, speech and language disorder and semantic memory in schizophrenia. In: Speech and Language Disorders in Psychiatry (A Sims, editor), Gaskell Press, pp. 70-80.

Shedlack K, Lee G, Sakuma M, Xie S-H, Kushner M, Pepple J, Fineer DL, Hoff AL, DeLisi LE (1997) Language processing and memory in ill and well siblings from multiplex families affected with schizophrenia. Schizophrenia Research 25: 43-52.

Tapley S and Bryden M (1985) A group test for the assessment of performance between the hands. Neuropsychologia 23: 215-221.

A short introduction to GRAEL
Grammar Adaptation, Evolution and Learning

Guy DE PAUW & Walter DAELEMANS

CNTS
UIA - University of Antwerp
depauw@uia.ac.be daelem@uia.ac.be

In this paper, we describe work in progress on the GRAEL (GRammar Adaptation, Evolution and Learning) environment, in which we aim to develop a suitable environment in which natural language grammars can interact and co-evolve according to underlying principles of genetic programming. From an engineering point-of-view, the GRAEL environment offers a grammar optimisation method that will help alleviate problems of grammar coverage and sparse data. The main focus is however more theoretical in nature: the modularity of the GRAEL system allows us to investigate the dynamics of grammar evolution and adaptation, as well as offering a wide range of extensions for future research.

Related research

The problematic domain of syntax has been of particular interest to NLP-researchers, prompting them to develop syntactic parsers that range from attempts to translate Chomskian insights into a computational context to highly efficient implementations of rather ad hoc grammatical systems.

So far, however, little syntactic ground has been covered in the evolutionary computing paradigm. With the notable exception of Losee's LUST system [Losee1995], in which an information retrieval system is powered by genetically evolving grammars, most GA syntactic research has focussed on non-linguistic data.

Smith and Witten [Smith and Witten 1996] use a Genetic Programming-related algorithm to adapt a population of hypothesis grammars towards a more effective model of language structure. Smith and Witten use tree structures as their syntactic representation, with nodes labelled as either AND or OR, which can be interchanged during mutation. Fitness of an individual is measured by counting its grammar size and its ability to parse test strings.

Wyard [Wyard 1991] and Blasband [Blasband 1998] also use Genetic Programming to induce and optimise grammars.

Closely related to this line of work, is Antonisse's effort [Rawlins 1991] in which grammar-based crossover is implemented, a feature lacking in [Smith and Witten1996]. Crossover occurs by randomly splitting a sentence into 2 sentence fragments and interchanging them with 2 fragments of another sentence.

Antonisse tries to develop a grammar-based genetic algorithm for all types of grammars, rather than to investigate the behaviour of linguistic grammars in an evolving context, Smith and Witten's linguistic analysis of the generated data suffers from the fact that their linguistic representation is too weak to offer any insight in this matter. Losee's LUST-system, which is powered by genetically evolving grammars, is very much applied to its information retrieval task and does not offer any theoretical insights on the dynamics of grammar adaptation itself.

GRAEL

The GRAEL environment is explicitly geared towards the goal of investigating the behaviour of grammars for natural language in an evolutionary context. Presented here is the grammatical backbone of the environment, as well as the general methodology, prototypical results and future extensions to this line of research.

Taken from the field of statistical parsing, Data Oriented Parsing (DOP, [Bod 1998]) fits the evolutionary computing paradigm very well because of its emphasis on substitution of parts of syntactic trees (so-called *substructures*), representing derivations, and its absence of an explicit abstract rule-based grammar. Parsing a sentence with the DOP-model involves finding the most probable combination of substructures, stored in a *treebank*, induced from an annotated corpus.

In the GRAEL-environment, the principles of the DOP-model are translated as follows: an annotated treebank i.e. a collection of pre-parsed sentences (currently ATIS [Santorini 1993]), is distributed over a number of individuals. An individual therefore consists of a number of tree structures, representing the language experience of a language user. The grammars are used to power a memory-based implementation of the DOP-model [De Pauw 2000] for parsing:

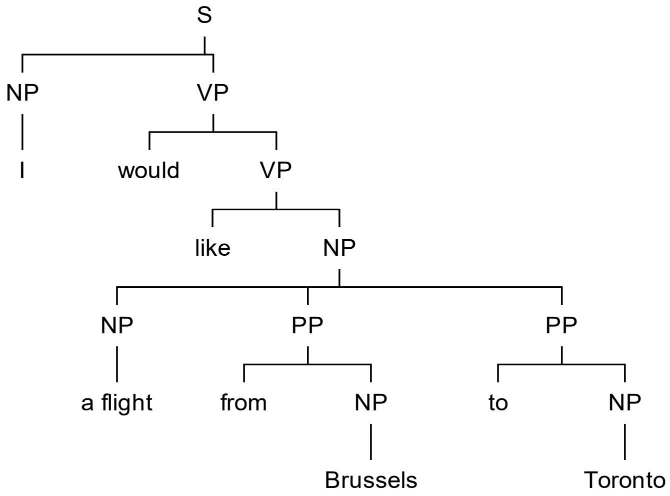
A weighted combination of metrics expresses an individual's fitness:
efficiency of the grammar: the number of nodes in an individual's "forest" (collection of trees), translating the notion that a smaller grammar is a better grammar
the *accuracy* with which an individual's forest is able to parse a held-out set of test strings
language games (methodology/concept borrowed from ARTI-research [Steels 1996]):
the ability with which the grammar induced from an individual's forest is able to parse another individual's sentences (*understanding*)
the success-rate with which another individual is able to parse the individual's generated sentences (*understandability*)

Neo-Darwinism is introduced by preferring fitter individuals for "reproduction". Reproduction currently does not create new individuals as such, but extends an individual's knowledge through *crossover* and *mutation* operations.

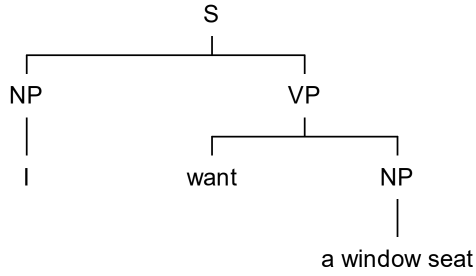
Crossover

One individual is randomly selected for "reproduction". Crossover occurs when a random node in one of the tree structures of this individual is exchanged for a random node (carrying the same label) of another individual. For example:

sentence n1 (individual 1)

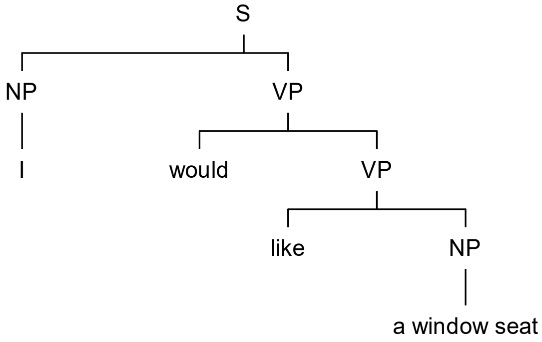


sentence n2 (individual 2)

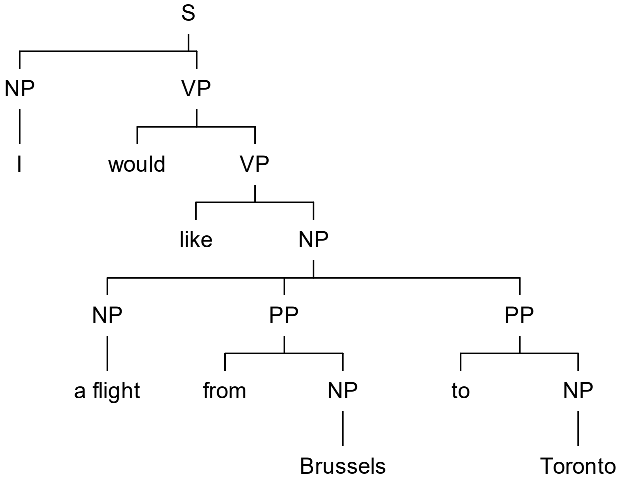


Individual 1 may now exchange the NP *a flight from Brussels to Toronto* with individual 2's NP *a window seat*, yielding the following two new tree structures (sentences) for the respective individuals:

sentence n1' (individual 1)



sentence n2' (individual 2)



In the above example, two new grammatical sentences have been generated. Exchanging VP-nodes however would yield the sentence *I want like a window seat from Brussels to Toronto*. The crossover operation yields sentences, varying in grammaticality. Fitness functions (2) and (3) will express a preference for individuals containing grammatical tree structures. Furthermore, when an individual's forest reaches a certain threshold, it is pruned using probability and entropy of the tree structures as relevant metrics.

Mutation

[De Pauw 2000] has also reported problematic parsing on the NP-level with the ATIS-corpus. When using a blind-testing method (testing on held-out test data), we notice that most of the unparsable sentences are problematic on the lowest NP-level, i.e. NP's only consisting of lexical items. An NP like “Flight 447 A K nine” would have the following NP-structure (NP → noun number noun noun number), a constituent understandably not found in the training data.

Through mutation, items (i.e. branches) are added or deleted from NP-structures in the constituents. This causes a lot of grammatical overhead in the environment, but the new structures are maintained or discarded, using the above-mentioned fitness functions.

Prototypical Results

The initial corpus-based experiments are based on the very small ATIS-corpus. The grammars distributed over the individuals are therefore extremely limited in generative power. The “reproduction” and interaction with other individuals, however, extends the grammar very fast and adapts it to a specific use, be it understanding other individuals, or parsing the test set. We will present how well the environment is able to do just that and we will present how a faulty grammar can improve itself by interacting with other (faulty) grammars in an evolutionary context.

We also evaluate this approach by comparing the accuracy on the held-out testset of the “full” grammar (i.e. the grammar induced from ALL individuals) to the accuracy of the grammar of the fittest individual. When we limit “reproduction” to crossover, we notice a tendency in the individuals to level around 250 rules (from a total rules of 2600 rules), indicating that the full corpus contains many marginal rules not needed for efficient and accurate parsing. The mutation operator rapidly extends grammar size at the cost of parser efficiency, yet the desired effect of grammatical smoothing takes place with some individuals developing grammars that parse sentences that the full grammar was not able to.

More theoretically, we will look at the evolution and intricacies of each individual’s syntactic component for particular (combinations of) fitness-functions. We will describe what a grammar looks like when it’s been genetically improved for efficiency (fitness function (1)), for accuracy (2), understanding (3a) and understandability (3b).

Future work and directions

A proper implementation of reproduction (in which new individuals are created) will be added to the GRAEL environment. Also, the limitation on crossover nodes (cf. only nodes with the same label can be exchanged) will be relaxed. Mutation will be extended to other categories than NP.

With respect to the data, we will be using other corpora in the GRAEL-environment, as well as combining different corpora, in order to investigate notions of *understanding* and *understandability* between individuals made up of tree structures from different corpora.

The corpus-based experiments however only serve as a first test drive of the GRAEL environment and the main focus in future work will be on the unsupervised emergence of grammatical principles. A grammar development module will serve as a bootstrap for the GRAEL framework, so that given a set of random strings, a grammatical systems can be created which will evolve using the principles mentioned above.

The current corpus-based experiments serve the initial purpose to understand the mechanics of grammatical evolution, as well as give us pointers for the above-mentioned extensions to the GRAEL-environment

References

[Antonisse 1991] Antonisse, H.J.. 1991. A Grammar-Based Genetic Algorithm. In [Rawlins 1991]

[Blasband 1998] Blasband, M. 1998. Gag genetic algorithms for grammars. Technical Report Compuleer

[Bod 1998] R. Bod, 1998. Beyond Grammar: An Experience-Based Theory of Language, CSLI Publications, Cambridge University Press. 184pp.

[De Pauw2000] De Pauw, G. 2000. Probabilistic Parsers - Pattern-matching aspects of Data-Oriented Parsing [in Dutch]. Antwerp Papers in Linguistics (forthcoming)

[Losee11995] Losee, R.M. 1995. Learning syntactic rules and tags with genetic algorithms for information retrieval and filtering: An empirical basis for grammatical rules. In Information Processing and Management.

[Rawlins 1991] Rawlins, G.J.E. 1991 Foundations of genetic algorithms. M. Kaufmann: San Mateo, Calif.,1991

[Smith and Witten 1996] Smith, T.C. and I.H. Witten, 1996. Learning Language using genetic algorithms. In Connectionist, statistical and symbolic approaches to learning for natural language processing.

[Santorini 1993] Santorinit et al: *Building A Large Annotated Corpus Of English: The Penn Treebank.* Computational Linguistics, Vol. 19, 1993.

Steels, L. (1996) Synthesising the origins of language and meaning using co-evolution, self-organisation and level formation. In: Hurford, J., C. Knight and M. Studdert-Kennedy (ed.) (1997) Evolution of Human Language. Edinburgh Univ. Press. Edinburgh.

[Wyard 1991] Wyard, P. 1991. Context free grammar induction using genetic algorithms. In *Proceedings of the 4th International Conference on Genetic Algorithms*, 514-518.

Two stages in the evolution of language use

Jean-Louis DESSALLES

Ecole Nationale Supérieure des Télécommunications, Paris
dessalles@enst.fr

Introduction

The study of language use, usually called *pragmatics*, reveals that the competence of speakers is not monolithic. It can be split into two quite distinct behaviors. The first one deals with salient events; the second one deals with problematic situations. We claim that the second ability emerged long after the first one in hominid evolutionary history. A consistent scenario is that communication about salient events is what the protolanguage hypothesized by Bickerton (1990) was used for. The detection and collective processing of problematic situations can be understood as an additional ability which gave rise to modern language.

The importance of studying language use

Language is marginally used to give orders or make promises. Its main use is conversation. According to Dunbar (1996), people in various cultures spend 20% of their awake time in conversation. If we want to understand the function of language, we cannot ignore the way it is actually used in spontaneous conversations. Some of the reasons why language evolved in the first place are to be found there, in the chatter resulting from everyday social interactions. Several authors have addressed the issue of the emergence of language from the point of view of phonology, syntax or semantics (Hurford *et al.* 1998, parts II & III). Others, like Dunbar, have considered the social implications of language. Surprisingly, few authors have considered that much could be learned about language origin by studying its use.

The study of spontaneous conversation shows that interlocutors behave in a few systematic ways. One of them is to draw attention to salient events, as in the following example:

[translated from French]
A1 – Who?
B1 – People suffering from Alzheimer's disease. They walk aimlessly like that.
A2 – Alzheimer patients are not the only ones to do.
B2 – No, but Alzheimer patients walk miles and miles each day... It's incredible! They're exhausted.
A3 – Really?
B3 – Yes I swear! Some of them… They put a… they tested and everything. One of them, she walked at least a hundred kilometres each day. I promise, she was exhausted!
C1 – Yes, it's amazing!

The behavior of Alzheimer patients is unusual enough to be signalled by A. This type of verbal interaction contrasts with another one, which is characterised by an initial puzzle and the subsequent search for explanation:

[translated from French]
D1 – It's strange, the ticket is more expensive when you buy it in the suburbs
E1 – No, but when you purchase it in Paris, it's a journey extension
D2 – No, I think it's because it is a temporary booth. They charge 4F more because they are especially there.

Here D expresses his surprise because he paid two different prices for the same travel. Notice that D already holds a tentative explanation (D2) and that his first utterance D1 does not really aim at solving his puzzlement, but rather at expressing it. As we will see now, there are fundamental reasons to contrast the conversational behaviours shown in the two preceding excerpts.

Pointing to salient events

The human behavior which consists in pointing to salient events can be quite basic. It can be performed with simple deictic gestures meaning "Look at this!". As soon as an improbable event occurs in the vicinity (say, an elephant walking in the street in Paris), we draw our friends' attention to it. This behavior is almost a reflex, as it is difficult to refrain from performing it when the salience of the event is high. It seems to be deeply rooted in our nature. By the age of nine months, the child begins to point to salient events like a dancing doll used for experimental purposes, or to his father arriving (Carpenter, Nagell & Tomasello 1998). What is new at that age is that the child does not desire to get the object itself, but wants to draw the mother's attention to that object. Tomasello notices that apes are not showing such behavior. Even if their attention is drawn by salient events, they never try to *share* this attention with others.

What is salience? As the study of conversation reveals, it can have two sources (Dessalles 1992). Events considered being *a priori* improbable are perceived as salient, and their occurrence is communicated. Events considered being *a priori* desirable or undesirable, *i.e.* offering a stake, are communicated as well. Technically, salience can be defined as a combination of probability and desirability. The salience of an event of probability p and desirability d (d between –1, the most undesirable, and 1, the most desirable) is defined as:

$$S = Log\left(\frac{1}{p(1-|d|)}\right)$$

This is a straightforward generalisation of Shannon's definition of information. An event of good probability (p above, say, 0.1) and which does not matter too much ($|d|$ small) has a poor salience. This formula provides us with a way to predict which events are likely to be reported in this conversation mode (Dessalles 1992).

How do interlocutors react to the mention of a salient event? C1 and A2 illustrate the two possibilities. C1 acknowledges the salience. After this kind of reaction, the conversation may stop or change to another topic. The effect of A2 is conversely to lower the salience of the reported event. A common way to do this is to mention another event similar to the reported one. A could have done this by mentioning another illness bringing patients to walk. The effect on salience is well predicted by conditional probabilities: if the events mentioned in the initial utterance and in the reply can be considered as two occurrences of the same generic event, then the reply increases the probability of this generic event and thus diminishes its salience. This second strategy may give rise to a recursive process: since the new event has to be as close as possible to the first one, it appears as salient itself. So-called story rounds (Tannen 1984) may ensue, in which the role of each story is to diminish the salience of the preceding one. This is one explanation of why conversations involve more than one or two moves.

The maxim of this conversational mode is simple to state: (1) in the first utterance, try to report a salient event; (2) when replying evaluate the salience of what is said, either (2a) directly or (2b) by reporting a salient event similar to the first one. We expect such conversations to alternate between salient events and evaluations or, because of the similarity between (1) and (2b), to turn into story rounds. The conversational moves (1) and (2) do not require all our sophisticated linguistic abilities in order to be performed in a functional way. In particular, this mode of verbal interaction should have been within the reach of proto-human individuals who relied on protolanguage to communicate.

Detecting problematic issues

The second conversational mode, by contrast, requires from interlocutors the ability to detect logical incompatibilities and to find ways to solve them. In D1, D presents as inconsistent the fact that the fare is not identical in one way and the other. E1 and D2 then appear as attempts to solve these inconsistencies. Inconsistencies are not always epistemic, as in this excerpt. They can oppose incompatible desires. In the following example, F wants his doors to look nice. A solution is to remove the old paint.

[translated from French]
F1 – I have to repaint my doors. I've burned off the old paint. [...] It's really tough work!
[...]
G1 – You have to use a wire brush
F2 – Yes, but that wrecks the wood

Removing the old paint creates a new problem: the tough work. G’s solution, the wire brush, creates a new problem: the wrecking of the wood. The alternation between problems and solutions is typical of this conversational mode. The initial problem may be to *explain* or to *avoid* some state of affairs, the mechanism is the same. We modelled this alternation between problems and tentative solutions as a surface structure generated by a simple recursive procedure (Dessalles 1998). Notice that we have here a second and independent account of the fact that verbal interactions are not limited to one or two moves, but develop into potentially infinite sequences that we call conversations.

The conversational mode illustrated in this section may be called the *argumentative* mode. It requires from participants two main abilities: (1) to detect that events are inconsistent (*e.g.* an effect without cause, a state in contradiction with your desires) and (2) to generate solutions through abduction. Both abilities can be argued to be beyond the reach of apes and, presumably, of proto-humans. This suggests that the argumentative mode emerged when our ancestors were already communicating by reporting and evaluating salient events. There are several further arguments suggesting that there were two stages in the evolution of human communication and that the argumentative is the more recent.

Discussion: Why two conversational modes ?

It is never a good thing to postulate two systems when one is sufficient. Why should we distinguish two conversational modes, one about salient events and the other consisting of an argumentation about problematic issues? The main reason is that these two modes are to be observed in the human conversational behavior, in which they can be isolated as functional parts. Moreover, there are reasons to think that one mode was used before the other in the course of hominid evolution.

According to the protolanguage hypothesis, language in its full-fledged form, *i.e.* with constituent syntax and flexional morphology, emerged with *homo sapiens* (Bickerton 1990; Jackendoff, *to appear*). The main role of syntax is to encode the thematic roles, either by the position or through marking, in the phrases which constitutes a given sentence. Thematic roles can be shown to be necessary for the conception and the expression of inconsistencies. For example, the paradox signalled in D1 may result from a theme, the price of the ticket, which moves (changes value) with no cause (causes, in a thematic representation, may take the form of abstract forces (Talmy 1988)). By contrast, a fully functional communication system based on salience can rely on sole protolanguage, without proper syntax and without thematic roles (Dessalles, *to appear*). This suggests that argumentative abilities emerged together with syntactic abilities with our species, while the previous species relied on protolanguage to communicate about salient events only.

Another reason to see argumentation as a new feature arising in an already established conversational ability is the functional role that can be assigned to the possibility of detecting inconsistencies. As argued in (Dessalles 2000), the ability to detect inconsistencies may have been selected as a protection against lies. In the salient mode, the evaluation procedure, which consists in making comparisons (*e.g.* "*Alzheimer patients are not the only ones to do*"), is a protection against overestimated salience. In the scenario proposed here, modern humans evolved the possibility to check the reliability of reports by means of thematic analysis and detection of inconsistencies. The argumentative ability would be the evolutionary outcome of this protection device.

References

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

Carpenter, M., Nagell, K. & Tomasello, M. (1998). "Social cognition, joint attention, and communicative competence from 9 to 15 months of age". *Monographs of the Society for Research in Child Development*, *255*.

Dessalles, J-L. (2000). *Aux origines du langage – Une histoire naturelle de la parole*. Paris: Hermès

Dessalles, J-L. (1992). *Logical Constraints on Spontaneous Conversation*. Paris : TELECOM-Paris technical report ENST 92-D-011.

Dessalles, J-L. (1998). "The interplay of desire and necessity in dialogue". In J. Hulstijn & A. Nijholt (ed), *Formal semantics and pragmatics of dialogue*. Enschede : University of Twente, TWLT-13, 89-97.

Dessalles, J-L. (2000). "Language and hominid politics". In C. Knight, J. R. Hurford & M. Studdert-Kennedy (ed), *The evolutionary emergence of language: social function and the origin of linguistic form (to appear)*. Cambridge : Cambridge University Press.

Dunbar, R. I. M. (1996). *Grooming, gossip, and the evolution of language*. Cambridge : Harvard University Press.

Hurford, J. R., Studdert-Kennedy, M. & Knight, C. (1998). *Approaches to the Evolution of Language - Social and Cognitive Bases*. Cambridge University Press.

Jackendoff, R. (*to appear*). "Possible stages in the evolution of the language capacity". *Trends in cognitive science.*

Talmy, L. (1988). "Force dynamics in language and thought". *Cognitive Science*, *12*, 49-100.

Tannen, D. (1984). *Conversational Style - Analyzing Talk Among Friends*. Norwood : Ablex Publishing Corporation.

Language from scratch?
Prerequisites shared with monkeys and apes

Frans B. M. de Waal

Living Links, Yerkes Regional Primate Research Center,
and Psychology Department, Emory University, Atlanta, USA.

ABSTRACT

Human language is a unique system of communication characterized by arbitrary symbols with context-defined meanings that are freely combined in hierarchical structures, called syntax. I will focus on the last two characteristics of language with regards to the behavior of nonhuman primates, and argue that they are no strangers to the free combination of elements, and hierarchically structured solutions to problems, both social and technical. We know since Wolfgang Köhler's work of the 1920s that the free combination of old knowledge into new, adaptive solutions is a hallmark of pongid intelligence. We also can view ape tool-use, political strategies, even plant processing as organized in clusters and subclusters of if-then routines. I will further review evidence for reciprocal altruism and exchange of favors among chimpanzees, which is an interindividual if-then routine. My arguments will parallel those that I have made before with regards to the evolution of morality: the full-blown mechanism appears absent in apes, but many of the building blocks can be found. In the same way, humans didn't evolve language from scratch, but exploited many classificatory, combinatorial, and organizational mental abilities that we share with our closest relatives.

A non-modular suggestion about the origin of symbols

Thomas E. DICKINS

Department of Psychology, London Guildhall University,
London E1 7NT, UK
E-mail: dickins@lgu.ac.uk

Preamble

Over the last decade a brand of nativism has taken hold of behavioural science in the form of Evolutionary Psychology. Advocates of this theoretical approach argue that peripheral and central cognition is organised as a modular architecture, each module reflecting a solution to an ancestral problem from the Pleistocene (Cosmides and Tooby, 1992; Pinker, 1997). This Evolutionary Nativism is at odds with recent ontogenetic work that argues for a canalized development whereby constraints guide and shape a more general associative learning system to achieve specific ends (Elman, Bates, Johnson, Karmiloff-Smith, Parisi and Plunkett, 1996; Karmiloff-Smith, 1992). As such one can talk about emergent modularization where ontogenetic development leads to an adult end-state that is module like. This recent perspective has the potential to alter current modes of adaptationist thinking within psychology. In order to explore this potential a theoretical model of symbol origins is proposed that relies on specific constraints canalizing expectancy and operant learning. It is proposed that this system of constraints underwent Baldwinian selection thereby instantiating, in modern *Homo sapiens sapiens*, the ability to learn symbols rapidly. This accords with specific ontogenetic theories of word acquisition (Bloom and Markson, 1998, Markson and Bloom, 1997).

Definition

In this paper I assume that symbols are arbitrarily and symmetrically related to their referent, and in this way symbols are afforded the property of displaced reference. The relation is arbitrary because the symbol has no direct causal relationship with the referent. For example, there is nothing “rabbit-like” about the word <rabbit>. <Rabbit> is attached to rabbits by social convention (Deacon, 1997). Furthermore, if someone points to an array of objects and says <rabbit> we can pick out the rabbit from this array. Equally, we can say <rabbit> when someone presents us with a rabbit. In this way the symbol <rabbit> is symmetrically related to its referent.

Communication is to be regarded as the reliable transfer of information from actor to reactor, resulting in a behavioural change. Symbols are to be seen as one method of information transfer. There are at least four methods of such transfer, shown below in Table 1:

Information Type:	**Feature:**	**Example:**
Cues	Always on	Yellow & black stripes of wasp
Signs/Indexicals	Indicate presence of something	Footprints
Signals	Can be on or off	Alarm calls
Symbols	Displaced reference	Words

Table 1: *Summary of Information Categories*

It is worth noting that signals and symbols are to some extent similar. Signals, such as the vervet monkeys' repertoire of alarm calls, indicate specific objects, events or states of affairs. In the vervet case signals indicate specific predators and subsequently stimulate appropriate behavioural response in other vervets (Cheney and Seyfarth, 1985, 1988). Symbols also indicate specific objects, event and states of affairs and trigger behavioural responses. None the less, the relationship between signal and referent is not symmetrical and is based on affective response so nor is it entirely arbitrary.

It is the properties of arbitrary and symmetrical relationship between symbol and referent that have to be explained by any symbol origins story.

Argument

I argue that in order to construct a model of the origins of symbols it is instructive to think about the systems and abilities which might have preceded them, and I suggest the following:

1. Classical and operant learning;
2. Some order of signaling system;
3. A bias to categorize at the whole object level.

There is good reason to assume, both *a priori* and from comparative data, that these abilities have a more ancient provenance than symbols and symbolic behaviour.

Once we have determined what systems might have been in place in ancestral organisms we need to determine the minimum alterations to environment and/or cognition that would allow for the emergence of symbols. Given that most mutations are not good news for an organism it is wise to avoid catastrophic models of symbol origin of the same order that Bickerton (1990, 1996) has proposed for the emergence of syntax (cf. Studdert-Kennedy, 1992). Bickerton argued for a massive mutation that significantly and fortuitously reorganized cortical architecture. Such an argument to explain the origins of symbolic behaviour would also rely on great fortune rather than minor changes and variance being selected for more cautiously (albeit blindly by Natural Selection).

We could possibly use arguments based on the conception of pre-adaptation (or exaptation as Gould and Vrba, 1982 referred to it).

In order to think about putative models we also need to think about the defining characteristics of symbols as they are used in human language. I argue that the following things are crucial:

1. *The vocal medium* – most languages are spoken (Locke, 1998) which strongly suggests this was the medium in which ancestral linguistic abilities emerged. Written and sign languages are historical inventions that have arisen after spoken languages so, given this, we must produce a possible explanation of how symbols arose within a vocal medium in the first instance.
2. *Social cognition* – modern *Homo sapiens* is a social species, just like many other primate species. From comparative evidence we can hypothesize that our hominid ancestors were a social and hierarchically organized species. Any communication will have happened within this social order. Given the manipulative nature of communication we would expect such communication to reflect social complexity.
3. *Ostensive behaviour* – there is comparative data to suggest that non-human primates engage in ostensive behaviour – either directly pointing with their hands or utilizing the direction of gaze (Gomez, 1998a,b; Leavens and Hopkins, 1998). Also modern humans use gesture to augment spoken language. This proto-referential[1] behaviour is a communication in its own right.

Dunbar's (1993, 1996) hypothetical scenario for the emergence of a vocal language as a form of allogrooming is used as a model for the selection of increased vocalization, rather than language per se. I argue (i) there is a need to clearly indicate to whom specific vocalizations are directed; (ii) it is likely that as vocalizations became predominant they were initially coupled with specific aspects of the physical grooming interaction; and (iii) this coupling was achieved through expectancy learning. Thus the hypothesis can be stated as follows:

Ostensive systems, such as pointing and gaze[2], might have acted to direct the attention of hominid ancestors to the object of vocalizations. First this might have occurred solely within the environment of vocal grooming, as postulated by Dunbar, but once a flexible vocal system was capable of being directed to specific grooming interactions object discriminations would easily follow. The main impact of social grooming would have been vocal control, the main impact of ostension would be that of guiding other behaviours and rendering them communicative. Once firm associations between vocalization and object, event or state of affairs had been established through ostension the specific ostensive act would be unnecessary. This would liberate the vocalization, which would, to all intents and purposes, be arbitrarily linked to its referent. Other symbol properties of displaced reference and symmetry would necessarily follow.

This hypothesis leaves us with a number of options for the implementation of symmetry. Symmetry might come for free. Equally, it might be the result of a two-way learning procedure (i.e. learning <rabbit> to rabbit and vice versa; Place, 1995/6). Alternatively, a variant of Hurford's Saussurean Sign argument could be made (Hurford, 1989; cf. Oliphant, 1996, 1997).

Conclusion

This hypothesis suggests some specific constraints acting to guide learning toward symbol acquisition. The process so far described represents an historical progression and therefore leaves one remaining question – how did symbol acquisition become systematized in *modern Homo sapiens*?

---

[1] By which I mean non-symbolic reference – i.e. not symmetrical or arbitrary.

[2] It is worth noting at this point that humans have very distinctive, species specific, whites of their eyes. It is very likely that this coloration is useful in indicating the direction of gaze. The distinctive nature of the whites further suggests that they have been selected for – perhaps precisely for the role that Gomez is indicating. I am indebted to Professor Richard Bentall for reminding me of this point.

I argue that the ability to use these directing constraints was selected for in a Baldwinian fashion thereby instantiating in the modern human infant the ability to rapidly acquire symbols. Such an argument fits well with the work of Bloom and Markson (1998; Markson and Bloom, 1997) who claim there is no dedicated word learning mechanism in the human infant but instead a canalized predisposition to learn words. Canalization implies the directing of more general (but not necessarily totally general) systems toward specific goals thereby obviating the need for domain specific computations.

References

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

Bickerton, D. (1996). *Language & Human Behaviour.* London: UCL Press.

Bloom, P. & Markson, L. (1998) Capacities underlying word learning. *Trends in Cognitive Science*, *2* (2), 67 - 73.

Cheney, D. L. & R. M. Seyfarth (1985). Vervet monkey alarm calls: manipulation through shared information? *Behaviour 94*: 150 - 166.

Cheney & Seyfarth (1988) Assessment of meaning & the detection of unreliable signals by vervet monkeys. *Animal Behaviour*, *36*, 477 - 486

Cosmides, L. & Tooby, J. (1992). Cognitive Adaptations for Social Exchange. In: *The Adapted Mind: Evolutionary Psychology & the Generation of Culture.* Barkow, J. H., Cosmides L. & Tooby J. New York: Oxford University Press.

Deacon, T. (1997) *The Symbolic Species: The co-evolution of language & the human brain.* London: Penguin: Allen Lane.

Dunbar, R. I. M. (1993). "Coevolution of neocortical size, group size & language in humans." *Behavioural & Brain Sciences, 16*, 681 - 735.

Dunbar, R.I.M. (1996) *Grooming, Gossip & the Evolution of Language*. London: Faber & Faber.

Elman, J.L., Bates, E.A., Johnson, M.H., Karmiloff-Smith, A., Parisi, D. & Plunkett, K. (1997) *Rethinking Innateness: A Connectionist Perspective on Development*. London: MIT Press.

Gomez, J-C (1998a) "Ostensive behaviour in great apes: The role of eye contact." In: Russon, A.E., Bard, K.A. & Taylor Parker, S. (Eds.) *Reaching into Thought: The Minds of the Great Apes*. Cambridge: Cambridge University Press.

Gomez, J-C (1998b) Some thoughts about the evolution of LADS: With special reference to TOM & SAM. In Carruthers, P. & Boucher, J. (Eds.) *Language & Thought: Interdisciplinary themes*. Cambridge: Cambridge University Press.

Gould, S. J. & Vrba, E. S. (1982). "Exaptation - a missing term in the science of form." *Paleobiology 8* (1): 4 - 15.

Hurford, J. (1989) Biological evolution of the Saussurean sign as a component of the language acquisition device. *Lingua*, *77*, 187 – 222.

Karmiloff-Smith, A. (1992) *Beyond Modularity: A Developmental Perspective on Cognitive Science*. London: MIT Press.

Leavens, D.A. & Hopkins, W.D. (1998) Intentional Communication by Chimpanzees: A Cross-Sectional Study of the Use of Referential Gestures. *Developmental Psychology*, *34* (5), 813 – 822.

Locke, J.L. (1998) Social sound-making as a precursor to spoken language. In: Hurford, J.R., Studdert-Kennedy, M. & Knight, C. (Eds.) *Approaches to the Evolution of Language*. Cambridge: Cambridge University Press.

Markson, L. & Bloom, P. (1997) Evidence Against a Dedicated System for Word Learning in Children" *Nature*, *385* (6619), 813 – 815.

Oliphant, M. (1996) The dilemma of Saussurean communication. *Biosystems*, *37* (1-2), 31 – 38.

Oliphant, M. (1997) Formal approaches to innate and learned communication: Laying the foundation for language. Ph.D. Dissertation.

Pinker, S. (1997) *How the Mind Works*. Allen Lane: The Penguin Press.

Place, U. T. (1995/6). "Symbolic Processes & Stimulus Equivalence." *Behaviour & Philosophy 23/4*: 13 - 30.

Studdert-Kennedy, M. (1992) Leap of faith: A review of Language & Species. *Applied Psycholinguistics*, *13*, 515 – 527.

Why linguists and anthropologists should be interested in the aquatic ape theory for the origin of speech.

Karl DILLER

Professor of Linguistics
University of New Hampshire
Durham NH 03824 USA
karl.diller@unh.edu

Outline

- – Internal Reconstruction, the Comparative Method, and Hard Evidence.
- – The Aquatic Ape syndrome and speech.
- – The Preadaptation for Speech
- – Implications for the origin of language.

Internal reconstruction, the comparative method, and hard evidence.

If "human language is an embarrassment for evolutionary theory", as stated in the epigraph at the head of the Web Page for this conference, it is because linguists and anthropologists have not taken seriously the method of internal reconstruction and the evidence it gives for the Aquatic Ape Theory.

Linguists are brought up on the story of how the young Ferdinand de Saussure, in his doctoral dissertation, reconstructed a series of pharyngeal consonants for the proto-Indo-European language, even though these consonants were not present in Sanskrit, Greek, Latin, or any known Indo-European language. The pharyngeal consonants were reconstructed not by comparative evidence but by the method of internal reconstruction. Various alternations in the vowel systems of Indo-European languages could be explained if there had been an original series of pharyngeal consonants.

How reliable is internal reconstruction? Don't we need harder evidence? Classicists and philologists wedded to the text were highly skeptical. Then a few years later the Hittite language was deciphered – it was an Indo-European language with a series of consonants right where de Saussure had posited them.

Morphophonemic rules which are formulated by internal reconstruction to take account of sound alternations in present day speech end up recapitulating the history of sound changes, as seen in Bloomfield's classic paper on Menomini morphophonemics and Chomsky and Halle's synchronic reconstruction of the English vowel shift in their *Sound Patterns of English*. Bloomfield's rules mirrored those derived from comparative Algonquin data; Chomsky and Halle's rules were like those derived from historic texts.

The distinction between the method of internal reconstruction and the comparative method is also seen in molecular biology and genetics. Comparing mitochondrial DNA in various populations to find the source of the human population in an African Eve is a use of the comparative method. The recent reconstruction by Lahn and Page (1999) of the history of how the human Y chromosome came to diverge from its originally homologous X chromosome is a very interesting example of internal reconstruction, involving four separate inversions of DNA on the Y chromosome, ordered in a clear chronological order. The time scale was then made more precise by comparative genetic data and by using the fossil dating of species divergence. The first event, which marked the beginning of X-Y divergence occurred about 240 to 320 million years ago, shortly after the separation of mammals and birds from their common ancestor – pretty far back in the pre-history of humans! The most recent inversion occurred between 30 and 50 million years ago, during the period of primate evolution but before the split between Old and New World monkeys. This is a model study for the relative importance of internal reconstruction and hard evidence. Fossils are incapable of yielding any evidence about the evolution of the Y chromosome. Internal reconstruction defined the problem and provided the explanation. Then comparative evidence and hard fossil evidence provided refinements and certain kinds of precision.

The Aquatic Ape syndrome and speech.

In an inspired example of internal reconstruction, Oxford marine biologist Alister Hardy (1960) argued that several unusual aspects of human anatomy and physiology can be explained if a human ancestor went through an aquatic or semi-aquatic evolutionary phase.

The biological preadaptation of speech – the voluntary control of one's breathing and vocalization – is a feature of aquatic mammals, as Elaine Morgan has pointed out in three books since 1972, and it fits well with the cluster of other human aquatic features initially pointed out by Hardy. Hardy's list includes reduced body hair, the orientation of the hair tracts, the layer of subcutaneous fat, sweat glands, upright posture, flexibility of the spine, swimming and diving capabilities (including the slowed metabolism of the diving reflex), and the form of the hands and feet. Morgan adds the features of salt tears, the ability to frown, the shape of the nose, the form of the buttocks and female breasts, face to face copulation, and the foundations of speech.

These are features we share with aquatic mammals and which we do not share with chimpanzees. In particular, the ability to pronounce words of human language is almost completely beyond the capability of the chimpanzee, as is obvious to anyone who has seen films of the Hayes's chimpanzee Vicky trying with great effort to produce the repertoire of four words she learned to mimic after six years of behaviorist training. The words were *mama, papa, cup,* and *up*, pronounced without voicing, with great contortion of the face and upper body, requiring even the help of the hand to keep the lips closed for the whispered /m/ in *mama*. A chimpanzee would not be a good candidate to initiate the evolution of speech.

Hardy had just returned in 1927 from a two year excursion to study the biology of the Antarctic Seas, when he found this statement in Wood-Jones's *Man's Place among the Mammals*: "The peculiar relation of the skin to the underlying superficial fascia is a very real distinction, familiar enough to everyone who has repeatedly skinned both human subjects and any other members of the Primates. The bed of subcutaneous fat adherent to the skin, so conspicuous in Man, is possibly related to his apparent hair reduction; though it is difficult to see why, if no other factor is invoked, there should be such a basal difference between Man and the Chimpanzee." Hardy immediately recognized the "other factor" that was involved: an aquatic environment. Nearly all aquatic warm-blooded vertebrates have a layer of subcutaneous fat, from ducks and penguins to dolphins and whales. And mammals who have permanently returned to the aquatic environment from land have generally lost their hair, as whales and dolphins have.

Nearly every major group of vertebrates from the time of the dinosaurs has had a member species return to an aquatic or semi-aquatic environment. Paleontologists have no trouble identifying the skeletons of aquatic dinosaurs, the ichthyosaurs and plesiosaurs. Since humans never became as completely aquatic as the whales, we do not have a completely aquatic-looking skeleton. But note the elongated streamlined body, which is well adapted to swimming and which requires upright posture on land.

The theory is that the human ancestor was forced into an aquatic environment – perhaps marooned on an island – and because of this geographical isolation, diverged into a new species of primate. Presumably this animal always slept on land, but spent a major part of its waking hours in the water, much of the time wading upright with its head out of the water, searching for shellfish and other marine creatures for a diet that came to be increasingly derived from animal sources. This diet, rich in Omega-6 and Omega-3 fatty acids, would have permitted the increase in brain growth that was initiated at this time.

The proposals are actually quite specific. Leon P. LaLumiere first located the island to what is now the Danakil Alps in the Afar region of northeastern Ethiopia near the fossil bed of Lucy, the *australopithecus afarensis*. The Danakil Alps are surrounded by salt plains, and were isolated by the Afar Sea at the time of the divergence of the human line from chimpanzees. Morgan quotes the geologist Paul Mohr, that "What is more or less agreed upon is that by the late Miocene [seven million years ago] a marine basin had become established over northern Afar...and these conditions persisted until the isolation and desiccation of the salt plain arm of the sea some 70,000 years ago (CNR-CNRS team, 1973)" (quoted in Morgan 1990, 51). This area was a forested primate habitat at the time of the inundation, and would have been an obvious setting for the separation of the human line from the other apes, and for the shaping of the human line in a forced aquatic setting.

The implication is that whether or not the Australopithecines were direct descendants of humans, or whether they escaped from the aquatic environment at a divergent time, they were shaped in the same or in a similar aquatic setting. Along with their bipedal gait, the Australopithecines presumably had most of the other aquatic features of human anatomy including hairlessness and the layer of subcutaneous fat and the other anatomical features that separate us so markedly from Chimpanzees. But of course if the Australopithecines were not in the direct ancestral line of humans, they may not have developed all these aquatic features as much as the human line did.

The preadaptation for speech

Mammals forced from land into aquatic environments develop voluntary control of breathing and specific adaptations to prevent water from entering the lungs through the nose. Elephants, excellent swimmers and another species with an aquatic interlude before returning to land, developed an extremely long nose which they can hold up out of the water. The human nose is not so spectacular, but it is impressively large compared with Chimpanzees, and is streamlined appropriately for a swimmer. The human nose is similar to that of the Proboscis monkey, a semi-aquatic primate living in the mangrove swamps of Southeast Asian islands. Seals developed valvular nostrils; humans developed the velum which closes off the back of the nose. Human diving capabilities are formidable, as seen for example by traditional pearl divers in Japan and Korea, and Hardy points out that humans have a partial diving reflex in which metabolism is slowed. A swimming and diving creature needs to be able to take in quickly a large amount of air, and for humans this is facilitated by the descent of the larynx which enables and enhances the taking in of large breaths through the mouth. (See also the contribution by Verhaegen and Munro in this conference proceedings on the aquatic preadaptations for speech).

Implications for the origin of language.

The disreputable theories of language origin, the bow-wow theory, the ding-dong theory and the like, were hung up on the problem of how voluntarily controlled speech could have emerged from instinctive stimulus-bound cries. This problem disappears if the earliest bipedal ancestors of humans were preadapted for voluntary control of vocalization. If these human ancestors also had the ability of present-day chimpanzees to use manual and visual symbols and to understand vocal symbols, we are at a very good starting point for a system of vocal language without a lot of hand-waving.

With preadaptation, we have a very long timescale to work with. Swadesh (1971) argued that there were eoglottic, paleoglottic, and neoglottic eras in the prehistory of language, corresponding to the eolithic, paleolithic, and neolithic eras of toolmaking. The eoglottic/eolithic period, he argued, went back to the australopithecines, and was a very long period in which the elements of language were as limited and primitive as the hand tools. Preadaptation makes it plausible that pre-humans actually used that entire period of time in developing language. There need not be any direct connection between toolmaking and language to suggest that language development might have taken as long in pre-human culture as toolmaking ability did. If so, we presumably have a situation where the evolution of language was longer and more graded than in the model of two relatively quick and almost discontinuous stages of protolanguage and full language. We may have had eo-protolanguage, paleo-protolanguage, and neo-protolanguage. Given the long timescale and what we know about grammaticalization, we also may have had an appreciable amount of proto-grammar in those earlier eras. After all, a preadaptation for voluntary use of the vocal tract implies the existence of a basic functional Broca's area in the motor association cortex, and just as agrammatism is associated with Broca's Aphasia, basic grammar ability may well be associated with this elementary but functional Broca's area.

References

Hardy, Alister (1960). Was man more aquatic in the past? *New Scientist* 642-654.

Lahn, Bruce T., and David C. Page (1999). Four evolutionary strata on the human X chromosome. *Science* 286:5441, 964-967.

Morgan, Elaine (1972). *The Descent of Woman.* New York: Stein and Day.

_______(1982). *The Aquatic Ape*. Souvenir Press. (1990).

_______ (1990). *The Scars of Evolution*. Oxford: Oxford University Press.

Swadesh, Morris (1971). *The Origin and Diversification of Language*. Chicago: Aldine-Atherton.

Verhaegen, Marc and Stephen Munro (2000). The Origin of Phonetic Abilities: A Study of Comparative Data with Reference to the Aquatic Theory. Conference on the Evolution of Language, April 4-6, 2000, Paris.

Evolution of dissociable processing of function *vs.* content words for syntactic comprehension

Peter Ford DOMINEY

Institut des Sciences Cognitives, CNRS UPR 9075
67, Boulevard Pinel, 69675 BRON Cedex, France

Introduction

A central function in syntactic analysis or syntactic comprehension is the assignment of thematic roles to noun phrases in sentences. In a simplified manner, we can consider that in languages like French and English, there is a default or *canonical* order in which thematic roles are assigned (e.g. “Agent Object Recipient” in English for the canonical sentence “John gave the ball to Mary.”) However, in non-canonical sentences (e.g. “The ball was given to Mary by John.”), this ordering is transformed, and thematic role assignment is guided, in part, by function items (e.g. closed class words, including prepositions “to” and “by”, or grammatical morphemes).

The objective of this paper is to describe the possible evolutionary development of a capacity for thematic role assignment in syntactic analysis, based on dissociable processing of function and content words. The method will be to first identify a quantifiable behavioral measure of thematic role assignment, and to then propose a plausible model that yields this identified target behavior. The model will be constrained by the requirement that to the largest extent possible its functional components should correspond to behavioral or computational capabilities whose underlying neurophysiology is known. The system’s development will thus represent a minimal cost, largely attributable to the recombination of existing computational capabilities. The model will then be tested against the established behavioral criteria, and will be further validated to predict specific analysis impairments in agrammatic adults, and also the presence of primitive analysis capabilities that should be present in infants. The ensemble of behavioral and simulation results will be discussed in the context of economy in evolution of syntactic analysis capabilities.

The ability to assign thematic roles has been quantified in different clinical tests used to asses agrammatism in aphasia. A well know version developed by Caplan et al (*Cognition*, 1985) consists of 9 sentence types of varying syntactic complexity, five canonical and four non-canonical, and will serve as our target problem. Five sentences of each type are used for a total of 45 sentences. The sentences are constructed so that no semantic interpretation can contribute to thematic role assignment, which must proceed entirely as guided by function items. Sentences are real aloud to subjects in a pseudo-random order, and after each sentence, the subject should indicate by pointing at photographs “who did what to whom”. Interestingly, a rather significant subgroup of aphasics with left-hemisphere lesions that include the peri-sylvian cortex in and around Broca’s area demonstrate a thematic role assignment deficit that is highly selective for non-canonical sentences. The specific vulnerability of this comprehension factor indicates that it is a specialized capability, not directly related to other aspects of syntactic analysis.

In order to realize this thematic role assignment task, a system should first be capable of distinguishing function words (or morphemes) from content words. Numerous behavioral and event-related brain potential studies indicate that indeed, adults process function words and content words in a dissociated manner. The system should also be able to store the content words in a form of working memory, and then to access the contents of this memory in a non-standard order (i.e. different from the input order) guided by the function items. This capability to transform the order of content words, guided by the function words, provides the basis for the assignment of thematic roles.

The model

We have recently developed such a model for the processing of serial surface structure, and abstract transformational rules for non-linguistic sensorimotor sequences (Dominey et al. 1998). The model was developed to simulate human behavior in artificial grammar learning tasks. In these tasks, two isomorphic sequences ABCBAC and DEFEDF have different surface structure but share the same abstract grammatical structure 123213 in which the second triplet is obtained via a systematic transformation of the first. We can thus consider the abstract structure 123213 to be non-canonical, whereas 123123 would be canonical. We demonstrated that while one relatively simple "surface" system (a recurrent network (Dominey et al. 1995)) can learn and reproduce individual sequences, the "abstract" system requires the addition of the working memory processes described above in order to acquire and transfer knowledge of the abstract structure to isomorphic sequences. This resulted in the development of a dual-system model with separate sub-systems for surface and abstract structure respectively. We argue that the functional components – recurrent connections and working memory – exist in the non-human primate and are thus evolutionarily plausible building blocks, and that the important evolutionary step may involve a novel recombination of these capacities.

We have recently demonstrated (Dominey & Ramus 2000) that in its untrained initial state, the "surface" system of the dual stream model is sensitive to the temporal or rhythmic structure of language as demonstrated in newborn infants (Nazzi et al. 1998), as well as to the serial structure and statistical regularities of speechlike sound sequences as demonstrated in 8 month olds (Saffran et al. 1996). Similarly the "abstract" system is capable of learning within 2 minutes to discriminate between sentence with the abstract structures AAB vs ABB vs ABA as demonstrated in 8 month olds (Marcus al. 1999).

We now examine the behavior of the resulting dual-system model in syntactic comprehension of canonical and non-canonical sentences (Dominey 1999). The crucial "evolutionary" modification is that function and content items are treated in dissociated processing streams. Function items are represented by the "surface" system and guide the application of transformational rules that are represented by the "abstract" system. We can demonstrate that after training on a "supervised" version of the Caplan task in which the correct thematic role assignment is provided, the model can then perform the standard task correctly, including the generalization to new sentences. Interestingly, when rendered agrammatic by disruption of the representation of the function items within the recurrent network, the model fails in the processing of transformational structure both for lingustic and non-linguistic sequences. With respect to this prediction, we have recently demonstrated that agrammatic patients fail to process non-canonically ordered sentences both in natural and artificial grammars (Dominey & Lelekov 2000).

Together these observations allow us to propose a neural network model that simulates the combination of simple functions that allow the realization of a reduced form of thematic role assignment in syntactic analysis. The model explains data on normal and aphasic performance, and also predicts performance of infants and aphasics in a novel artificial grammar learning task. While clearly a long way from a complete model of syntactic analysis, this simplified model provides a point of departure for discussions of how syntactic analysis might and might not be implemented. Recurrent cortico-cortical connections provide the basis for context encoding necessary for the representation of syntactic organization of function words. Likewise cortico-subcortical working memory circuits provdie the necessary storage for on-line processing of open class items. Both of these functions have been demonstrated in non-human primates. What they lack is the coordinated cooperation of these two functions so that sequential context can guide the application of transformations on items stored in the working memory. This suggests that while monkeys may posses capabilities for sequence learning and working memory, they are separated from man in part by an evolutionary step providing the innate specification of these systems' interaction required for this aspect of syntactic analysis.

References

Caplan D, Baker C, Dehaut F (1985) Syntactic determinants of sentence comprehension in aphasia. *Cognition*, 21, 117-175,

Dominey PF, Arbib MA, Joseph JP (1995) A Model of Cortico-Striatal Plasticity for Learning Oculomotor Associations and Sequences, *J Cog Neuroscience*, 7:3, 311-336

Dominey PF, Lelekov T, Ventre-Dominey J, Jeannerod M (1998) Dissociable Processes for Learning the surface and abstract structure sensorimotor sequences. *Journal of Cognitive Neuroscience*, 10 :6 734-751.

Dominey PF, Ramus F (2000) Neural network processing of natural language I : Sensitivity to serial, temporal and abstract structure of language in the infant. *Language and Cognitive Processes*. In Press

Dominey PF (1999) Neural network processing of natural language. II: Syntactic Comprehension via Dissociable Processing of Content and Function Words in the Normal and Agrammatic Adult, ISC Working Paper 99-9.

Marcus GF, Vijayan S, Bandi Rao S, Vishton PM (1999) Rule learning by seven-month-old infants. Science Jan 1;283(5398):77-80

Nazzi T, Bertoncini J, Mehler J (1998) Language discrimination by newborns: Towards an understanding of the role of rhythm. *Journal of Exp Psych. Human Percept & Perform*, *24*(3), 1-11.

Saffran JR, Aslin RN, Newport EL (1996b) Statistical learning by 8-month-old infants. *Science*, Vol. 274, 1926-1928.

Correlation between genetic and linguistic differentiation of human populations: The specific action of linguistic boundaries on gene flow ?

Isabelle DUPANLOUP de CEUNINCK[1], André LANGANEY[1,2], Laurent EXCOFFIER[1]

[1] *Laboratoire de Génétique et Biométrie, Département d'Anthropologie,*
Université de Genève, Suisse.
isabelle.dupanloup@anthro.unige.ch
[2] *Laboratoire d'Anthropologie Biologique, Musée de l'Homme,*
Paris, France

Introduction

Several linguists have proposed recently to classify the 5'000 actual world's languages in only slightly more than 20 families and a few isolates (Bomhard and Kerns 1994, Fleming 1987, Ruhlen 1987, Ruhlen 1994, Shevoroshkin 1989, Starostin 1989). Among these linguists, several authors support the hypothesis of a common origin of these linguistic families and isolates. But this proposal, and also the proposal of the existence of intra and inter-continental families of languages, is far from gaining a large support in the field of genetic linguistics.

The human population geneticists have tried recently to compare their findings to those of the linguists. In the last decades, several studies have revealed a strong correlation between the genetic and linguistic (of high level) differentiation of human populations at both world-wide scale (Cavalli-Sforza, et al. 1992, Cavalli-Sforza, et al. 1988) and continental scale (Excoffier, et al. 1991, Excoffier, et al. 1987, Greenberg, et al. 1986). These results can be interpreted as the trace of two alternative or concommitant processes: the coevolution of genetic and linguistic structures and/or the drastic reduction of the exchange of genes between human populations by linguistic boundaries.

The study of human genetic polymorphisms seems to indicate a relatively recent coalescence of the genealogies of all human populations. The actual human populations would have diverged recently from a small ancestral population in which probably a low level of linguistic diversity, in the sense of the number of independent linguistic lineages, would have been observed. Thus, even though the validity of the existence of large families of languages can not be discussed in the genetic linguistics field, the genetic of human populations seems favorable to the hypothesis of a common origin of the actual human languages. Moreover, the correlation observed between the genetic and linguistic differentiation of human populations can be interpreted as the results of common divergence of genes and languages in the history of these populations.

This correlation could nevertheless have been induced by the action of linguistic boundaries, i.e. the transition zones between the repartition areas of languages and families of languages, on the exchange of genes between populations. If these frontiers effectively reduced the exchange of migrants and thus the gene flow between populations, the genetic differentiation of populations speaking different languages would be more important than those between populations speaking the same language because the last ones echanged more genes than the former ones. We will thus observe a strong correlation between linguistic and genetic diversity.

We have developped an original methodology to test if linguistic frontiers correspond to barriers of genetic contacts between populations (Dupanloup de Ceuninck, et al. submitted). We do not reject the hypothesis of a synchronisation of the genetic and linguistic differentiation of human populations. This hypothesis, although really difficult to test, is not incompatible with the proposal that certain frontiers between languages are not permeable to free gene flow.

We propose here to present our new method and the results of its application to the linguistic boundary between afro-asiatic and indo-european populations tested by classical and molecular markers (RH system (Sanchez-Mazas 1990), Y chromosome-specific p49a,f/TaqI restriction polymorphisms (Poloni, et al. 1997)).

Linguistic boundaries: Segmentation and evaluation

The aim of our method is to evaluate the "permeability" of linguistic boundaries that is to say to estimate the reduction of gene flow between populations exerted by these cultural boundaries. We determine if populations speaking different languages are more differentiated genetically than populations belonging to the same linguistic group. The hypothesis tested here is: do linguistic boundaries correspond to genetic boundaries ?

Our method is based principally on the comparison of the genetic distances between populations of the same linguistic group and between populations located on each side of the linguistic boundary under evaluation. We used an isolation by distance model to estimate the value added by the linguistic frontier under study to the genetic distance expected between populations located on each side of this frontier taking into account the geographical distribution of the populations. This genetic distance added by the linguistic boundary, called δa statistic, is tested by permuting the populations on each side of this frontier and re-computing this statistic after each permutation round to get its null distribution.

We also test whether the distribution of gene frequencies representing the groups separated by the frontier under study differ from each other by use of the analysis-of-molecular-variance approach (AMOVA) (Excoffier, et al. 1992). The variance component F_{CT} due to the variation between the groups of populations separated by the linguistic frontier is estimated and its significance is tested by use of a nonparametric permutational procedure (Excoffier, et al. 1992, Schneider, et al. 1997).

As the processes acting on different portions of the frontier may be heterogeneous, we divide the frontier into segments of arbitrary sizes and then evaluate the "permeability" of each segment independently. Some portions of the linguistic boundary may indeed act as strong genetic barriers whether or not they correspond to an ecological frontier; some others may alternatively not enhance genetic differentiation. The goal of the segmentation of the linguistic boundary under evaluation is to understand at a finer scale the genetic processes at work along this boundary. As for the whole frontier, we associate to each segment analyzed, an F_{CT} value to estimate the genetic variation between the groups of populations on each side of the segment and the value of the genetic distance added by the segment to the genetic distance expected between populations located on each side of this segment.

Application of our methodology to the afro-asiatic/indo-european case

We have chosen to study the impact on gene flow of the linguistic frontier separating the afro-asiatic and indo-european populations which are well characterized for classical as well as molecular markers. The afro-asiatic and indo-european language families are associated according to several linguists (Bomhard and Kerns 1994, Illich-Svitych 1971-1984) in the nostratic super-phylum. These two groups are separated in the western part of their repartion area by the Mediterranean sea and join in the eastern zone in the Middle-East. Acccording to a study of the Rhesus and GM polymorphisms in the language families defined by (Ruhlen 1987)

(Dupanloup de Ceuninck 1999), we notice that these two families show close genetic peculiarities, the genetic proximity of the afro-asiatic populations of North Africa and the indo-european populations of Europe, the differentiation of the afro-asiatic populations of East Africa in the sense of populations of other african linguistic groups (khoisan, nilo-saharan and niger-kordofanian language families) and a large diversity of the indo-european populations of western Asia and India.

The whole afro-asiatic and indo-european linguistic boundary seems to have reduced the gene flow between populations (RH system, Y chromosome-specific p49a,f/TaqI restriction polymorphisms) but the impact of this frontier on population differentiation seems quite heterogeneous, especially in the middle section of the frontier (RH system). The Mediterranean Sea does not constitute a real barrier to gene flow, as is documented for historical times, with the development of commercial routes between the northern and southern part of the Mediterranean Sea. This result is also compatible with the independent colonization of the northern and southern coasts from the Middle East (Renfrew 1991, Renfrew 1994) with following subsequent contacts between populations of the two immigration waves.

References

Bomhard, A. R. and Kerns, J. C., 1994. *The Nostratic Macrofamily, a study in distant linguistic relationship*. Berlin, New York: Mouton de Gruyter.

Cavalli-Sforza, L. L., Minch, E. and Mountain, J. L., 1992. Coevolution of genes and languages revisited. *Proceedings of the National Academy of Science* 89, 5620-4.

Cavalli-Sforza, L. L., Piazza, A., Menozzi, P. and Mountain, J., 1988. Reconstruction of human evolution: Bringing together genetic, archaeological and linguistic data. *Proceedings of the National Academy of Science* 85, 6002-6.

Dupanloup de Ceuninck, I., 1999. Evaluation et synthèse des contributions de la linguistique et de la génétique à l'étude de la différenciation des populations humaines pendant la préhistoire récente *Genetics and Biometry Laboratory*. University of Geneva, Switzerland.

Dupanloup de Ceuninck, I., Schneider, S., Langaney, A. and Excoffier, L., submitted. Inferring the impact of linguistic boundaries on population differentiation: application to the Afro-Asiatic – Indo-European case. *European Journal of Human Genetics* ,

Excoffier, L., Harding, R. M., Sokal, R. R., Pellegrini, B. and Sanchez-Mazas, A., 1991. Spatial differentiation of Rhesus and GM haplotype frequencies in Sub-Saharan Africa and its relation to linguistic affinities. *Human Biology* 63(3), 273-307.

Excoffier, L., Pellegrini, P., Sanchez-Mazas, A., Simon, C. and Langaney, A., 1987. Genetics and history of Sub-Saharan Africa. *Y. Phys. Anthrop.* 30, 151-94.

Excoffier, L., Smouse, P. and Quattro, J. M., 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes : application to human mitochondrial DNA restriction data. *Genetics* 131, 479-91.

Fleming, H. C., 1987. Toward a definitive classification of the world's languages. *Diachronica* IV, 159-223.

Greenberg, J. H., Turner, C. G. and Zegura, S. L., 1986. The Settlement of the Americas: a comparison of the linguistic, dental and genetic evidence. *Current Anthropology* 27,no5, 477-97.

Illich-Svitych, V., 1971-1984. *A comparison of the Nostratic languages*. Vol. 1(1971), 2(1976), 3(1984). Moscow.

Poloni, E. S., Semino, O., Passarino, G., Santachiara-Benerecetti, A. S., Dupanloup, I., Langaney, A. and Excoffier, L., 1997. Human genetic affinities for Y-chromosome P49a,f/Taql haplotypes show strong correspondance with linguistics. *American Journal of Human Genetics* 61, 1015-35.

Renfrew, C., 1991. Before Babel: Speculations on the Origins of Linguistic Diversity. *Cambridge Archeological Journal* 1, 3-23.

Renfrew, C., 1994. World Linguistic Diversity. *Scientific American* , 104-10.

Ruhlen, M., 1987. *A Guide to the World's Languages*. Vol. 1: Classification. London: Edward Arnold.

Ruhlen, M., 1994. *The origin of language: tracing the evolution of the mother tongue*. New York: John Wiley & Sons, Inc.

Sanchez-Mazas, A., 1990. Polymorphisme des systèmes immunologiques Rhésus, GM et HLA et histoire du peuplement humain *Genetics and Biometry Laboratory*. University of Geneva, Switzerland

Schneider, S., Kueffer, J. M., Roessli, D. and Excoffier, L., 1997. Arlequin ver. 1.1: A software for population genetic data analysis. University of Geneva, Switzerland: Genetics and Biometry Laboratory

Shevoroshkin, V., 1989. Methods in Interphyletic Comparisons. *Ural-Altaic Yearbook* 61, 1-26.

Starostin, S., 1989. Nostratic and Sino-Caucasian. In: V. Shevoroshkin (ed) *Explorations in Language Macrofamilies, first International Interdisciplinary Symposium on Language and Prehistory (Ann Harbor, 1989)*: Bochum: Brockmeyer, 42-66.

The evolution of subjacency without universal grammar: Evidence from artificial language learning

Michelle R. Ellefson & Morten H. Christiansen

Southern Illinois University

*{ellefson, morten}@siu.edu*mailto:morten%7D@siu.edu

The acquisition and processing of language is governed by a number of universal constraints. Undoubtedly, many of these constraints derive from innate properties of the human brain. Theories of language evolution seek to explain how these constraints evolved in the hominid lineage. Some theories suggest that the evolution of a Chomskyan universal grammar (UG) underlies these universal constraints. More recently, an alternative perspective is gaining ground. This approach advocates a refocus in evolutionary thinking; stressing the adaptation of linguistic structures to the human brain rather than vice versa (e.g., Christiansen, 1994; Kirby, 1998). On this account, many language universals may reflect non-linguistic, cognitive constraints on learning and processing of sequential structure rather than innate UG. If this is correct, it should be possible to uncover the source of some linguistic universal in human performance on sequential learning tasks. This prediction has been borne out in previous work by Christiansen (2000) in terms of an explanation of basic word order universals. In this paper, we take a similar approach to one of the classic linguistic universals: subjacency.

Why subjacency?

According to Pinker and Bloom (1990), subjacency is one of the classic examples of an arbitrary linguistic constraint that makes sense only from a linguistic perspective. Informally, “Subjacency, in effect, keeps rules from relating elements that are ‘too far apart from each other’, where the distance apart is defined in term of the number of designated nodes that there are between them” (Newmeyer, 1991, p. 12). Consider the sentences in Table 1. According to the subjacency principle, sentences 3 and 6 are ungrammatical because too many boundary nodes are placed between the interrogative pronouns and their respective 'gaps'. In the remainder of this paper, we explore an alternative explanation which suggests that subjacency violations are avoided, not because of a biological adaptation incorporating the subjacency principle, but because language *itself* has undergone adaptations to root out such violations in response to non-linguistic constraints on sequential learning.

Artificial language experiment

We created two artificial languages, natural (NAT) and unnatural (UNNAT), consisting of letter strings, derived from a basis of 6 different constructions (see Table 2). Each training set consisted of 30 items. In NAT training, 10 items were grammatical complement structures involving complex extractions in accordance with subjacency (SUB) (5 and 6 in Table 2). For UNNAT training, the 10 SUB items involved subjacency violations (5* and 6*). The 20 remaining training items were general grammatical structures (GEN) that were the same for both groups (1–4 in Table 2). The test set contained 60 novel strings, 30 grammatical and 30 ungrammatical for each group. Twenty-eight novel SUB items, 14 each, grammatical and ungrammatical complex extraction structures were created. For UNNAT, ungrammatical SUB items were scored as grammatical and grammatical SUB items were scored as ungrammatical. The reverse was true for NAT. We created 16 novel grammatical GEN items. Sixteen ungrammatical GEN items were created by changing a single letter in each grammatical item, except for those letters in the first or last position. Both training and test items were controlled for length across conditions and balanced according to different types of frequency information.

1. Sara asked why everyone likes cats. N V Wh N V N	4. Sara heard (the) news that everybody likes cats. N V N Comp N V N
2. Who (did) Sara ask why everyone likes cats? Wh N V Wh N V N	5. What (did) Sara hear that everybody likes? Wh N V Comp N V
3. *What (did) Sara ask why everyone likes? Wh N V Wh N V	6. *What (did) Sara hear (the) news that everybody likes? Wh N V N Comp N V

Table 1. *Examples of Grammatical and Ungrammatical NP- and Wh-Complements*

In total, 60 adults participated in this experiment, 20 in each of three conditions (NAT, UNNAT, and CONTROL). NAT and UNNAT learned the natural and unnatural languages, respectively. CONTROL completed only the test session. During training, individual letter strings were presented briefly on a computer. After each presentation, participants were prompted to enter the letter string using the keyboard. Training consisted of 2 blocks of the 30 items, presented randomly. During the test session, with 2 blocks of the 60 randomly presented items, participants decided if the test items were created by the same (grammatical) or different (ungrammatical) rules as the training items.

NAT		UNNAT	
Sentence	Letter String Example	Sentence	Letter String Example
1. N V N	Z V X	1. N V N	Z V X
2. Wh N V	Q Z M	2. Wh N V	Q Z M
3. N V N comp N V N	Q X M S X V	3. N V N comp N V N	Q X M S X V
4. N V Wh N V N	X M Q X M X	4. N V Wh N V N	X M Q X M X
5. Wh N V comp N V	Q X V S Z M	5*. Wh N V N comp N V	Q X V X S Z M
6. Wh N V Wh N V N	Q Z V Q Z V Z	6*. Wh N V Wh N V	Q Z V Q Z V

Note: Nouns (N) = {Z, X}; Verbs (V) = {V, M}; comp = S; Wh = Q.

Table 2. *The Structure of the Natural and Unnatural Languages (with Examples)*

Results and discussion

Controls. Since the test items were the same for all groups, but scored differently depending on training condition, the control data was scored from the viewpoint of both the natural and unnatural languages. Differences between correct and incorrect classification from both language perspectives were non-significant with all t-values <1 (range of correct classification: 59%–61%). Thus, there was no inherent bias in the test stimuli toward either language.

Experimental group. An overall t-test indicated that NAT (59%) learned the language significantly better than UNNAT (54%) ($t(38)=3.27$, $p<.01$). This result indicates that the UNNAT was more difficult to learn than the NAT. Both groups were able to differentiate the grammatical and ungrammatical items (NAT: $t(38)=4.67$, $p<.001$; UNNAT: $t(38)=2.07$, $p<.05$). NAT correctly classified 70% of the grammatical and 51% of the ungrammatical items. UNNAT correctly classified 61% of the grammatical and 47% of the ungrammatical items. NAT (66%) exceeded UNNAT (59%) at classifying the common GEN items ($t(38)=2.80$, $p<.01$). Although marginal, NAT (52%) was also better than UNNAT (50%) at classifying SUB items ($t(38)=1.86$, $p=.071$). Note that the presence of the SUB items affected the learning of the GEN items. Even though both groups were tested on exactly the same GEN items, the UNNAT performed significantly worse on these items. Thus, the presence of the subjacency violations in the UNNAT language affected the learning of the *language as a whole*, not just the SUB items. From the viewpoint of language evolution, languages such as UNNAT would loose out in competition with other languages such as NAT because the latter is easier to learn.

Computational model

In principle, one could object that the reason why we found differences between the NAT and the UNNAT groups is because the NAT group is in some way tapping into an innately specified subjacency principle when learning the language. To counter this possible objection and to support our suggestion that the difference in learnability between the two languages is brought about by constraints arising from sequential learning, we present a set of connectionist simulations of our human data.

For the simulations, we used simple recurrent networks (SRNs; Elman, 1991) because they have been successfully applied in the modeling of both non-linguistic sequential learning (e.g., Cleeremans, 1993) and language processing (e.g., Christiansen, 1994; Elman, 1991). SRNs are standard feed-forward neural networks equipped with an extra layer of so-called context units. The SRNs used in our simulations had 7 input/output units (corresponding to each of the 6 consonants plus an end of sentence marker) as well as 8 hidden units and 8 context units. At a particular time step t, an input pattern is propagated through the hidden unit layer to the output layer. At the next time step, $t+1$, the activation of the hidden unit layer at time t is copied back to the context layer and paired with the current input. This means that the current state of the hidden units can influence the processing of subsequent inputs, providing an ability to deal with integrated sequences of input presented successively.

Forty networks with different initial weight randomizations (within ± .5) were trained to predict the next consonant in a sequence. The networks were randomly assigned to the NAT and UNNAT training conditions, and given 20 pass through a random ordering of the 30 training items appropriate for a given condition. The learning rate was set to .1 and the momentum to .95. Following training, the networks were tested separately on the 30 grammatical and 30 ungrammatical items (again, according to their respective grammar). Performance was measured in terms of how well the networks were able to approximate the correct probability distribution given previous context. The results are therefor reported in terms of the Mean Squared Error (MSE) between network predictions for a test set and the empirically derived, full conditional probabilities given the training set (Elman, 1991).

Results and discussion

The results show that the NAT networks had a significantly lower MSE (.185; SD: .021) than the UNNAT networks (.206; SD: .023) on the grammatical items ($t(38)=2.85$, $p<.01$). On the ungrammatical items, the NAT nets had a slightly higher error (.258; SD: .036) compared with the UNNAT nets (.246; SD: .034), but this difference was not significant (*t<1*). This pattern resembles the performance of the human subjects where the NAT group was 11% better than the UNNAT group at classifying the grammatical items, though this difference only approached significance (*t(38)=1.10, p=.279*). The difference was only <3% in favor of the NAT group for the ungrammatical items (*t=1*). Also similarly to the human subjects, there was a significant difference between the MSE on the grammatical and the ungrammatical items for both the NAT nets (*t(38)=7.69, p<.001*) and the UNNAT nets (*t(38)=4.33, p<.001*). If one assume that greater the difference between the MSE on the grammatical (low error) and the ungrammatical (higher error) items, the easier it should be to distinguish between the two types of items, then the NAT networks would have a significantly better basis for making such decisions than the UNNAT networks (.072 vs. .040; *t(38)=4.31, p<.001*). Thus, the simulation results closely mimic the behavioral results, corroborating our suggestion that constraints on the learning and processing of sequential structure can explain why subjacency violations tend to be avoided: They were weeded out because they made the sequential structure of language too difficult to learn.

Conclusions

The artificial language learning results show that not only are constructions involving subjacency violations hard to learn in and by themselves, but their presence also makes the language as a whole harder to learn. The connectionist simulations further corroborated these results, emphasizing that the observed learning difficulties in relation to the UNNAT language arise from non-linguistic constraints on sequential learning. When language itself is viewed as a dynamic system sensitive to adaptive pressures, natural selection will favor combinations of linguistic constructions that can be acquired relatively easily given existing learning and processing mechanisms. Consequently, difficult to learn language fragments such as UNNAT will tend to disappear. In conclusion, rather than having an innate UG principle to rule out subjacency violations, we suggest that they may have been eliminated altogether through an evolutionary process of linguistic adaptation constrained by prior cognitive limitations on sequential learning and processing.

References

Christiansen, M. H. (1994). *Infinite languages, finite minds: Connectionism, learning and linguistic structure*. Unpublished doctoral dissertation, Centre for Cognitive Science, University of Edinburgh, U. K.

Christiansen, M. H. (2000). *Using artificial language learning to study language evolution: Exploring the emergence of word order universals*. Paper to be presented at the Third Conference on the Evolution of Language, Paris, France.

Cleeremans, A. (1993). *Mechanisms of implicit learning: Connectionist models of sequence processing*. Cambridge, MA: MIT Press.

Elman, J.L. (1991). Distributed representation, simple recurrent networks, and grammatical structure. *Machine Learning, 7*, 195–225.

Kirby, S. (1998). *Language evolution without natural selection: From vocabulary to syntax in a population of learners*. Edinburgh Occasional Paper in Linguistics, EOPL-98-1.

Newmeyer, F. (1991). Functional explanation in linguistics and the origins of language. *Language and Communication*, *11*(1/2), 3–28.

Pinker, S., & Bloom, P. (1990). Natural language and natural selection. *Brain and Behavioral Sciences*, *13*(4), 707–727.

Vocal production in nonhuman mammals: Implications for the evolution of speech

W. Tecumseh FITCH

Dept. of Organismic & Evolutionary Biology
Harvard University

ABSTRACT

Human speech is characterized by rapid, precise movements of vocal tract articulators (lips, tongue, jaw, velum, larynx). The resulting changes in the shape of the supralaryngeal vocal tract (specifically its cross-sectional area function) leads to the dynamic pattern of formant variation which typifies human vocal communication. Several previous studies based on video analysis of lip movements suggest that animal vocalizations also involve such movements. However, anatomical studies comparing the vocal tract morphology of humans with non-human mammals suggest that the human vocal tract is fundamentally different from that of all other mammals. In particular, the resting location of the standard mammal larynx is high in the throat, and typically engaged in the nasopharynx, allowing animals to swallow fluids and breathe simultaneously. This position, and ability, also typifies human newborns.

In contrast, the resting position of the larynx in adult humans is much lower in the throat. While this makes it impossible for us to engage the larynx in the nasopharynx, and thus to breathe and swallow simultaneously, it does appear to make possible a wider variety of vocal tract shapes, and thus speech patterns, than would otherwise be attainable. In particular, the “descent of the larynx” that occurs in human ontogeny, gives adults a vocal tract with a horizontal oral tube and a vertical pharyngeal one. This two-tube vocal tract allows the production of quantal vowels such as /i/, /a/ and /u/, that feature in the vowel systems of most human languages. Because the anatomical data suggest that animals do not have such a two-tube vocal tract, it has long been believed that they are anatomically incapable of producing the vowels that occur in human languages. Thus, it has been argued that the “decent of the larynx” was a key innovation that occurred during human evolution, without which the evolution of spoken language would have been impossible. However, the whole argument is based on observations of the static anatomy of dead specimens, and there has previously been little information on the capabilities of mammal vocal tracts in the living, vocalizing animal.

In this talk I present the results of cineradiographic (x-ray video) analysis of mammal vocalizations which suggest that animal vocal tract movements are more pronounced than was previously suspected. In particular, all of the mammals we have examined (dogs, pigs, monkeys and goats) lower the larynx into the oral cavity during loud vocalizations, and in some cases (e.g. dogs) this dynamic “descent of the larynx” is substantial, and results in a long and well-defined pharyngeal tube in the vocal tract. Thus, during vocalization, these mammals have a vocal configuration much more similar to those of adult humans than is the resting configuration. In particular, the capacity to form a two-tube vocal tract, with a well-defined pharyngeal tube, appears to be a much more general in mammals than previously imagined. In contrast to some earlier claims, all of our subjects also appear capable of closing of the nasal cavity by raising the velum, and thus producing non-nasal vocalizations. Thus, these data suggest that 1) the anatomy of dead animals does not provide a reliable guide to their dynamic

vocal capabilities, and 2) the anatomical constraints placed on animal vocalizations by their vocal anatomy have been overemphasized.

Of course, dogs and pigs do not speak, nor do they produce vowels such as /i/. I suggest that this results from deficiencies at the motor control level rather than their static anatomy. Further, human resting anatomy is different from that of most other mammals, raising the question of how and why this unusual low position (with its attendant cost of increased risk of choking) evolved. I suggest that the laryngeal lowering that appears to typify most mammals represents a "preadaptation" to speech, and that early hominids dynamically lowered the larynx to vocalize just as do most other mammals. It is possible that laryngeal lowering played a role in exaggerating the vocal impression of size conveyed by the voice, as appears to be the case in roaring vocalizations of deer. The increased formant range made possible by this dynamically-created two-tube vocal tract, under pressure for increased vocabulary size (and increased distinctiveness of phones), was then exploited via simultaneous movements of the tongue body. This dynamic two-tube vocal tract thus represents a putative intermediate stage in the evolution of speech. Finally, under pressure for increasingly detailed control (e.g. for fricatives, flaps, etc), or alternatively due to increased costs of laryngeal lowering by more and more frequent speech, the larynx gradually assumed its present day resting position low in the throat.

These data suggest that, rather than representing an abrupt Rubicon for spoken language, the descended larynx of humans is probably the end product of a long and gradual process of refinement under selection for increasing vocal complexity and vocabulary size. By making use of a pre-existing vocal potential, present in our mammalian ancestors, the evolution of the human speech apparatus would thus be a nice example of the type of gradualistic exploitation of "preadaptations" envisioned by Darwin. Furthermore, these data indicate that to better understand human evolution, we need a much richer understanding of the capabilities of our nonhuman relatives. There are myriad similarities and differences between human language and animal vocal communication, and a broad comparative perspective is a prerequisite for understanding which differences really make a difference.

Evolving the linguistic mind

Keith Frankish

The Open University, UK.
k.frankish@open.ac.uk

There has been growing interest recently in the so-called *cognitive conception* of language – the idea that some human thought processes constitutively involve the representational resources of the language faculty – or, more colloquially, that we can *think in* natural language. This view is, I think, very attractive: there are theoretical reasons for endorsing it and introspection supports it, too. However, I shall not be defending the cognitive conception here. Instead, I shall be asking how a language-based cognitive system – a *linguistic mind*, as it were – might have *evolved*. I begin by outlining some evolutionary criteria which a satisfactory version of the cognitive conception must meet. I then look at some recent versions of the doctrine and ask which best satisfies them. This will give us a reason for preferring that version, should we decide to endorse the doctrine in the first place.

An evolutionary perspective

Many questions arise once we adopt an evolutionary perspective. For the present, however, I shall concentrate on just two:

1) How did natural language become involved in central cognition? It is widely accepted that the language faculty is a modularized, peripheral system, which is relatively encapsulated from the rest of cognition and which originally evolved for communicative purposes. How did such a system come to play a role in *central* cognition (that is, in flexible, intelligent, nonencapsulated, conscious thought)?

2) When did language-based cognition evolve? Most writers agree that grammatical language evolved some time within the last quarter of a million years. This does not leave much time for the *subsequent* development of language-based cognitive mechanisms – not, at any rate, if this would have involved alterations in neural anatomy.

Any satisfactory version of the cognitive conception must address these questions and any version which conspicuously lacks the resources to do so can be ruled out in advance. So, for example, we can rule out the hypothesis that a hardwired language-based cognitive processor evolved subsequently to the emergence of language. There would simply have been no time for such a system to develop. There remain, however, a number of viable candidates and the bulk of this paper is devoted to assessing them.

Language and the off-line mind

According to Derek Bickerton, language and central cognition *co-evolved*.[1] Structured non-demonstrative thought ('off-line thought' as Bickerton calls it) requires a system of schematic representations and a set of combinatorial principles defined over those representations. Language, too, requires a set of schematic representations and a combinatorial syntax, and parsimony suggests that the same neural resources play both roles.

---

[1] D. Bickerton, *Language and Human Behaviour* (Seattle: University of Washington Press, 1995).

This suggestion has attractions. If Bickerton is right, then there is no problem of *how* language got involved in central cognition: it *is* central cognition – at least, in so far as central cognition involves off-line thinking. The 'when' question is also dispelled. There was no need for further adaptation *after* the development of language; the development of fully grammatical language *was* the development of structured off-line thought.

The proposal also has some serious drawbacks, however. I will mention two. First, it is unlikely that language possessed cognitive functions from the very start. For it would have been simpler to build a purely *communicative* language system than to build one which had both communicative *and* cognitive functions. A communicative language system requires only syntax, phonology, and comprehension systems, together with a lexicon. And while these could be adapted to play a role in cognition, they would not in themselves *constitute* a cognitive system. Additional subsystems would be required – in particular, some sort of central processor. (Syntax alone might give you structured *thoughts,* but not structure-sensitive *thought-processing*.) But if it would have been easier to construct a merely communicative language system than one that had cognitive functions, then we should expect the former to emerge *before* the latter – and not to co-evolve with it. Secondly, Bickerton ignores evidence for modular structure *within* central cognition. He tends to view the whole of human central cognition as language-based – and thus to suppose that it is a fairly recent system, with little inherited structure. Yet there is mounting evidence that we have lots of innate cognitive competences, realized in functionally distinct, partially encapsulated modules which have developed gradually over last million years or so, in many cases predating language. Such evidence tends to undermine Bickerton's solution to the 'how' question. If the linguistic mind is not the whole of central cognition, then how is it related to the rest of it?

Language and the modular mind

Is the cognitive conception compatible a modularist view of central cognition? Peter Carruthers suggests so, building on suggestions by Steve Mithen.[1] According to Mithen, the human mind developed in three phases. In phase 1, it consisted of a rudimentary general purpose problem-solving system. In phase 2, this was supplemented by a number of self-contained domain-specific modules, which were fast but inflexible, and did not communicate with each other or with general intelligence. Finally, in phase 3, there was a growth of 'cognitive fluidity'. The previously isolated central intelligences began to communicate with each other and with general intelligence – either through direct channels, or through the mediation of a metarepresentation supermodule. Carruthers argues that this picture naturally supports a version of the cognitive conception. Even while remaining internally isolated, he points out, the central modules would have formed input-output links with the language faculty. Natural language would then have been the obvious vehicle for *inter-modular* information transfer, once the internal barriers started to come down. Natural language would thus have come to serve as a cognitive *lingua franca*.

I have two worries about this proposal. First, it does not amount to a full-blooded vindication of the cognitive conception – not, at least, if we take that doctrine to involve the claim that language can act as a medium of *inference* as well as *thought*. In Carruthers's scenario all the real inferential work done within modules, in their own internal representational media, and natural language functions merely as a conduit between them. Secondly, the proposal does not

---

[1] S.Mithen, *The Prehistory of the Mind* (London: Thames and Hudson, 1996). P.Carruthers, 'Thinking in Language: Evolution and a Modularist Possibility' in P.Carruthers and J.Boucher, *Language and Thought: Interdisciplinary Themes* (Cambridge: CUP, 1998).

fully resolve the 'how' question. A neural *lingua franca* may be a *necessary* condition for inter-modular co-operation, but it is not *a sufficient* one. Coherent trains of thought do not just spring into existence spontaneously – a problem has to be identified and the various modular resources deployed intelligently to its solution. Some kind of *executive* would thus be needed to marshal the problem-solving resources of the different modules and to co-ordinate their outputs. Moreover, this system would need to process sentences in a way that was sensitive to their semantic properties. But then it starts to look like the sort of hard-wired language-based cognitive processor whose existence we have already ruled out.

Language and the virtual mind

We been thinking of the linguistic mind and the processing mechanisms that support it as *part* of the brain. But perhaps this is wrong. Perhaps it is more like a *program* running on the brain – a feature of our mental software, not of our neural hardware. Such a view has been defended by Daniel Dennett.[1] The modern human mind, Dennett claims, is not a biological system at all, but a *virtual machine* – the product of learned behaviours (Dennett calls them 'good tricks' or 'memes') which have reprogrammed our biological brains. The behaviours in question, Dennett suggests, are linguistic. We acquire virtual minds by talking to ourselves – by producing, rehearsing and rearranging sentences in overt or silent soliloquy. This stream of inner verbalization transforms the activity of the biological brain, causing its parallel multi-track hardware to simulate the behaviour of a serial, single-track processor, operating upon natural language sentences. Dennett calls this softwired system the *Joycean machine*.

This story is particularly attractive from our current perspective. There is no special problem about how or when the virtual mind evolved: its development involved a process of *memetic* or *cultural* evolution, not the emergence of new neural structures. Still, the story will not do as it stands. The problem lies in the way the Joycean machine is supposed to work. According to Dennett, the key mechanism is one of *self-stimulation*. Inner speech is channelled through a feedback loop from speech production to speech comprehension. Internally generated sentences are then processed by the comprehension system just like externally produced ones, often evoking similar responses. So, for example, questioning yourself may prompt an instinctive verbal reply, containing information that would otherwise have been inaccessible. Dennett suggests that neural subsystems routinely compete for control of the vocal system and the self-stimulatory mechanisms it supports. As a result, the Joycean machine comes to act both as a bulletin board, where locally stored information is made globally available, and also as sort of virtual executive, focusing attention, marshalling resources and co-ordinating the activities of different subsystems.

It is likely that inner verbalization does have a self-stimulatory function of this sort.[2] But self-stimulation cannot be all there is to the linguistic mind. For one thing, it is doubtful that it could generate sustained trains of intelligent thought. Self-stimulation might help to produce some regularity and consistency in one's inner verbalizations, but it is hard to see how it could yield coherent inferential sequences of the kind involved in dedicated problem-solving. In such cases, it seems, our inferential subsystems are not merely *competing* for vocal control, each shouting out its favoured (and not particularly bright) solution. Rather they *co-operating*, each subordinating its activity to a global objective. And it is hard to see how they could be induced to do this without executive supervision of some sort. Secondly, self-stimulations will not have the cognitive role typical of linguistic thought. Consider the sort of cases that lend intuitive

---

[1] D.C.Dennett, *Consciousness Explained* (Boston: Little, Brown, 1991).

[2] R.M.Diaz and L.E.Berk (eds), *Private Speech* (Hillsdale, NJ: Lawrence Erlbuam, 1992).

support to the idea that we can think in language. I notice that the steering on my car is uneven and say to myself 'The wheel alignment needs checking.' Here, it seems, I am not *instructing* or *encouraging* myself to think that the alignment needs checking; I am *judging* that it does. And this judgement may have long-term effects – say, in getting me to the garage the following day – which a transient self-stimulation would lack.

Language and the supermind

Dennett gives us part of the story, then, but important aspects of the linguistic mind remain unaccounted for. Can we complicate his picture in order accommodate these features? I think so.[1] The trick is to think of linguistic reasoning as, to some extent, under *personal control*. The linguistic mind, I suggest, is indeed a virtual one, constructed by the discovery and transmission of good tricks. But these include, not just inner verbalization, but also various *meta-cognitive* and *meta-linguistic* skills. We do not only speak to ourselves, I suggest; we also adopt attitudes *towards* our inner verbalizations and perform explicit inferential operations *upon* them. In particular, we adopt some as explicit premises and others as explicit goals, and manipulate them so as to construct chains of reasoning, using learned inferential skills. I have suggested elsewhere that these activities constitute a distinct level of mentality, which is intentionally formed and sustained, and which constitutively involves natural language. I call this the *supermind*.

This proposal retains all the advantages of Dennett's. Like the Joycean machine, the supermind is the product of memetic and cultural evolution, rather than changes in the brain. Indeed, there is a plausible story to tell about just how it emerged. The metalinguistic and metacognitvie skills needed to develop a supermind – the ability to think about one's thoughts and words and to articulate cogent trains of argument – are just the skills needed for engaging in reasoned argument with one's peers. And as such, their possession would have carried huge benefits in early human society – for example in securing an equitable distribution of resources and in attracting mates (the 'big man' hypothesis). There would thus have been strong independent pressure for their development. And supermental abilities would have followed naturally in their wake. Humans would have begun to develop linguistic minds as soon as they started to *internalize* their skills in interpersonal argument, reasoning and debating with themselves. (The supermind, we might say, is not so much a *Joycean* machine as a *Beckettian* one.)

And supermind theory can, I believe, resolve the problems facing Dennett's account. Take the question of executive control. For self-conscious agents, equipped with metacongitive skills, problem-solving will assume a dual aspect. They will be able to think, not only about the first-order problem of what to do or think next, but also about the *metaproblem* of how to solve that problem. Suppose they have some general ideas about how to solve this metaproblem. So, they want to evaluate candidate hypotheses as they occur to them, preferring those that harmonize well with premises and goals they have previously endorsed, and rejecting those that conflict with them. And as a subgoal they want to trace out the implications of each hypothesis, searching for data that might confirm or refute it, or indicate how it should be revised. These desires then drive their subsequent attempts to tackle the first-order problem. As various candidate hypotheses occur to them (thrown up, let us suppose, by submodules), they set to work evaluating them, in line with their metacognitive goals. They persist in this, refining and

---

[1] The following sketch draws on my 'Natural Language and Virtual Belief' in P.Carruthers and J.Boucher (eds), *Language and Thought: Interdisciplinary Themes* (Cambridge: CUP, 1998) and on my 'A Matter of Opinion' *Philosophical Psychology* 11 (1998): 423-42.

complicating their hypotheses, until they reach a solution that satisfies their goals. In short, their metacongitive attitudes *regulate* their first-order problem-solving activities. In effect, they can act as their *own* central executives, marshalling and directing their low-level cognitive resources.

Supermind theory can also explain how inner verbalizations can assume a direct cognitive role, rather than a merely self-stimulatory one. We can *decide* that an inner verbalization will have the role of a thought by deciding to adopt it as premise and to use it as a basis for inference and action. Executing such a decision will involve using one's explicit reasoning skills to make sure that the verbalized proposition has the appropriate inferential role – for example, by using it as a premise in one's syllogisms. Language-based reasoning will thus be genuinely computational – though the computations in question will be carried out at an explicit, personal level.

Of the candidates reviewed, then, I suggest that supermind theory is the best placed to provide a full-blooded and neurologically plausible defence of cognitive conception. If we have indeed developed a linguistic mind, then *this* is form it is most likely to have taken, given the constraints on its evolution.

Creolization as the superposing of functional structure on lexical projections

Karl Erland GADELII

Göteborg University
gadelii@african.gu.se

This paper argues that Bickerton (1990, 1996,1998) is basically correct when he claims that the development from proto-language to "modern language" equals the development of pidgin language into creole.

Basing myself on a number of constructions in French pidgins, creoles and Hexagonal French, I argue that pidginization of Hexagonal French, leading to French pidgins, involves restructuring of the linguistic system to the extent that grammatical morphemes, free and inflectional, are lost. The resulting pidgin has a more or less topic – comment-like structure, where highly polysemous lexical items are concatenated to each other and communicative success highly depends on use of contextual information. It is not unreasonable to assume, as Bickerton does, that human proto-language must have largely resembled such a pidgin.

When creolization takes place, the pidgin is enriched with a grammatical apparatus so as to be able to fulfil the functions required of any natively spoken first language. In the terms of chomskyan grammatical theory, creolization can be seen as the superposing of a hierarchy of functional projections on top of ordinary lexical projections. Consequently, pidgins can be thought to have lexical projections only, upon which the functional structure emerging during creolization is built. Now functional projections in languages can be "visible" in two ways, namely by having linguistic material base-generated in them or by having lexical items moved into them. I will here call the first phenomenon "lexicalization" and the second "feature checking by movement". The main argument of the paper is that creoles never initially use movement to check grammatical features, but rely on "lexicalization". This means that functional meanings are expressed by free grammatical morphemes in creoles. As examples of these, we may mention classical creole phenomena such as preverbal TMA particles and various kinds of determiners, which by the way in most cases have a more articulate structure than in the original lexifier language. It is reasonable to assume that Man‘s first "modern language" also had this quality, i.e. that grammatical meanings were expressed by free morphemes and not inflectional ones, and that grammatical features were checked by "lexicalization" of these free morphemes and not by movement. It should however be noted that not all function words which are lost in pidginization immediately reappear in creoles. For example, it would seem that some prepositions and complementizers present in the superstratum are absent in the creole. This difference between these items on the one hand and TMA-markers and determiners on the other will be further analyzed.

Once creolization has taken place, the creole will develop as any natural language does, i.e. with the gradual erosion of free morphemes into bound ones and the increasing need for verb and noun movement to check morphosyntactic features. Incidentally, the picture sketched above denies grammaticalization as an ingredient increolization: it is rather believed that grammatical items such as the tense marker *bin* in English creoles or *té* in French ones are directly inserted in the functional projections in question, i.e. the position is "lexicalized" with a functional item whose linguistic provenance is basically irrelevant, but which has not developed out of a previously lexical item.

References

Bickerton, Derek. 1990.*Language and species*. Chicago: University ofChicago Press.

Bickerton, Derek. 1996. *Language and humanbehaviour*. London: UCL Press Unlimited.

Bickerton, Derek. 1998. Catastrophicevolution: The case for a single step from protolanguage to full humanlanguage. In *Approaches to the evolution oflanguage*. Hurford, James, Michael Stuart-Kennedy & Chris Knight, 341–358.Cambridge: Cambridge University Press.

The evolution of syntactic categories

David GIL

ABSTRACT

What did primitive languages look like? One way to answer this question is to examine patterns of cross-linguistic diversity exhibited by modern languages of the world.

There is now general agreement that today's languages are all equally advanced on the evolutionary scale; there are no more primitive languages left. There is also consensus that, when viewed holistically, all languages are of roughly equal complexity. Nevertheless, wihin particular grammatical domains, it is indeed possible to identify different levels of complexity across languages. To cite one well-known example: by definition, the morphology of isolating languages is less complex than that of inflectional languages, which in turn is less complex than that of polysynthetic languages. A common assumption is that languages that appear to be simple in some domains compensate for this by being more complex in other domains.

In this paper, it is suggested that cross-linguistic variation in complexity within particular grammatical domains may provide insights into possible paths of language evolution. Specifically, it is argued that within any given grammatical domain, if one finds a situation of greater simplicity in one language than in another, then that simpler situation may, under certain conditions, be considered to be characteristic of a prior stage in the evolution of language.

In this paper, the above approach is applied to the domain of syntactic categories, or parts of speech. In particular, it is claimed that languages differ greatly in their inventory of syntactic categories. More specifically, some languages exhibit an extremely simple, or impoverished syntactic category inventory. Such inventories, it is argued, may be taken to reflect the situation that prevailed in an earlier stage of language evolution.

The first part of this paper sketches the outline of a theory of syntactic categories designed to facilitate the formulation of cross-linguistic generalizations governing possible syntactic category inventories. As in categorial grammar, syntactic categories are derived from other syntactic categories by means of category-formation operators. Unlike categorial grammar, however, the present theory posits the existence of a single primitive syntactic category S^0, from which all other syntactic categories are derived, by recursive application of two category-formation operators: a unary operator “kernel”, forming category X^{n+1} (the “kernel of X”) from category X^n, plus the familiar binary operator “slash”, forming category X/Y (“X slash Y”) from categories X and Y.

The above framework permits the formulation of typological generalizations governing possible inventories of syntactic categories in different languages. Specifically, for any language L:

(4) If X is a category in L, then all the categories in the derivational history of X are categories in L. Similarly, if X is an open category in L, then all the categories in the derivational history of X are open categories in L.

(5) If X and Y are categories in L, then L possesses constructions formed with X and Y.

(6) If X/Y is an open category in L, then Y is the kernel category of X.

The second part of this paper surveys the evidence to the effect that some languages possess almost maximally impoverished inventories of syntactic categories. For example, Tagalog has just one open syntactic category S^0 and one closed syntactic category S^0/S^0, while Vietnamese has just one open syntactic category S^0 and two closed syntactic categories S^1 and S^0/S^1.

The results of this paper thus show that some contemporary languages can get by with a very simple inventory of syntactic categories. In doing so, they suggest what the inventory of syntactic categories might have looked like at a primitive stage in the evolution of language. More generally, the categorial-grammar based theory of syntactic categories presented in this paper provides for a family tree of syntactic categories of increasing complexity: a scheme which may be considered as a recapitulation of the process by which these categories evolved in human language. Thus, this paper shows how, by examining the diversity of languages spoken today, it is possible to infer how language evolved from simple beginnings to contemporary complexity.

The evolution of communicative interaction systems: A formal semantics perspective[1]

Jonathan GINZBURG

Computational Linguistics Programme
Dept. of Philosophy
King's College, London
The Strand, London WC2R 2LS, UK

Linguistics Programme
Dept. of English
Hebrew University of Jerusalem,
Mount Scopus Jerusalem 91905, Israel

jonathan.ginzburg@kcl.ac.uk

Introduction

Despite the fact that significant advances have taken place over the last 25 years in the study and understanding of the semantics and pragmatics of natural languages (NL's) using tools from formal logic and theoretical computer science (for a survey see e.g. Lappin 1996) such work has to date had little impact on the study of evolution of language. This is not surprising, given the focus in such work on classifying semantic properties in terms designed for disembodied formal languages where notions of interaction and context dependence had little place. However, in recent years in the wake of the development of the *dynamic* perspective on meaning, there has been significant interest in applying formal semantic techniques to the study of human dialogue. As a result various linguistic phenomena have emerged whose description and analysis require taking a perspective on language which, in a nutshell, views its use in interaction as basic. These include *Collaborative utterances*(Clark 1996), *the Turn Taking Puzzle* (Ginzburg 1997), and *Cross speaker anaphora* (Dekker 1997).[2]

In this paper, I will argue that this perspective and the techniques emerging from the formal semantic study of the structure of human cognitive states utilized in communicative interaction shed new light on the issue of discontinuity between the human communicative interaction system (CIS) and the CIS's of other species. I will propose two measures of complexity for a CIS. One based on *attitudinal complexity,* that indicated by the type of cognitive attitude predicates structuring the information states apparently utilized *inter alia* in communicative interaction. The other based on the complexity of the messages employed in the CIS, a measure I will dub *contextual complexity* . I will suggest that both measures provide at least *prima facie* evidence for continuity between the CISs of primates and those of humans. More importantly, this way of looking at CISs offers prospects for locating criterial dividing points between the

---

[1] Supported by the UK Economic and Social Research Council (grant number R000222969 Phrasal Utterance Resolution in Dialogue

[2] Collaborative utterances are exemplified in (i). The turn taking puzzle is exemplified in (ii): a context where one speaker keeps the turn results in different ellipsis resolution possibilities from a context in which speakers change across turns. Cross speaker anaphora is illustrated in (iii): the phenomenon vitiates syntactic views of anaphora, which involve some notion of copying of a previously used expression.

(i) A: Did she B: leave the room?

(ii) A: Who attacked Bill? When? (=when did whoever attack Bill do so); A: Who attacked Bill? B: When? (= Please clarify your question: when did the attack you're asking about take place?

(iii) A: An actor was sitting on the bench. B: He's not an actor, he's a tramp.

currently existing adult NL–based CIS and simpler human-neonate, proto-human, cetacean, and primate systems.

Cognitive states and attitudinal complexity

What is the structure of the cognitive states used in human conversational interaction? Simplifying somewhat, a tentative answer to this question which has been emerging in recent formal semantic and pragmatic work on dialogue (see e.g. Ginzburg 1996,1997,1998, Traum et al 1999) is to model these as feature structures of the form:

$$(1)\ \left[\begin{array}{ll} \text{FACTS} & \textit{set propositions} \\ \text{QUD} & \textit{partially ordered set questions} \\ \text{LATEST - MOVE} & \textit{action} \end{array}\right]$$

Here *facts* represents the knowledge that accumulates in the context during a conversation, *qud* is a partially ordered set of questions[1], which represents the issues under discussion at a particular point in a dialogue, and l*atest-move* represents the most recent conversational move undertaken in the dialogue. Using this cognitive architecture, allows one to view conversational interaction dynamically as a sequence of cognitive states which are updated as a consequence of *dialogue moves*. These include queries, assertions, and various moves such as acknowledgements, corrections, and clarifications that relate to ``metalinguistic" interaction about the grounding process of utterances, the feedback which conversationalists provide each other about whether an utterance has been understood or requires clarification (Clark 1996, Traum 1994, Ginzburg 1998). States of this kind can also be used to explain linguistic phenomena such as anaphora and ellipsis resolution possibilities in dialogue.

What underlies cognitive states of the form sketched in (1) is the fact that human cognitive states can be structured in terms of a number of distinct attitudes to the external environment, including minimally *belief, wonder, and plan.* Here *belief* is the familiar attitude predicate, relating an agent to a proposition, which if true provides some descriptive condition on an external situation *.wonder* is the attitude which relates an agent to a question, that semantic unit which encapsulates in a consistent way mutually exclusive ways of describing an external situation (see refs in footnote 3), and *plan* is the attitude predicate relating an agent to a sequence of actions (identifiable for current purposes with a temporally ordered sequence of (propositional) state descriptions.). I dub a CIS which exploits states as in (1) a *Discursive Informing System (DIS)*: agents who communicate within a DIS can inform each other of facts but also discuss questions and dispute claims.

How might one explain the evolutionary chain that has lead to the emergence of DISs? One possible answer is to view the complexity of CISs as correlating with the *abstractness* or *defeasibility* of the information processed by agents in a given CIS. At the lowest end one would locate *Action Registration Systems* (ARS). In such a system information is entirely encapsulated in the act per-formed by the utterer such as an act of greeting or threatening. Such systems are well known even among non-primate mammals such as geese or wolves, as documented already in early ethological literature (e.g. Lorenz). The basic architecture required for communication using such states is the possession of ability to correlate message tokens with discrimination of situations into various distinct types.

---

[1] The notion of question appealed to here is one that has emerged in semantic research concerning interrogative sentences (see Ginzburg 1996, Groenendijk and Stokhof 1997). These are logical entities which *inter alia* encapsulate the various mutually exclusive answers which can resolve a question.

A level above ARSs are *Pure Informing Systems* (PUS). In such a system the acts available involve messages which classify the situation in which utterer and addressee find themselves, most prototypically classifying it as dangerous in some way. Such a system involves one level of abstractness in that of course the reliability of the information depends on the agent providing it. This introduces the potential for providing incorrect information and, consequently, for inconsistency. Inconsistency between the information arising in the message and information arriving from a distinct source. For an agent in a PUS it seems apposite to attribute a cognitive attitude of *confusion* if a situation is encountered which is apparently classifiable in incompatible ways, then the animal behaves as if something is very wrong (its cognitive system (temporarily)``crashes'', various symptoms of panic are evinced etc).

A significant increase in the complexity of an information state is one that has evolved from merely encoding confusion, to one that manifests *wondering*: the ability to consistently represent the existence of incompatible ways of categorizing a situation, that which, as noted above, is represented by a question. wondering is the key to the evolution of a notion of reliability of a signal: if a signal arrives on the basis of which the hearer is supposed to classify the context situation with type T_1 but the hearer also has information requiring her to classify the context with type T_2, incompatible with T_1 if her information state is advanced enough to encode wondering, the hearer can react rationally, and weigh whether to accept T_1 or to reject/ignore it. From the information provided by e.g. Cheney and Seyfarth (Cheney& Seyfarth 1988}), we can conclude, for instance, that vervets' information states encode some notion of wondering. Similarly for Chimpanzees, given evidence for their ability to reason about deception (e.g. de Waal 1986). A CIS where the agent can wonder will be called a *Pondered Informing Systems* (PIS).

Qualitatively, the difference between a PIS and a DIS is that in the latter the agent can go further than simply wonder, they can actually externalize their wondering and engage in discussion of a given question, e.g. in whether information provided by an act of informing is correct or not. A parameter which can distinguish the complexity of two DISs is the cardinality of the *qud* (questions under discussion) attribute they carry: it is straightforward to demonstrate from conversations that the cardinality of *qud* for adult humans can be higher than 3. Agents with simpler DISs such as adult neonates seem to be limited to a *qud* of at most 2.

Evaluating a CIS by message complexity

Attitudinal complexity is a hopefully a measure which can provide some help in characterizing the complexity of CISs. The increase in complexity in the hierarchy sketched above (ARS < PUS < PIS < DIS) requires an increase in cognitive complexity and equally involves enhancements that are clearly adaptive. Attitudinal complexity is, nonetheless, only one component in evaluating a CIS. An additional component, which is at least partly independent, is a measure of the complexity of the *messages* communicated within a given CIS, taking as a starting point the increasingly detailed descriptions achieved by semantic analysis of the human CIS.

The measure of message complexity I develop in the extended version of this paper is one I call *contextual complexity*. It is motivated by the need to (a) define message complexity in a way which does not presuppose syntactic complexity (cf. the suggestive evidence for syntactic complexity in relatively `simple' CISs such as that of gibbons (Ujheyli 1998)) (b) the importance of integrating contextual factors in the calculation of message import. In linguistic semantics an influential way of thinking of meanings deriving originally from the work of David Kaplan and situation semantics (e.g. Kaplan 1989, Barwise and Perry 1983) is as relations between an utterance situation and potentially other situations in which certain parameters (the contextually dependent parameters) get instantiated and the content communicated.With this view of meanings, we can define a measure of complexity on meanings correlating increased complexity with the increase in situations used in calculating the meaning. A *homo-situational* message system involves merely contextual relativization to the utterance situation. This is exemplified in the following messages:

(2) *hi*: given spkr,hearer, spkr-greets-hearer;*danger*: given situation, danger(situation) *eagle*: given situation, $danger_{eagle}$(situation);*what's up*: given situation, λP.P(situation)

In a *bi-situational* message system there is one additional situation relative to which contextual relativization is defined:

(3) *Pogo is absent:* given situation u and reference to p in s, absent(p) is true of u

In a *dynamic-situational* message system relativization can crucially involve the previous utterance situation. In other words, calculating the values of a meaning require a buffer which involves the previous utterance situation. Such a system is needed, for instance, to handle linguistic phenomena such as anaphora and feedback such as clarification and correction.

References

Barwise, J. and J. Perry.: 1983, *Situations and Attitudes* MIT Press, Cambridge

Cheney, D. and Seyfarth, R.: 1988, "Assessment of meaning and the detection of unreliable signals by vervet monkeys", *Animal Behaviour,* 36, 477-486

Clark, H.: 1996, *Using Language*, Cambridge University Press, Cambridge

de Waal, F.: 1986, "Deception in natural communication of chimpanzees" In: R. Mitchell and N. Thompson (eds.), *Deception: Perspectives on Human and Nonhuman Deceit,* SUNY Press, Albany

Dekker, P.: 1997, "First Order Information Exchange", In: G. Jaeger and A. Benz (eds.),*Proceedings of MunDial 97* (Technical Report 97-106), Universitaet Muenchen Centrum fuer Informations- und Sprachverarbeitung, Muenchen

Ginzburg, J.: 1996, "Interrogatives: Questions, Facts, and Dialogue", In: S. Lappin (ed.), *Handbook of Contemporary Semantic Theory*, Blackwell, Oxford

Ginzburg, J.: 1997, *Structural Mismatch in Dialogue*, In: G. Jaeger and A. Benz (eds.), *Proceedings of MunDial 97*, (Technical Report 97-106)}}, pp 59-80, Universitaet Muenchen Centrum fuer Informations- und Sprachverarbeitung, Muenchen

Ginzburg, J.: 1998, "Clarifying Utterances" In J. Hulstijn and A. Nijholt (eds.), *Proceedings of TwenDial 98, 13th Twente workshop on Language Technology*, pp 11–30, Twente University, Twente

Groenendijk, J. and Stokhof, M.: 1997, "Questions", In J. van benthem and A. ter Meulen (eds.), *Handbook of Logic and Linguistics* , North Holland, Amsterdam

David Kaplan.: 1989 "Demonstratives: An Essay on the Semantics, Logic, Metaphysics, and Epistemology of Demonstratives and Other Indexicals", In: J. Almog et al (eds.) *Themes from Kaplan*, Oxford University Press, New York,

Lappin, S. (ed.): 1996, *Handbook of Contemporary Semantic Theory*, Blackwell, Oxford

David Traum.: 1994, *A Computational Theory of Grounding in Natural Language Conversations*. Unpublished PhD thesis, University of Rochester, David Traum and Johan Bos and Robin Cooper and Staffan Larsson, and Ian Lewin and Colin Matheson and Massimo Poesio 1999, "A model of dialogue moves and information state revision", TRINDI Deliverable 2.1, University of Gothenburg", Available from http://www.ling.gu.se/research/projects/trindi

Ujhelyi, M. "Long-call structure in apes as a precursor for language. In: J. Hurford, M. Studdert-Kennedy, and C. Knight (eds.), *Approaches to the Evolution of Language*, Cambridge University Press

Worden, R.: 1998, "The Evolution of Language from social intelligence", In: J. Hurford, M. Studdert-Kennedy, and C. Knight (eds.), *Approaches to the Evolution of Language*, Cambridge University Press.

From sensorimotor praxis and pantomime to symbolic propositions

Stevan HARNAD

Department of Electronics and Computer Science
University of Southampton
Highfield, Southampton
SO17 1BJ UNITED KINGDOM
harnad@soton.ac.uk
http://cogsci.soton.ac.uk/~harnad/

What lies on the two sides of the linguistic divide is fairly clear: On one side, you have organisms buffeted about to varying degrees, depending on their degree of autonomy and plasticity, by the states of affairs in the world they live in. On the other side, you have organisms capable of describing and explaining the states of affairs in the world they live in. Language is what distinguishes one side from the other. How did we get here from there? In principle, one can tell a seamless story about how inborn, involuntary communicative signals and voluntary instrumental praxis could have been shaped gradually, through feedback from their consequences, first into analog pantomime with communicative intent, and then into arbitrary category names combined into all-powerful, truth-value-bearing propositions, freed from the iconic "shape" of their referents and able to tell all.

The attendant increase in speed and scope in acquiring and sharing information can be demonstrated in simple artificial life simulations that place the old and new means into direct competition: Symbolic theft always beats sensorimotor toil, and the strategy is evolutionarily stable, as long as the bottom-level categories are grounded in sensorimotor toil.

If you have a preference for "hard evidence" you approach the problem of language origins at your own peril. It has been suggested that language is a kind of organ (Chomsky 1972; Pinker & Bloom 1990), but even if it is, it must share the fate of that other organ of which any language organ is surely only a part, the brain, namely, that it leaves no trace after its demise, or at least no trace that anyone has yet given a convincing functional (let alone cognitive) interpretation (Holloway 1970, Wilkins & Wakefield 1995).

So in pondering the origins of the language organ, we cannot expect much help from the fossil record. But if soft organs deny us hard evidence, surely their behavioral functions are even more evanescent: "*Verba volunt, Scripta manent*," to be sure, but writing surely arrived on the scene too late in the day to help us (Harnad 1991).

Some have turned, in desperation, to other traces: tools, weapons, drawings (Isaac 1987). They have hypothesised that either language was necessary to make and/or use these artifacts or that there is some formal or functional commonality or co-dependency between the capacity to make or use them and the capacity to speak. But hypotheses are hardly evidence, let alone hard evidence, and in and of themselves, artifacts are just artifacts.

Others have looked to contemporary species, rather than ancient ones, for a clue (Greenfield 1992), but hunting and tool-making and -use can hardly be said to be garrulous activities today, so they are not very compelling evidence for loquacity, let alone its origins, long ago. Moreover, contemporary evidence for rudimentary tool-making and use in animals that have neither language nor the ability to acquire it goes against the idea that these functions have much in common.

The other prominent functional commonality that has been proposed is between language and consciousness (Rolls 1997, 2000). Some have thought language was a prerequisite for consciousness; but if this were true, it would deny pain to all creatures who were incapable of expressing it in words, and surely that's wrong, although, again, there is no “hard” evidence for it, short of BEING the mute animal feeling the pain (Harnad 2000).

The impossibility of determining empirically whether or not a non-speaking creature is conscious – otherwise known as the “other minds” problem – does suggest one genuine point of commonality with the problem of language origins, however: Consider the problem of “demoting” explanations, when you try to attribute consciousness to an animal or a machine: You can always interpret a system “as if” it were conscious, even an inert, nonbehaving system like a book on a table. You can say the book knows it's on the table, wants to be on the table, etc. Who's to say otherwise? The data are perfectly compatible with that interpretation, it's just that the interpretation seems both unnecessary and wrong. We can understand the book's being on the table without having to infer that it is conscious. If anyone claims the consciousness is necessary, we can ask why? There is no physical reason why a book needs to know it's on a table in order to be on a table.

Demoting conscious interpretations of inert systems is trivial. It's only a bit harder to demote dynamic, performing systems: Physics and engineering do not require, nor can they use, a conscious explanation of the functioning of a thermostat (“it feels hot,” “it wants to turn off the furnace”) or a car, an airplane or a computer. The performance of each of these systems is fully explained by mindless mechanisms.

The trick is to show where the need for a conscious explanation kicks in: What functional capacity CANNOT be explained without recourse to consciousness? Behaviour-analytic psychology gave us operant and Pavlovian conditioning, but do these require consciousness? It is easy to build devices, not very different from thermostats, whose performance is shaped by their history of associations or their history of reinforcements, especially in this era of neural nets and other computational models. So any behavior that can be reduced to an operant or a Pavlovian explanation has been demoted to a mindless explanation.

And no behavior seems to be immune to this sort of demotion: Show me a creature, human or nonhuman, that you think is managing to do something only with the aid of consciousness and it will invariably be easy to show that the same thing could be accomplished by a mindless mechanism merely shaped by the consequences of its behavior, just as Skinner would have said: Gorillas rubbing off from their foreheads yellow spots that they have seen in their mirror image (Gallup 1970)? All the data are there for a mindless learning mechanism to learn that correlation from the sensorimotor interactions with its mirror image: Happens too fast? Well then the same correlation, or a generalization of it, could have been “prepared” by evolution, likewise a mindless process.

The point is not to deny that gorillas are conscious and do recognize themselves in the mirror. It is the causal role of the consciousness in the explanation that keeps on turning out to be unnecessary, hence the demotion: The rule seems to be, whatever we happen to do mindfully, could just as well have been done mindlessly.

Now this is not a conference on the origin of consciousness. I have only introduced the ever-ready, mindless demoting explanation by way of analogy, for it has an exact counterpart in the case of language. Here is the heart of the analogy: Just as it is impossible to show that THIS is where what you can do with a mindless mechanism ends, and beyond this you can only go with a mind – there is no such point, and hence no functional explanation for why consciousness should ever kick in (though it certainly does) – there is likewise no point at which nonlinguistic praxis and pantomime end and linguistic propositions take over.

Let's define terms. “Praxis” just refers to our sensorimotor skills, the things we and other species need to be able to do in order to live and act out our lives: Finding food, eliminating wastes, avoiding predators, finding mates, etc., each according to the demands of our own species-specific niche. Nonlinguistic creatures share this nonverbal portion of our praxic repertoires (including the capacity to learn by operant and Pavlovian conditioning), and that covers a lot.

“Pantomime” is a special subset of praxis: It is social. Behavioral imitation (as opposed to anatomical mimicry) largely occurs between living creatures, not between living creatures and inert objects (Byrne & Russon 1998). But pantomime includes both automatic and deliberate imitation. A songbird may imitate the tune of its conspecifics mindlessly (who knows), but people at least, and perhaps other species as well, can “act out” in ways that are intended to get you to do and even think something. We know this in the case of people (I might cover my mouth and point when I mean for you to know that someone else is present), but we know from the demoting explanation of gorilla mirror-recognition that even pantomime can be explained as mindlessly as any other form of praxis.

So far, this is not the analogy I was promising; it is still the problem of consciousness. To appreciate the analogy, we first have to pass from pantomime to propositions. A few critical differences have to be borne in mind: A pantomime, like a picture, cannot be true or false. It can only be more or less like whatever it is a picture or pantomime of. Praxic gestures, whether pantomimic or just plain instrumental acts (Catania & Harnad 1988; Harnad 1996b), do not have truth values. They can of course be CONSTRUED as having truth values, but then they are construed as propositions. Propositions propose something. That's why they have truth values: What they claim or propose to be the case may or may not be the case. If not, then they are false.

Now when I put my hand on my mouth and pointed, you could have construed that as the proposition “There is someone else in the room,” and, if upon inspection, it turned out there was no one else in the room, you might want to say I had lied (Whiten & Byrne 1988). But in a court of law, so accused, I could claim that all I had done was put my hand on my mouth and pointed. I had never said – never uttered a verbal contract – to the effect that there was someone in the room. That was just your interpretation. I was merely performing a pantomime. I never intended you to construe it as a proposition.

By way of contrast, if I were in a crowded theater and I yelled “Fire,” I could be held liable for causing a stampede and causing injury if it turned out I had been crying wolf. It would not do me much good to claim (though it might be true) that I had merely been saying out loud words that rhymed with “dire,” and that I had not intended it to be construed as a proposition. If this distinction sounds legalistic rather than objective and empirical, then you are beginning to catch my drift, but there is nevertheless a way, if not to draw the line, then to make the nonpropositional construal more and more improbable:

“Fire” uttered in isolation is, by accepted social convention, a shorthand way of uttering the proposition (say) “There is a fire in here.” It is true that one might have been enumerating the words that rhyme with “dire,” but that is unlikely, and if it is unlikely for a monosyllabic proposition like “fire,” it is still more unlikely for the longhand version “There is a fire in here.” As the utterance becomes more complex, it becomes more far-fetched to construe it in any other way than propositionally.

So with complex propositions, we are in a performance domain that is radically different from praxis and pantomime, for propositions DO have truth value. Moreover, they seem to have the power to express any truth: This hypothesis, in the form of the “effability” hypothesis, to the effect that anything that is the case can be described in words, was put forward by Jerry Katz (1976) in the NY Academy of Sciences Conference that helped re-open the language origins question in our century. In the same volume, an independent variant of the effability hypothesis was proposed by Steklis and Harnad (1976) in the form of the “translatability” criterion, to the effect that all natural languages are fully intertranslatable. That turns out to be logically equivalent to the effability hypothesis, but it may be the more suggestive version for our purposes here, because it highlights the “cryptographic” aspects of language's expressive power: There are of course a limitless number of propositions, but every language has at least one way of expressing them all: What are the chances of coming out with just the right string of symbols, but without intending that proposition? They quickly approach the chances of chimpanzees typing a passage from Shakespeare (Harnad 1996a).

So here we are faced with a profound divide: One one side, is the world of praxis, with its objects, events and states of affairs. It has some limited resources for aping itself: I can try to draw or imitate a tree in the wind or perhaps even a rainbow over a horizon. But my praxic repertoire quickly runs out of resources in relation to all the objects and events and states there are in the world: How to mime that “all men are mortal,” or that “a continuous function is everywhere differentiable?” Those are states of affairs that can only be described in words. Moreover, as I suggested before, even pictures and mime are not DESCRIPTIONS: they are merely other states of affairs that happen to have some similarity to whatever they are pictures of. So even to construe pantomime as anything other than “being there” – like a poem, that should not “mean” but “be” – is already to construe it propositionally.

And here is where the analogy with the demotion of consciousness comes in: For case by case, practically, or rather praxically speaking, every instance of praxis and pantomime could be acted on instrumentally: The person who covers his mouth and points could be a correlate and hence a predictor, like the yellow spot on the gorilla's forehead, of the presence of someone else in the room. I need not have recourse to a proposition for that; hence a propositional construal need not be posited for the pantomime itself. Where does the proposition need to kick in? And in what does its kicking in consist? What can we do with language that we can't do with praxis and pantomime?

Candidates immediately come to mind: It's hard to mime things that one does not have the equipment to depict. It's hard to mime in the dark. In their presence, we can point to objects we'd have trouble mimicking, but in their absence? It's hard to mime either/or relations, or conditional relations – hard to mime relations themselves, or features or properties. The more abstract something is, the harder it is to mime it, because miming is concrete and particular.

Can we even mime KINDS, as opposed to specific instances? We've conceded that the proposition “All men are mortal” might defy miming, but could we even mime “mortal”? Sure, we can show someone or something dying; and then maybe show another, entirely different thing dying, and hope that in providing this panorama of concrete instances, the abstract category will somehow be picked out. But how would we MARK that abstract category that we had laboriously acted out? And how would we carry it into the more complex proposition “I am mortal,” much less “All men are mortal.” Be careful to distinguish “I am going to die,” which is relatively easy to mime, from “I am mortal,” which is not.

So we need to be able to MARK abstract categories, such as “mortal.” At the very least, one of the concrete depictions of some dying creature would have to do double duty for THAT dying creature, and for mortality in general. Now notice that for its concrete role of depicting a particular dying creature, the resemblance between the depiction and the object depicted is the kind of nonarbitrary, analog relation that psychologists have called “iconic.” Saussure stressed the arbitrariness of linguistic signs, in contrast to this. Why? I'm not sure whether Saussure intuited the property that Jorge-Luis Borges (1969) singled out in his “Funes the Memorious.”

Funes was a man who once fell off a horse and after that time he could never forget anything. He had an infinite rote memory for every concrete particular he ever encountered. His memory was so good that he gave all the integers unique proper names – Fred, Jeff, Charlie – all the way up into the hundreds of thousands till he got bored. Yet he had the greatest difficulty understanding why the rest of us, ordinary mortals with frail memories, insisted on calling (what we referred to as) that dog “Fido” in that particular position at that particular instant by the same name as what we insisted was the same dog, “Fido” at another instant, in another position. For to Funes, these were all infinitely unique and different experiences. His memory faithfully mimed and saved them all. What it couldn't do was forget or ignore any of it. Hence it could not abstract. Hence he couldn't mark all those instances of “Fido” with the arbitrary sign “Fido.” They were all infinitely different and unique to him. So of course if Borges had been completely consistent, he could not have portrayed Funes as speaking at all, for to speak he would have had to have gained a command of those arbitrary names for abstract categories that would have required forgetting or ignoring all the differences that are preserved in a faithful copy. Instead, all he had was the nonarbitrary icons, each unique to its specific instance (Harnad 1987).

Now of course it's not just a speaking Funes that is impossible; even a nonverbal Funes could not survive for a day if he could not abstract. The abstraction would not have to be marked by an arbitrary sign; it could be marked by a nonarbitrary praxic response such as sitting only on those things that afforded sitability-upon, and so on. A repertoire of evolutionarily prepared as well as learned feature detectors that subserved praxis would serve creatures nearly as complex and capable as ourselves quite adequately. Where does the added power of the arbitrary sign and the proposition kick in? What can species NOT do by praxis and pantomime alone?

My main objective here was to suggest that this is the real question at the heart of the problem of the origin of language. The origin of language is the origin of marking categories with arbitrary “signs” (symbols) and stringing those symbols together into descriptive propositions that far outstrip the possibilities of praxis and pantomime. What would be the survival value, the adaptive advantage, of propositions over mere praxis and pantomime?

In a series of artificial life simulations (Cangelosi & Harnad 2000; Cangelosi, Greco & Harnad 2000) we have tried to show that this advantage can be thought of as the advantage of (symbolic/propositional) “theft” over honest (sensorimotor) “toil.” I will close with a sketch of how this would work. “Honest toil” is good old trial-and-error operant learning guided by feedback from the consequences of one's behaviour, as in learning to distinguish edible mushrooms from toadstools: An organism samples mushrooms, tastes them, see whether it gets sick or gets nourished by them, and eventually, if the category is learnable, learns to tell apart the ones that afford nourishment from the ones that are toxic. Those are categories an organism has earned by honest toil.

Note that the foregoing is just a description, not an explanation. An explanation requires a causal mechanism for HOW the organism managed to learn to tell apart edible and toxic mushrooms by honest toil: how its brain managed to find the critical features that reliably distinguish the shadows cast on its senses by mushrooms from those cast by toadstools.

Neural nets are one natural candidate for such a feature-detecting, category-learning mechanism (Tijsseling & Harnad 1997). Mushroom-sorting (Cangelosi & Paris 1988) is of course not a realistic paradigm for category learning; it is just a “toy” problem. (For one thing, the timing is unrealistic: If telling apart mushrooms and toadstools were hard, then how could a creature in a mushroom world afford to sample them by trial error long enough to learn which kinds are which without starving itself to death?)

But contrast this with the inborn internal feature detectors of the frog, who already "knows" what kind of thing to flick his tongue out at from birth, or rather, from the time of metamorphosing from a tadpole into an air-breathing frog. Let us say that the frog has come by his bug-detectors not by honest toil, as in the case of the hypothetical mushroom-detectors, but by Darwinian theft: He was born with already prepared detectors; he got them "for free." Of course this too leaves out the critical part, for nothing comes for free. If the frog did not perform the honest toil, involving variation and selection on the basis of the consequences of trial and error, then the "Blind Watchmaker" (evolution by natural selection) must have done it for him.

But it is not Darwinian theft that I meant when I spoke of the virtues of theft over honest toil. To understand what form of theft I had in mind, we have to go back to the mushroom world: Suppose the bleaker scenario I mentioned were the actual one: Suppose there was not enough time in the day to sample toadstools and mushrooms until you had them safely sorted out: If you had to rely on honest toil alone, you stood a good chance of starving to death or perhaps getting poisoned. But suppose there were others of your kind who already had the detectors (by some means or other): If they could just DESCRIBE to you in words the features of the safe and unsafe mushrooms, perhaps supplemented with some pointing to examples, they could save you an awful lot of honest toil. (Biederman & Shiffrar [1987] have provided evidence that a verbal description of the winning features, together with some good examples, can fast-forward a novice to 90% of grandmaster performance level in newborn chicken-sexing, a level that normally requires months of honest toil at the feet of black-belted masters.)

Now notice that in a realistic scenario "theft" is a misnomer here, for, all else being equal, symbolic theft (hearsay) is a victimless crime. If you know something I don't, you are not in general any the poorer for telling me about it and saving me the time and trouble of learning it the hard way. Of course, we have managed to put a price tag on everything, and perhaps in is only in our contemporary information society that this kind of "gift" (as opposed to theft) is becoming the COMMERCE it was always destined to be, but gift, barter, theft or commerce, it is clear that it is language that has conferred on us the power of bypassing countless hours of honest toil.

Cangelosi & Parisi (1988) tried to show the adaptive advantages of symbolic theft in artificial life simulations. They put the theft/toil strategies into competition: A population of virtual mushroom-foragers (back-propagation nets) learned to distinguish edible from inedible the hard way, through trial-and-error sensorimotor toil, supervised by feedback from errors; another population learned it the easy way, by overhearing the toilers vocalize "edible" and "inedible." The thieves did not have to learn the features, because the toilers had already done it for them. And there were plenty of mushrooms, so vocalizing freely did not deprive anyone of anything. Within a few generations the thieves were out-surviving and out-reproducing the toilers. The theft strategy defeated the toil strategy, and demonstrated the adaptive advantage of language.

Or did it? Would such a strategy be "evolutionarily stable," or would it, like certain forms of cheating and parasitism, eventually play itself out? For consider that the theft strategy only works while there are toilers in the know within earshot; without the guidance of hearsay, the thief is lost, not having learned the critical features. So if anything, a competition of this sort, continued across generations, could at best induce an oscillation, with thieves at an advantage over toilers while there are plenty of toilers about, vocalizing their hard-won knowledge, but as the toilers' numbers shrink in favour of the thieves, the thieves become increasingly clueless and their own numbers accordingly shrink in favour of the toilers.

Such an oscillation is evolutionarily possible, but there is certainly no evidence of it today: We are all thieves. How did this come to pass? Theft is parasitic on toil: it is "ungrounded" (Harnad 1987, 1990) unless there are toilers about too. So perhaps we all do the groundwork, acquiring certain "basic" categories the old way, by direct sensorimotor toil, and then the rest can be acquired the new way, by the Biederman/Shiffrar strategy of symbolic description, stringing together the grounded names of the basic categories into propositions describing higher-order categories by Boolean combinations (Harnad 1996a). This is indeed what our mushroom-simulations have shown (Cangelosi & Harnad 2000): Everyone learned the ground-level categories by toil, but for higher-order categories, the theft strategy beats the toil strategy, and it is evolutionarily stable.

This is still a toy simulation, however; it remains to be seen whether the this model for the adaptive advantage of language will scale up to lifesize ecological settings.

References

Andrews, J., Livingston, K. & Harnad, S. (1998). "Categorical Perception Effects Induced by Category Learning". *Journal of Experimental Psychology: Learning, Memory, and Cognition* 24(3) 732-753.

Biederman, I. & Shiffrar, M. M. (1987) "Sexing day-old chicks: A case study and expert systems analysis of a difficult perceptual-learning task". *Journal of Experimental Psychology: Learning, Memory, & Cognition* 13: 640 - 645.

Borges, J. L. (1969) "Funes the memorious". In Yates, D. & Irby, J. E. (eds.) *Labyrinths*. New York: New Directions.

Byrne, R.W. & Russon, A.E. (1998). "Learning by imitation: A hierarchical approach". *Behavioral and Brain Sciences*, 21 (5): 667-684.

Cangelosi, A. & Harnad, S. (2000) "The Adaptive Advantage of Symbolic Theft Over Sensorimotor Toil: Grounding Language in Perceptual Categories". *Evolution of Communication* (Special Issue on Grounding) http://www.cogsci.soton.ac.uk/~harnad/Papers/Harnad/harnad00.language.html

Cangelosi A., Greco A. & Harnad S. (2000). "From robotic toil to symbolic theft: Grounding transfer from entry-level to higher-level categories". *Connection Science* (in press) http://www.cogsci.soton.ac.uk/~harnad/Papers/Harnad/cangelosi-connsci2.ps

Cangelosi A. and D. Parisi. (1998) "The emergence of a 'language' in an evolving population of neural networks". *Connection Science*, 10:83-97.

Catania, A.C. & Harnad, S. (eds.) (1988) *The Selection of Behavior. The Operant Behaviorism of BF Skinner: Comments and Consequences*. New York: Cambridge University Press.

Chomsky, N. (1972) *Language and mind*. Harcourt, Brace, and World.

Gallup, G. G. (1970) "Chimpanzees: self-recognition". *Science* 167:86-87

Greenfield, Patricia M. (1992) "Language, tools, and brain: The ontogeny and phylogeny of hierarchically organized sequential behavior". *Behavioral and Brain Sciences* 14:531- 95.

Harnad, S. (1987) "The induction and representation of categories". In: Harnad, S. (ed.) (1987) *Categorical Perception: The Groundwork of Cognition*. New York: Cambridge University Press. http://www.cogsci.soton.ac.uk/~harnad/Papers/Harnad/harnad87.categorization.html

Harnad, S. (1990) "The Symbol Grounding Problem". *Physica D* 42: 335-346. http://www.cogsci.soton.ac.uk/~harnad/Papers/Harnad/harnad90.sgproblem.html

Harnad, S. (1991) “Post-Gutenberg Galaxy: The Fourth Revolution in the Means of Production of Knowledge”. *Public-Access Computer Systems Review* 2 (1): 39 - 53 http://www.cogsci.soton.ac.uk/~harnad/Papers/Harnad/harnad91.postgutenberg.html

Harnad, S. (1996a) “The Origin of Words: A Psychophysical Hypothesis.” In Velichkovsky B & Rumbaugh, D. (Eds.) *Communicating Meaning: Evolution and Development of Language*. NJ: Erlbaum: pp 27-44. http://www.cogsci.soton.ac.uk/~harnad/Papers/Harnad/harnad96.word.origin.html

Harnad, S. (1996b) “Experimental Analysis of Naming Behavior Cannot Explain Naming Capacity.” *Journal of the Experimental Analysis of Behavior* 65: 262-264. http://www.cogsci.soton.ac.uk/~harnad/Papers/Harnad/harnad96.naming.html

Harnad, S. (2000) “Turing Indistinguishability and the Blind Watchmaker”. In: Fetzer, J. & Mulhauser, G. (ed.) *Evolving Consciousness* Amsterdam: John Benjamins (in press) http://www.cogsci.soton.ac.uk/~harnad/Papers/Harnad/harnad98.turing.evol.html

Holloway, Ralph L. (1970) “Neural parameters, hunting and the evolution of the human brain.” In: *Advances in primatology, Volume I: The primate brain*, ed. Charles R. Noback and William Montagna. Appleton-Century-Crofts.

Isaac, Barbara (1987) “Throwing and human evolution”. *The African Archaeological Review* 5:3-17.

Katz, J.J. (1976) “The Effability Hypothesis”. In: Harnad, S., Steklis, H. D. & Lancaster, J. B. (eds.) (1976) “Origins and Evolution of Language and Speech”. *Annals of the New York Academy of Sciences* 280.

Mesulam, M.M. (1998) “From sensation to cognition”. *Brain* 121: 1013-1052

Pinker, & Bloom, P. (1990). “Natural language and natural selection”. *Behavioral and Brain Sciences* 13 (4): 707-784.

Rolls, E.T. (1997) “Consciousness in neural networks?” *Neural Networks* 10: 1227-1240

Rolls, E.T. (2000) “Precis of *Brain and Emotion*.” *Behavioral and Brain Sciences* 23 (2): 177-234

Steklis, H.D. & Harnad, S. (1976) “From hand to mouth: Some critical stages in the evolution of language”. In: Harnad, S., Steklis, H. D. & Lancaster, J. B. (eds.) (1976) “Origins and Evolution of Language and Speech.” *Annals of the New York Academy of Sciences* 280, 445-455. http://www.cogsci.soton.ac.uk/~harnad/Papers/Harnad/harnad76.hand.html

Tijsseling, A. & Harnad, S. (1997) “Warping Similarity Space in Category Learning by Backprop Nets”. In: Ramscar, M., Hahn, U., Cambouropolos, E. & Pain, H. (Eds.) *Proceedings of SimCat 1997: Interdisciplinary Workshop on Similarity and Categorization*. Department of Artificial Intelligence, Edinburgh University: 263 - 269. http://www.cogsci.soton.ac.uk/~harnad/Papers/Harnad/tijsseling97.cpnets.html

Whiten, A. & Byrne, R. W. (1988) “Tactical deception in primates”, *Behavioral and Brain Sciences* 11:233-273.

Wilkins, W.K. & Wakefield, J. (1995). “Brain evolution and neurolinguistic preconditions”. *Behavioral and Brain Sciences* 18 (1): 161-226.

How can the study of limb preference be used to investigate the evolutionary emergence of language?

Rebecca HARRISON & Pia NYSTROM

Department of Archaeology and Prehistory, University of Sheffield, U.K.
r.m.harrison@sheffield.ac.uk p.nystrom@sheffield.ac.uk

ABSTRACT

Throughout primate evolution the brain has increased in both size and specialisation in order to accommodate an increasingly complex range of behaviours. Control of many of these behaviours has become represented asymmetrically in the cortex (Kaas 1987), for example, language capabilities and motor control of the dominant hand. It has been suggested that the evolution of language is related to increasingly lateralised use of the limbs due to the strong tendency for control of these behaviours to be located in the same hemisphere of the brain (e.g. Deacon 1992a; Gibson 1993). In this paper we examine the evidence which may support the perception that there is a relationship between the evolution of limb preference and language and whether it is really justified to draw a correlation between these two behaviours. We will also assess the potential benefits to the study of the origins of language if such a correlation exists.

Over time the language areas of the brain, notably Broca's and Wernicke's areas, have become increasingly lateralised, with the left hemisphere being dominant in the control of language functions in the majority of modern humans. Even though Broca's area is now considered to consist of several overlapping regions involving different capabilities (Aitchison 1998), it is still believed to perform important language functions in modern humans (Deacon 1992b). The planum temporale is an important component of Wernicke's language area for both the production and comprehension of spoken and gestural human languages. This structure, which is predominantly more expanded in the left hemisphere, also exhibits a correlation with handedness, notably a differential activation in listening tasks according to handedness (Tzourio *et al*. 1998). There is evidence for asymmetry of the planum temporale in fossil endocasts of *Australopithecus*, *Homo habilis*, *H. erectus* and *H. sapiens neandertalensis* (Holloway 1980; LeMay 1976; Tobias 1987). Apes, particularly chimpanzees, show a human-like asymmetry in the language areas, including the planum temporale, and in areas associated with cognition (Gannon and Kheck 1999; Gannon *et al*. 1998).

Thus, it appears that the neural framework for "chimpanzee language" (Gannon *et al*. 1998) may have been in place in the common ancestor of chimpanzees and humans, i.e. long before the appearance of fully developed human language. However, Corballis (1991) argues that there is a fundamental discontinuity between humans and other species, which relates primarily to a generative assembling device "responsible for constructing representations in generative fashion from small vocabularies of primitive units" (Corballis 1991: 219). In contrast, Aitchison (1998) states that language is composed of many constituents, each of which displays a different degree of continuity. The evolution of a complex adaptive trait, such as language, must be considered to have been a gradual process. Therefore, chimpanzee vocalisation may be considered as a precursor to the highly specialised form of vocalisation of modern humans, and could provide

insight into the early development of human language. It is possible that the same speculation can be made of handedness, with manual dexterity becoming increasingly complex over time.

The lateralisation of motor control of the dominant hand has been the subject of much investigation, with a special focus on nonhuman primate limb preference. Results from meta-analyses suggest that preferences exist at an individual rather than species-level (e.g. McGrew and Marchant 1996, 1997), and that individual preferences get increasingly stronger with certain tasks in higher grades of primates. Even though right hand preference in modern human populations appears to be cross-cultural, there is evidence that the strength of this preference varies both according to the task being performed and at an individual level (Healey *et al.* 1986; Marchant *et al.* 1995; Peters and Murphy 1992). Unfortunately, the focus of many studies on human handedness inventories and the performance of fine motor tasks such as writing does not facilitate an easy comparison with nonhuman primates.

However limb preference is not unique to *H. s. sapiens*. Preferences are found in other primate species, with higher grades of primates showing increasing levels of individual preferences. Even though modern humans also show variability in limb preference, in addition to a low but constant level of variability in the neural control of dominant limb and language (Ramussen and Milner 1977), it can be questioned if there is an advantage for modern humans to have both handedness and language represented in the same hemisphere (Corballis 1991). Many explanations can be offered for this widespread, though not exclusive, association. Possibly handedness and language control are located in the same hemisphere by random chance and there is no connection between them. Alternatively there may be a highly developed functional correlation which necessitates them being located in the same hemisphere of the brain. Asymmetries also occur in the areas of the cortex which relate to motor control of speech, the periSylvian cortex (Galaburda 1984), with the Sylvian fissure being longer in the right hemisphere, particularly in right handed individuals (Bradshaw 1980; LeMay and Culebras 1972). Thus, the motor control of language production might be related to the need to have a dominant hand capable of performing highly specialised tasks. A possible impetus behind such a specialisation is the manufacture and use of tools which, at its extreme levels, requires fine motor control of both the dominant hand in the manufacture of tools and of vocal communication in the transmission of information regarding the processes of production.

The evolution of grammatical structures

Bernd Heine

Center for Advanced Study in the Behavioral Sciences
75 Alta Road
Stanford, CA 94305-8090
bernd.heine@casbs.stanford.edu

ABSTRACT

There are a number of approaches that are available to the linguist to study earlier phases in the evolution of human language or languages. The present paper aims at discussing the potential that grammaticalization theory offers to the reconstruction of language evolution. Findings of this paradigm have been used in previous works to study language genesis (e.g. Sankoff 1979; Aitchison 1996), using in particular findings from pidgin and creole languages.

Grammaticalization theory relies on regularities in the evolution of linguistic forms, especially on the unidirectionality principle and the implications it has for the recustruction of earlier language states (Heine, Claudi & Hünnemeyer 1991; Hopper & Traugott 1993). The purpose of the paper is to show that there are certain classes of grammatical forms that can be assumed to presuppose other grammatical forms in time. In the course of the paper, an attempt is made to reconstruct sequences of grammatical evolution with a view establishing how language may have been structured at earlier stages of human evolution.

References

Aitchison, Jean. 1996. The seeds of speech: Language origin and evolution. Cambridge: Cambridge University Press.

Bybee, Joan L., Revere D. Perkins & William Pagliuca. 1994. The evolution of grammar: Tense, aspect, and modality in the languages of the world. Chicago: University of Chicago Press.

Comrie, Bernard. 1992. Before complexity. In Hawkins & Gell-Mann (eds.). 193-211.

Hawkins, John A. & Murray Gell-Mann (eds.). 1992. The evolution of human language. Proceedings of the Workshop on the Evolution of Human Languages, held August, 1989 in Santa Fe, New Mexico. Santa Fe: Addison-Wesley Publishing Company.

Heine, Bernd. 1997. Cognitive foundations of grammar. Oxford, New York: Oxford University Press.

Heine, Bernd, Ulrike Claudi & Friederike Hünnemeyer 1991. Grammaticalization: a conceptual framework. Chicago: University of Chicago Press.

Hill, Kenneth C. (ed.). 1979. The genesis of language. (The First Michigan Colloquium, 1979.) Ann Arbor: Karoma Publishers.

Hopper, Paul J. & Elizabeth C. Traugott. Grammaticalization. Cambridge: Cambridge University Press.

Sankoff, Gillian. 1979. The genesis of a language. In Hill (ed.) 1979. 23-47.

The time scale of language evolution

Jean-Marie HOMBERT & Charles LI

Dynamique du Langage *University of California*
Lyon *Santa Barbara*

ABSTRACT

If we plot the increase in language capacities as a function of time over the last 4 million years, several theoretical curve shapes are possible :

- a straight line representing a slow linear evolution of language through this long period of several million years
- a step-like representation illustrating a series of punctuated equilibrium
- a curve progressing very slowly for a long period of time followed by a sharp increase after an inflexion point relatively recent in time.

In this paper, we would like to summarize the existing evidence to argue for the third scenario. We will consider three phases :

Phase 1: “Slow evolution” period (from 4 My to 100 Ky)

Phase 2 : “Inflexion point” period (around 100 Ky)

Phase 3 : “Fast evolution” period (the last 100 Ky)

We will propose that the functions of our communicative system and the mechanisms at work to modify this system through time are different in phase 1 vs phases 2 and 3. In phase 1, our “language” is used to communicate our emotions, to maintain social ties and progressively to give information about the “here and now”. In phase 3, language is used to *transfer knowledge :* This new function will have an extremely strong snowball effect for the survival and development of our specie since social, technical, and geographical knowledge will not have to be “relearned” by each generation, it will be transmitted by articulated language. Knowledge will become *cumulative*. In order for phase 3 to be possible, it required a much more sophisticated linguistic system than the one used during phase 1. Drastic changes occurred in phase 2 : development of “double-articulation” (grouping of meaningless sounds to produce meaningful words- ie emergence of lexicon) and use of word sequencing (emergence of syntax). Because of the snowball effect mentioned, it is conceivable that this phase was relatively short (a few tens of thousands of years).

We will argue that if darwinian type of processes can account for early periods of language evolution (in order to create a more “efficient” communication system), recent evolution of linguistic systems are not explainable in these terms and are better accounted for by self organization processes.

Finally, we will suggest that a possible explanation for the disappearance of Neandertals could be that, unlike Homo sapiens, they did not get to a stage of “knowledge transfer”.

The roles of communication and representation in language evolution

James R HURFORD

Language Evolution and Computation Research Unit
Department of Theoretical and Applied Linguistics
University of Edinburgh
jim@ling.ed.ac.uk

ABSTRACT

``We should search for the ancestry of language not in prior systems of animal communication but in prior representational systems." (Bickerton, Derek, 1990, *Language and Species,* University of Chicago Press, Chicago. p.23.)

It can be agreed that nothing of the distinctively complex structure of modern languages can be attributed to ancestry in animal communication systems. But how much of the complex structure of modern languages can be attributed to ancestry in pre-linguistic representational systems? The essential differences between an internal (cognitive) representation system and a communication system are as follows.

A communication system maps external forms (such as speech sounds or manual signs), via mental structures, to meanings (where many, if not all, meanings relate to external objects, events or situations). A communication system is typically public, shared by many individuals{But this is not crucial, as for example the last living speaker of a dying language can still be said to possess a communication system..

A representation system lacks the mapping to external forms, and merely provides mental structures mapped to meanings. It is assumed that the meanings dealt with by any such representation system relate to, or denote, external situations. There would be no practical advantage in having a representation system which was not in some way related to the world outside the mind possessing it.

Thus a communication system properly includes a representation system. There are elements in a communication system that are not part of the inherent representation system. Any aspects of a communication system which pertain only to the mapping between external forms (sounds, signs) and the internal cognitive representation system are not part of the representation system per se.

Languages are very complex, highly structured communication systems. The question arises how much of the structure of a language is only part of the communicative aspects of the system, that is, how much of it does not belong to the inherent representation system. The view that linguistic structure derives from representation systems existing prior to language can only be sustained to the extent that there is no structure that is only part of the communicative aspect of a language system.

So how much of language structure is purely representational, and how much of it is part of the mapping to external forms? One cannot quantify such questions, but the answer is that almost all of the complex structure of languages belongs to their communicative aspect, and

very little to their purely representational aspect. This paper will survey the range of linguistic structures that relate only to the mapping between internal representations and external forms.

The fundamental universal structural characteristic known as ``duality of patterning'', whereby languages are organized at two levels of structure, namely phonology and morphosyntax, has no motivation in a purely representational system, but plausible arguments can be advanced for its communicative adaptiveness.

Obviously, all of phonological structure belongs in the communicative aspect of linguistic structure. All of the following aspects of structure play purely communicative roles, and no role in non-communicative representation: Phonemic structure, tonemic structure, phonotactics (which organizes the universal classes of vowels and consonants in patterns such as the sonority hierarchy), phonological processes of assimilation and dissimilation (including vowel harmony and consonant harmony), other allophonic processes, syllabic and moraic structure, intonation and rhythm. The phonological component of a language comprises a very significant proportion of its structure.

On the morphosyntactic side of the duality of patterning, the universal distinction between morphology and syntax (however that is drawn) plays no role in non-communicative representation. This distinction rests on the discrimination by languages of a level of words, which are small-to-middle-sized units distinct from both semantically atomic morphemes and higher level syntactic units such as phrases. Within morphology, various structural features, such as the layering of inflectional morphemes outside derivational morphemes, and the inventory of structural devices used in word-formation (affixation, suppletion, fusion, cliticization, reduplication, compounding) also play no purely representational role. These are processes affecting the external forms of languages.

Within syntax, many of the complex structural phenomena that have attracted study, such as case marking, anaphor-antecedent relationships, switch-reference devices, control by verbal predicates of the interpretation of their complement clauses, transformations of various sorts (e.g. passivization, topicalization, question formation) and the constraints on such processes, play no role in non-communicative representation. Linear ordering of elements, with which much of syntax is concerned, likewise plays no non-communicative role. Also fundamental to syntactic structure are lexical classes somewhat autonomous of semantics, such as Noun, Verb, Adjective and Preposition; to the extent that such classes are autonomous, they play no role in semantic representation. Other commonly found grammatico-lexical categories, such as grammatical gender (Noun classes), would seem to serve no representational purpose, although they may contribute to the redundancy of utterances, thereby serving a communicative purpose. Grammatical agreement (concord), which is widespread, also clearly plays no purely representational role.

Some aspects of linguistic structure may indeed plausibly be derived from non-linguistic, representational, structure. These include some (but not all) aspects of hierarchical organization in syntax. But the broad conclusion from the above survey of non-representational aspects of linguistic structure is that attempts to derive linguistic structure, in an evolutionary account, from previously existing cognitive representational structure must fail, for a large slice of linguistic structure. Correspondingly, we can seek evolutionary explanations (broadly conceived) for much (though not all) of the typical structure of languages in the demands of communication in the human environment.

Modeling discourse complexity

Ichiro IGARI & Takashi IKEGAMI

Institute of Physics, The Graduate School of Arts and Sciences
University of Tokyo
3-8-1 Komaba, Meguro-ku, Tokyo 153-8902
{igari, ikeg}@sacral.c.u-tokyo.ac.jp

ABSTRACT

We here try to unify the studies of discourse analysis and syntax analysis. There is a tradition started by Chomsky to take syntax and semantics to be independent of each other and to analyze syntax as a closed formal system. On the other hand, discourse analysis has focused more on the dynamic and open-ended aspects of a language system, such as entrainment and bifurcation of context flow during conversation. The unification of the two subjects is certainly required now (Steels, 1998).

Our approach towards the unified theory is based on dynamical systems theories and simulations. We propose coupled dynamical recognizers as a candidate model for studying discourse complexity theoretically. A dynamical recognizer is a class of recurrent neural network, which is capable of mimicking some formal language systems (Pollack, 1995). But at the same time, the dynamical recognizer has a rich structure as a novel dynamical system. It is now widely used as a basis for robot navigations (Tani and Fukumura 1994) and natural language processing (Elman 1995). A characteristic feature of the dynamical recognizer is not its computational ability but rather its dynamic nature of perception and manipulation of given information data sets. Due to a kind of chaotic attractor that exists in dynamical recognizers, we can formalize a language system not as a rigid formal system but as an autonomous evolving system. Coupling those dynamical recognizers adds a new level of complexity (Ikegami and Taiji 1998 1999, Taiji and Ikegami 1999). Our perspective has a close connection with recent developments of cognitive linguistics by Langacker (1987, 1991) and Lakoff (1987). In particular, Langacker's maximalism and his way of taking a growing structure of abstraction/extension network as “syntax” has stimulated our approach.

In the present model, as an initial setup, each dynamical recognizer is trained separately by a given set of words to learn certain syntax behind. A single dynamical recognizer is known to learn a given set of words not syntactically but “semantically”. To put it another way, words are not categorized in an alphabetical way or by living/non-living criteria. But they are learned as a set of elements that constitute the given context. For example, a set of words like cats, cheese and mouse have similar internal state patterns.

Then we study the conversation situation where two agents try to speak to each other by predicting what the other speaker expects to hear in the next turn by the given context. The prediction is based on its own dynamical recognizer. Namely, in each conversation step, agents update their dynamical recognizer structures to mimic the previous behaviors of the other agent. Since what one agent expects will perturb the other agent's dynamical recognizer structure mutually and indirectly, their dynamical recognizer tends to change their structures in time. We argue that all the complexity of the discourse pattern will be generated by this mutually predicting dynamics.

Our preliminary result shows that two dynamical recognizers are developed to have different syntactic structures from each other. In this sense, the mechanism which agent “learns” dynamically from another agent is not a mere entrainment. We call the new learning dynamics interactive learning. In some cases dynamical recognizers do not converge on static structures. This is consistent with our picture of language in which a language is an evolving system in its own right.

References

[1] Elman, J., Language as a Dynamical System. In R.F. Port and T. van Gelder(Eds.) Mind as Motion. MIT Press, 1995

[2] Ikegami, T., Taiji, M. Structures of Possible Worlds in a Game of Players with Internal Models. Acta Poly. Scan. Ma. 91. 283-292, 1998

[3] Ikegami, T., and Taiji, M., Imitation and Cooperation in Coupled Dynamical Recognizers In Advances in Artificial Life (Eds. D.Floreano et al. Springer, 1999) pp. 545-554

[4] Lakoff, G., Women, Fire, and Dangerous Things: What Categories Reveal About the Mind. The University of Chicago Press, 1987

[5] Langacker, R.W., Foundations of Cognitive Grammar, i, Theoretical

Prerequisites. Stanford University Press, 1987

[6] Langacker, R.W., Foundations of Cognitive Grammar, ii. Stanford

University Press, 1991

[7] Pollack, J.B., The Induction of Dynamical Recognizers. In R.F. Port and T. van Gelder(Eds.) Mind as Motion. MIT Press, 1995

[8] Steels, L., The origin of linguistic categories. In “The Evolution of Language” (Selected papers from 2nd International Conference on the Evolution of Language, London), 1998

[9] Tani, J. and Fukumura, N, Learning Goal-Directed Sensory-Based Navigation of a Mobile Robot. Neural Networks, vol.7, no.3, pp.553-563, 1994

[10] Taiji, M. and Ikegami, T., Dynamics of Internal Models in Game Players. Physica D 143 (1999) pp.253-266.

The control of affective expression: The origins of symbolic reference

Christine M. JOHNSON

Department of Cognitive Science
University of California San Diego
johnson@cogsci.ucsd.edu

Language is a highly complex process with a variety of interacting components. This paper will focus on one such component, the control of affective expression, examining its role in language development, and its possible evolutionary origins based on a comparative analysis of its rudiments in nonhuman primates.

The impact of this component is most starkly evident in very early language development when the infant, who has been babbling for some months, begins to use its first words (~1-1.5 years old). During this period, there are two apparent sources for these utterances, whose operation, I will argue, are inter-related. One is the (relatively faithful) imitation of externally generated sounds, such as the vocalizations of others. Much has been written about this sort of mimicry (e.g. Meltzoff 1988; Donald 1993) which clearly plays a vital role in the development of symbolic reference. For the purpose of this paper, however, I will focus more on the other source – the imitation, or modified repetition, of internally generated sounds. These may include "inventions" of the infant's articulatory system or utterances incorporated through the above-mentioned external imitation process. In either case, it is the manner in which such repetitions are modified that reveals the operation of the affect control mechanism.

These modifications may be best understood by approaching early language development from a Vygotskian perspective (e.g. Vygotsky, 1978). In this view, the child's early repertoire of utterances derive their communicative function from evoking the full, rich context of their original use. So, for example, "ball", when used by the 1 yr old, does not refer to just one particular (or class of) round object(s), but to the whole set of social and physical activities and surroundings that typically co-occur in the child's experience of that object. From this perspective, the utterance "ball" is just one of those activities, one that can be reproduced by the child even when some or most of the other aspects of the relevant situation are absent. It is by comparing such utterances with their original form that particular modifications become evident.

One modification involves a "flattening" of affect, where an utterance that was used with some emotional inflection in its original context, is now repeated without that emotional content. This can occur immediately after its affective use, while much of the originally evoking context is still in effect, or much later, when few, if any, elements, are present. In either case, the flattening of affect can, in itself, be seen as "detaching" the utterance, to some degree, in the sense of reducing the physical properties (e.g. inflection, loudness) it has in common with the original context. This detachment – both temporal and affective – also weakens associations with the reinforcement contingencies of the original context. That is, it fosters a separation, a de-contextualization, from the immediate consequences of that situation. This, in turn can facilitate, among other things, additional self control, by somewhat reducing the motivation for immediate reaction and providing time for a more studied response.

The other evidence of affect control at this early stage involves using a particular utterance – generally one of the common phonemes well-practiced during babbling, such as "da" or "ba" – in a variety of contexts, varying its inflection while holding its phonemic content constant. Different inflections tend to evoke different responses from more sophisticated interlocutors, and this provides the child with a basis for developing their strategic use. Learned patterns – such as, in English, a rising frequency sweep for the interrogative, a rise to a plateau and then fall for concurrence, etc. – can then also be applied to an array of phonemes. That is, the reproduction of such words with the context-appropriate inflection, in the absence of the original model, show that both their phonemic and affective content are eventually subject to the infant's control.

Both of these processes – the flattening of inflection and its context-dependent manipulation – demonstrate the child's burgeoning ability to control its affective expression. Both also play a role, I would argue, in the development of symbolic reference. That is, these affect control mechanisms help the child 1) to produce utterances that are disconnected from the particulars of the context in which they were learned, while still retaining some capacity to refer to (evoke) that context and 2) to apply them appropriately across a variety of contexts. Both are critical to an utterance's function as a symbol.

Given the very early stage of language development at which this affect control appears, and how much of subsequent language development (e.g. the growth of the lexicon, the combinatorics of syntax, the conceptual sophistication of narrative and metaphor) depend upon it, it would seem reasonable to expect that this capacity would likewise have arisen early in the evolution of language. This being the case, we might further expect that an assessment of communication in our nearest primate relatives could reveal rudiments of this fundamental ability and hints about its early adaptive functions.

The study of control in communication – and especially of the suppression of affect – in nonhuman primates presents a vexing challenge to the ethologist, however, because of the inherent ambiguity of a failure to respond. That is, an apparently passive animal may be repressing its response or it may be genuinely unmoved. Given no direct access to the mental states responsible for such behavior, the ethologist must rely on other behavioral cues to justify interpreting observations as cases of affect control versus indifference. Such cues include abrupt, radical changes in behavior, especially those linked to the attentional behavior of others, as well as apparent insusceptibility to emotional contagion, when nearby individuals are engaged in affective displays. In my research on social communication in bonobos at the San Diego Wild Animal Park (*Pan paniscus*) examples of such interactions have been captured on videotape, and will be presented here for consideration as evidence of a limited degree of control of affective expression in these animals.

The first example involves "charging displays" most often given (especially in our group) by the adult male. These dramatic displays involve pushing or dragging a branch or other large object directly past other individuals at high speed. It is a high arousal display, marked by pilo-erection, facial grimacing, and sometimes loud vocalizations. In response, younger animals give way or rush to their mothers, and at least one adult female in our group, who has a history of antagonism with the male, regularly flees with a fearful expression and sometimes screams as well. Of particular interest here, however, are the reactions of the other two adult females in the group, who typically become completely immobile and show no change in expression for the duration of the display. While the possibility – even likelihood – remains that they are genuinely less intimidated by such displays, which only rarely result in their actually being attacked, it may also be that their lack of overt response helps preclude such attacks. (Note that the female who flees is often chased.) Whatever the cause-and-effect relationship here, the "no-response" strategy has apparently been learned by these animals, the younger of whom was observed to flee such displays when she too was immature. The development of this behavior would seem to involve some capacity for the control of affective expression.

On a few occasions, the male in this group was recorded calmly walking to a distant site to pick up an object (a collapsed cardboard box) which he sometimes used in such displays. Only after carrying it to a place where he had a direct charging line to the other animals (in the dry moat that surrounds their island) did he begin his display and exhibit the pilo-erection, grimacing and vocalizations typically associated with that behavior. While it might be argued that these affective expressions were evoked by the combination of the proximity of the target animals and the armament of the displayer, his calm demeanor during the (clearly anticipatory) preceding activity is similar enough to the above described case of the human infant's flattened affect, in a somewhat removed context, to warrant comparison.

In another segment, an adult female was videotaped engaged in a (for her) highly atypical bout of somersaulting, tickling, and giggling with an adolescent female. The adult male noticed them from afar and gamboled toward them with a distinct "playface" and the wagging arms of an animal inciting play. However, as soon as he approached, the older female immediately sat up and assumed a demur, adult-like demeanor. This sudden change also brought both the other animals' play to an abrupt, although temporary, halt. Regardless of her motivation for, or conception of, this act, the female's behavior clearly demonstrated her capacity to exercise sudden control over emotional expression.

In a final segment, another of our bonobos was recorded making an aggressive but unsuccessful grab for some browse (food) held by another individual. She then sat very still, with a calm facial expression, directing her gaze away, until the second animal finally stopped monitoring her and looked away. At that point the first animal immediately turned and made another grab for the branch, this time successfully. (For other similar examples, see Goodall 1986; de Waal 1986; Whiten & Byrne 1988.)

Such behaviors in nonhuman primates are sometimes described anthropomorphically as "feigning indifference", "premeditation", "saving face" etc. Such characterizations are relevant to my topic in that I believe that the self control abilities that underlie such strategies in HUMANS are directly related to the control of affect in early language development. However, such attributions to the apes are not necessary to an evolutionary argument for the origins of these abilities. It is both easier, and more appropriate, to argue for the adaptive significance of functional behavioral traits than of the invisible "mental states" that may, or may not, underlie them (see Johnson, in press). Regardless of their construal by the animals, such interactions can provide insights into possible contexts where selective pressures for improved self control may have operated in our own prehistory.

It seems reasonable to suggest that, in the above described examples, the animals gained some advantage from so modifying their behavior. This can generally be seen in the behavior's impact on the response of others, whether in the inhibition of attack, avoiding dispersal from a premature warning that would decrease the effectiveness of a display, or reducing vigilance toward a contested (food) resource. The case of the adult female who abruptly stopped playing when the male showed up is perhaps somewhat more mysterious, but still suggests that "audience effects" were at issue.

Bonobos (like chimpanzees and, most likely, early hominids) live in fission/fusion societies in the wild, where subgroup membership changes frequently (see Kano, 1992). Such a social structure can offer selective advantages to those who can modify their behavior in such a way as to strategically vary the information they make available to their current audience. The control of affective communication would be particularly useful here, since affect-laden signals tend to be relatively conspicuous and are probably the ones most frequently used by others to predict subsequent behavior. To the extent that it pays off to effect the attention and predictions of others, and inasmuch as the ability to do so is mediated by some genetically influenced neuronal structure that enables the context-dependent control of affective expression, differences in these abilities could be translated into the differential reproduction that is critical to evolutionary change. A greater dependence on food sharing or other divisions of labor, such as have often been suggested for early hominids, might create just the sort of increased social stakes required to ratchet these abilities beyond those of the common primate ancestor.

Finally, in looking to the neuronal structures that may mediate such behavior, the most likely candidates probably lie in the prefrontal cortex. In general, the prefrontal cortex is considerably enlarged in humans relative to other primates and is implicated both in the inhibition of (for example, culturally inappropriate) behavior as well as in various planned activities, including language production. Portions of this cortical lobe (for example, what has been mapped, in the human brain, as "Area 13") are apparently involved in the direct inhibition, or modification, of lower brain centers (such as the amygdala and associated structures) of the limbic system, which play a major role in generating emotional expression. Recent comparative research has shown that this area of prefrontal cortex is generally less well developed in apes than in humans. Interestingly enough, however, bonobos DO show some additional development in this area – in terms of greater cell number and columnar differentiation – compared to the other apes (Semendeferri, et al. 1998). Reported behavioral differences in bonobos, compared to chimpanzees, in terms of the former's reduced tendency toward exaggerated affective responses (Savage-Rumbaugh 1984; de Waal 1989) is particularly interesting in light of these brain data.

Nonetheless, the above-described bonobo behaviors are relatively rare and clearly rudimentary relative to behavior that is common in even infant humans. Thus, both the brain and the behavior data confirm that our species has made significant strides in our evolution of affect control. Before any analysis of the more complex language abilities in humans can proceed, we perhaps need to acknowledge, and account for the emergence of, the fundamental role played by the control of affective expression.

References

Donald, M. (1993) Precis of "Origins of the modern mind: Three stages in the evolution of culture and cognition". *Behavioral and Brain Sciences 16*:737-791.

Goodall, J. (1986) *The chimpanzees of Gombe*. Harvard University Press, Cambridge MA.

Johnson, C.M. (In press) Distributed primate cognition. *Animal Cognition*.

Kano, T. (1992) *The last ape: Pygmy chimpanzee behavior and ecology*. Stanford University Press, Stanford CA.

Meltzoff, A.N. (1988) The human infant as *Homo imitans*. In Zentall, T.R. & Galef, J.B.G. (Eds) *Social learning: Psychological and biological perspectives*. Erlbaum, Hilldsale NJ.

Savage-Rumbaugh, E.S. (1984) *Pan paniscus* and *Pan troglodyytes*. Contrasts in preverbal communicative competence. In R.L. Sussman (Ed) *The pygmy chimpanzee*, pp: 131-177) New York: Plenum Press.

Semendeferri, K, Armstrong, E, Schleicher, A., Zilles, K & Van Hossen, G.W. (1998) Limbic frontal cortext in hominoids: A comparative study of area 13. *Am. J. Phys. Anth. 106*:129-155.

de Waal, F.B.M. (1986) Deception in the natural communication of chimpanzees. In: Mitchell, R.W. & Thompson, N.S. (Eds) *Deception: Perspectives on human and nonhuman deceit*. SUNY Press, Albany NY, pp 221-244.

de Waal, F.B.M. (1989) Behavioral contrasts between bonobo and chimpanzee. In P.G. Heltne & L.A. Marquardt (Eds) *Understanding chimpanzees*. Harvard University Press, Cambridge MA.

Whiten, A. & Byrne, R. (1988) Tactical deception in primates. *Behavioral and Brain Sciences 11*:233-244.

Vygotsky LS (1978) *Mind in society:The development of higher psychological processes*. Harvard Univeristy Press, Cambridge MA London.

Language and adaptation: How learning and laziness compete in the emergence of morphosyntax

Simon KIRBY

Language Evolution and Computation Research Unit
Department of Linguistics, University of Edinburgh
simon@ling.ed.ac.uk
http://www.ling.ed.ac.uk/~simon

ABSTRACT

Language is an adaptive system. In order for languages to persist over time, they must be transmittable from one generation of language users to another. The mechanisms of language transmission – i.e. use and acquisition – influence the relative transmittability of language variants, and therefore the emergent structure of language. An important question, obviously, is whether we can predict the structure of language given a particular hypothesis about these mechanisms. For example, it might be that we can simply read-off the properties of Universal Grammar from the structure of the language acquisition device (a view implicit in much of generative linguistics), or it may be that the link between language use and acquisition on the one hand, and the structure of syntax on the other, is rather more indirect.

In this paper, I explore these issues by examining the types of possible mappings between meaning space and signal space. Human languages typically employ mappings that are topographic. Informally, this means that structure in one space is preserved in the other. The most striking (and uniquely human) example of the topographic nature of language is recursive compositionality. Put simply, the meaning of an utterance tends to be a function of the meanings of parts of that utterance. Compositionality, however, is not a universal property of signals in human language. For example, highly frequent meanings seem to map onto signals in a rather different way from less frequent ones. They tend to be expressed irregularly (i.e. non-compositionally), and with shorter signals.

How are we to interpret these fundamental properties of the morphosyntax of human language? Should they be trivially predicted by our theory of the structure of the language faculty, suggesting that language users are directly constrained by their biology to acquire only languages structured in this way? Or does taking an adaptive systems perspective on languages make a less direct link (and more plausible language faculty) possible? Supporting the latter hypothesis, I will argue that the particular patterns of compositionality found in language are emergent adaptations by language itself to two simple properties of language users: generalisation, and avoidance of effort.

That this is the case can be demonstrated using a simple computational model of the process of language transmission (an extension of the model presented at the London evolution conference). It can be shown that, if the simulated language learners are predisposed to look for generalisations across examples, completely compositional encodings are inevitable "attractors" for evolving languages, even if initially, languages are completely non-compositional. This result, though interesting, is not entirely satisfactory, because it does not account for frequent meanings being short and irregular.

If a least-effort principle for the speakers is modeled, however, the mean length of utterance in the languages of the simulation tends to be much lower. Now, if the distribution of meanings in the simulation is skewed so that some are more frequent than others, the most frequent meanings do appear to behave differently to the others. They are shorter, and are expressed with either partial compositionality or as fully idiosyncratic strings. In other words, the combination of generalisation and least-effort on the part of learners and speakers in the simulation gives rise to meaning-signal patterns that look very natural.

This result can be explained in terms of adaptation. The mechanisms of acquisition and language-production pose adaptive problems for language to solve if it is to be faithfully transmitted from generation to generation. It turns out that the optimal language types for these two "problems" are actually different. In other words, the pressures from generalisation and least-effort are in competition in the evolution of language. Simplifying somewhat, the best language for learning is the most compositional encoding of the meaning space (as evidenced by the results from earlier simulations), whereas the best language for production is the minimal-length encoding of the meaning space. Generally, these cannot be the same.

The solution that language finds to this conflict of interests is interesting because the relative importance of the learning and production pressures changes with the frequency of meanings. Infrequent meanings must be expressed using a compositional syntax otherwise they are unlikely to survive the process of learning from a sub-sample. This is less true for frequent meanings. Conversely, the more frequent a meaning, the more it is subject to erosion from speakers, tending to make it shorter (and eventually unable to be compositional).

These results show us that the relationship between the structure of the human language faculty and the emergent structure of language can be quite indirect. In order to understand the origins of the particular features of human language, we should not simply look at the way humans have adapted to be better at learning language, but also the way language has adapted to being better at being passed on by us.

From 'nursing poke' to syntactical speech

Chris Knight

Department of Sociology and Anthropology
University of East London
Longbridge Road, Barking
Dagenham RM8 2AS, Essex, UK
c.knight@uel.ac.uk

Syntax is the hierarchical, recursive structuring of words or morphemes to form higher-order units such as phrases and sentences. While it is possible to conceive of words without syntax, the reverse possibility – syntax without words – is logically excluded. From this, it follows that in the evolution of language, words or 'protowords' must have arisen before syntax or at least concurrently with it.

Syntax cannot have preceded words. Yet it is not inconceivable that syntactical competence evolved prior to the emergence of vocal speech as we know it. One theoretical possibility is that it evolved to facilitate hierarchical, recursive structuring in some other medium such as manual gesture (Armstrong et al. 1994) or 'mimesis' (Donald 1991). In both primates and humans, gesture is a natural, intuitive medium for expressing thought (Tomasello and Call 1997; McNeill 1992); when iconic gestures become conventionalised, a repertoire of discrete, arbitrary signs is the result. Against this background, it is tempting to infer a similar dynamic at work in the early evolution of words (Steklis and Harnad 1978; Kendon 1991).

Recent authoritative accounts of language origins have denied any role for manual gesture in the early evolution of language. MacNeilage (1998: 238) adopts an extreme position, asserting that 'the vocal- auditory modality of spoken language was the first and only output mechanism for language'. This coincides with Dunbar's (1996: 141) view that gesture was never necessary -- 'it can all be done by voice'. MacNeilage's (1998) central argument is that if the vocal-auditory modality was adaptive during later stages of speech evolution, it must have been equally adaptive from the outset.

MacNeilage's argument would have force if it could be confirmed that the social contexts of language use remained invariant throughout the course of human evolution. But if changing social strategies are built into our models, there is no reason to suppose that a modality which is adaptive during an early period must remain equally adaptive at a later stage. Where social contexts are 'Machiavellian', as is the case among primates (Byrne & Whiten 1988), we know that constraints operate to obstruct the emergence of low-cost, conventional – in other words fakeable – signalling (Zahavi 1993, Zahavi & Zahavi 1997). The problem is that intentional signals are by definition cognitively controllable – yet any intentional manipulability in signals automatically undermines their intrinsic reliability. Because of the fundamental requirement for reliability, primate vocalisations are resistant to manipulation, remaining emotionally expressive, limbically governed and hence hard-to-fake. They must remain beyond intentional control as a condition of their efficacy in communication.

Precisely because the primate vocal modality has been designed by natural selection to serve a reliable, hard-to-fake system of communication, it is ill-suited for exaptation to serve novel linguistic functions. Indeed, of all aspects of primate expressive behaviour, vocalisations appear among the most constrained and hence the least qualified. They are not autonomous with respect to emotion, largely escaping cortical control. For the same reason – being unavailable for intentional 'picture- making' – they lack the iconic potential required of elementary conceptual communication (McNeill 1992). Being resistant to re-shaping or social learning, primate vocalisations lack the requisite plasticity to take on linguistic functions (Steklis and Harnad 1978: 447).

These features are not maladaptive: they reflect selection pressures intrinsic to communication in the animal world. Theorists unaware of Darwinian constraints might imagine that the evolution of cognitive complexity in primates will be matched by corresponding manipulability and complexity in communication. But this would be to overlook the requirement for reliability in signals (Ulbaek 1998; Zahavi 1993). As receivers seek to avoid costs of deception, they respond preferentially to signals which are 'hard-to-fake' (Burling 1993). Such pressures prevent primate vocal communication from coming under cognitive control.

Primate vocal communication, then, is a poor candidate for exaptation to serve linguistic functions. Much better qualified are the forelimbs, with their specialisation for prehensile and manipulative functions (Napier 1960). Manual gestures are manipulable precisely because the hands have evolved to perform technical, not communicative, functions. Complex bimanual activities such as food-extraction and tool manufacture intensify selection pressures for fine-tuned intentional co-ordination and control. Since such serial motor sequences are cognitively governed, it follows that on those rare occasions when primates need to convey details of cognition, they should rely more on gesture than on voice.

Among chimpanzees, conventionalised manual gestures occur only in highly restrictive social contexts. A well-documented case is the infant 'nursing poke' (Tomasello et al. 1994). This begins as a functional action: the infant pushes aside its mother's arm to reach the nipple. As mother and infant interact with one another over time, the poke becomes abbreviated and conventionalised. The end result is a learned, intentional, discrete shorthand, falling outside the normal species-specific repertoire of emotionally expressive, hard-to-fake gesture-calls (cf. Burling 1993). The 'nursing poke' suggests a plausible step in the direction of language.

What are the social conditions for such a step? At issue is the question of trust. A sign can become conventionalised only to the extent that the requisite trust can be assumed. Note that in acquiring its conventional form, a 'nursing poke' has no prospect of escaping the confines of the particular maternal relationship in which it has originated. Conventionalisation occurs because in this particular social context of communication, interests on both sides coincide. Infants have an interest in cutting the costs of requesting a feed, while mothers have a corresponding motive to satisfy their offspring and at the same time reduce the amount of poking endured.

Now, the question arises, where else within ape society might we find a comparable convergence of interests? Perhaps comparable social contexts do exist, but certainly they are few and far between. Ape social life and corresponding intelligence is 'Machiavellian' (Byrne and Whiten 1988). In such a social setting, volitional signalling between adults is likely to be manipulative. Whereas a nursing mother may not mind being 'manipulated' by her offspring, there are few contexts in which an adult chimpanzee has an interest in being manipulated by neighbouring conspecifics, even when kin-related. In ape society, community-wide relationships do not constitute extensions of the mother-infant bond. Where mother-infant-like trust is not extended throughout a coalition or community, associated conventions cannot be extended either. In short, apes in the wild have no use for low- cost conventional signs. This is not because they are cognitively deficient – Kanzi has disproved that idea. It is because arbitrary signs are low- cost signals and so lack intrinsic credibility (Zahavi 1993). Any successful use of such signals, whether vocal or gestural, will depend on exceptional levels of mutual trust and co-operation (Knight 1998).

If this is accepted, then the gap between nursing pokes and human language was not bridged thanks to the sudden evolution of special cognitive capacities. Rather, it was bridged thanks to the elaboration of low-cost, high-speed, convention-based strategies of communication – initially based on existing capacities for controlling temporal-sequential outputs – made possible by intensifying levels of ingroup social trust. Where sufficient trust existed, individuals would have been under pressure to develop corresponding expressive and communicative capacities. Kanzi and other trained chimps have demonstrated their ability to deploy signs and to act upon them provided there is some function to be served by doing so – in other words, provided human trainers (differing in this respect from chimp conspecifics) can be trusted to consistently reward such behaviour. In the case of evolving humans, there were of course no trainers to reward use of conventional signs. Rather, speakers and listeners had to place trust in one another to an extent which would be maladaptive under conditions of 'Machiavellian' primate politics.

Language is a low-cost system of conventional signs wholly dependent on social trust and bound up with the evolutionary development of stable kin- based coalitions (Dunbar 1996; Dessalles 1998; Nettle 1999; Power 1998; Knight 1991, 1996, 1998, 1999). Within these coalitions, status was determined by linguistic relevance (Dessalles 1998) rather than coercion, emotional manipulation or violence. Progressive elimination of emotional conflict from the sphere of ingroup relations allowed gestures to become increasingly dispassionate, low-cost, abbreviated and contrastive. Unlike costly displays, signals of this kind lend themselves naturally to sequential ordering and hierarchical, recursive structuring. There are no grounds for assuming that humans at any stage lacked the neural control capacities necessary to achieve such 'syntactical' output. Initially manual-gestural, and initially bound up with emotionally expressive 'mimesis' (Donald 1991, 1998), conventional signing dedicated to ingroup communication would have come under pressure to exploit and integrate whatever additional modalities best facilitated cost-cutting and efficiency. As ritually structured, trust-based coalitions became increasingly stable and emotionally homogenous, interlocutors became less interested in emotions or performance- based distinctions, more interested in underlying intentions and thought- processes. In such contexts, signallers came under pressure to develop shorthands, speed up transmission and string together increasingly elaborate sequences. For reasons of efficiency, this evolutionary dynamic drove progressive adoption of the vocal-auditory modality as the default medium for human language. Licensed by changed social circumstances, vocal speech exapted and developed for its own purposes the sophisticated neural control machinery originally developed to serve an earlier, more gesturally based system of communication.

References

Armstrong, D. F., W. C. Stokoe and S. E. Wilcox 1994. Signs of the origin of syntax. *Current Anthropology 35*: 349-368.

Burling, R. 1993. Primate calls, human language, and nonverbal communication. *Current Anthropology 34*: 25-53.

Byrne, R. and A. Whiten 1988. *Machiavellian Intelligence. Social expertise and the evolution of intellect in monkeys, apes, and humans*. Oxford: Clarendon Press.

Dessalles, J.-L. 1998. Altruism, status and the origin of relevance. In J. R. Hurford, M. Studdert-Kennedy and C. Knight (eds), *Approaches to the Evolution of Language. Social and cognitive bases*. Cambridge: Cambridge University Press, pp. 130-147.

Donald, M. 1991. *Origins of the Modern Mind. Three stages in the evolution of culture and cognition*. Cambridge, MA.: Harvard University Press.

Donald, M. 1998. Mimesis and the Executive Suite: missing links in language evolution. In J. R. Hurford, M. Studdert-Kennedy and C. Knight (eds), *Approaches to the Evolution of Language. Social and cognitive bases*. Cambridge: Cambridge University Press, pp. 44-67.

Dunbar, R. I. M. 1996. *Grooming, Gossip and the Evolution of Language*. London: Faber and Faber.

Kendon, A. 1991. Some considerations for a theory of language origins. *Man* (N.S.) *26*: 199-221.

Knight, C. 1991. *Blood Relations. Menstruation and the origins of culture*. New Haven & London: Yale University Press.

Knight, C. 1998. Ritual/speech coevolution: a solution to the problem of deception. In J. R. Hurford, M. Studdert-Kennedy & C. Knight (eds), *Approaches to the Evolution of Language: Social and Cognitive Bases*. Cambridge, Cambridge University Press, pp. 68-91.

Knight, C. 1999. Sex and language as pretend-play. In R. Dunbar, C. Knight & C. Power (eds), *The Evolution of Culture*. Edinburgh: Edinburgh University Press, pp. 228-229.

McNeil, D. 1992. *Hand and Mind. What gestures reveal about thought*. Chicago & London: University of Chicago Press.

MacNeilage, P. 1998. Evolution of the mechanism of language output: comparative neurobiology of vocal and manual communication. In J. R. Hurford, M. Studdert-Kennedy and C. Knight (eds), *Approaches to the Evolution of Language*. Cambridge: Cambridge University Press, pp. 222-41.

Napier, J. R. 1960. Studies of the hands of living primates. *Proceedings of the Zoological Society of London 134*: 647-57.

Nettle, D. 1999. Linguistic diversity and the evolution of societies. In: R. I. M. Dunbar, C. D. Knight and C. Power (eds), *The Evolution of Culture*. Edinburgh: Edinburgh University Press.

Power, C. 1998. Old wives' tales: the gossip hypothesis and the reliability of cheap signals. In J. R. Hurford, M. Studdert-Kennedy and C. Knight (eds), *Approaches to the Evolution of Language: Social and Cognitive Bases*. Cambridge: Cambridge University Press, pp. 111-29.

Steklis, H. D. & S. Harnad 1976. From hand to mouth: Some critical stages in the evolution of language. In S. Harnad, H. D. Steklis & J. B. Lancaster (eds), *Origins and Evolution of language and Speech*. Annals of the New York Academy of Sciences 280: 445- 455.

Tomasello, M., J. Call, K. Nagell, R. Olguin and M. Carpenter 1994. The learning and use of gestural signals by young chimpanzees: a trans- generational study. *Primates 35*: 137- 54.

Tomasello, M. & J. Call 1997. *Primate Cognition*. New York & Oxford: Oxford University Press.

Ulbaek, I. 1998. The origin of language and cognition. In J. R. Hurford, M. Studdert- Kennedy and C. Knight (eds), *Approaches to the Evolution of Language*. Cambridge: Cambridge University Press, pp. 30-43.

Zahavi, A. 1993. The fallacy of conventional signalling. *Philosophical Transactions of the Royal Society of London 340*: 227-230.

Zahavi, A. & A. Zahavi 1997. *The Handicap Principle. A missing piece in Darwin's puzzle*. New York & Oxford: Oxford University Press.

A modified-neutral theory for the evolution of linguistic diversity

Daniel LIVINGSTONE

Applied Computational Intelligence Research Unit
Dept. CIS, University of Paisley, PA1 2BE, UK
daniel.livingstone@paisley.ac.uk

Introduction

With the application of evolutionary theory to cultural evolution, it has become somewhat more common to view language change as the result of evolutionary processes on languages themselves. Such a viewpoint typically ignores the role played by individual and social motivations in creating or progressing language changes, but does not necessarily deny that such motivations have some influence (Lass 1997).

We ask if it is possible that the cultural evolution of languages can be adaptively neutral – that language change does not necessarily result in functionally better languages or provide adaptive benefit to the communities using them. We argue against theories that hold that such benefits are required for linguistic diversity and dialects to emerge.

Opposing views are held by, amongst others, Lass (1997) and Nettle (1999). Lass argues against functional explanations of language change at some length. In contrast Nettle argues that, without any social benefits, the neutral evolution of language would be unable to produce significant linguistic diversity or language change. Nettle also builds a computational model to support his case.

We review Nettle's argument and model, highlighting weaknesses of both, and counter with further arguments to support the neutral evolution of languages. We show that significant and sustained change and diversity can exist without apparent means of selection. Like Nettle, we support our argument with a computational model. In the model, diversity emerges and is maintained over time, while the abstract 'languages' continually change and evolve – despite the model lacking any means of selecting for diversity. A brief review is also given of other models which support the notion of adaptively neutral linguistic evolution.

Finally, we consider whether neutral evolution excludes other explanations or causes of language change, and discuss the relationship between these different causes.

Neutral theories of linguistic evolution

Both Lass and Nettle draw on the work of Kimura (1983), who showed that it was possible for evolution to occur without any apparent selective forces at work. However, they give very different explanations of why the neutral-evolution of language should, or should not, be sufficient to cause change and diversity in human languages.

Lass' proposal is based on the observation that languages are imperfectly replicating systems, within which elements of linguistic 'junk' and other 'marginal' features exist. This provides ample room for variation, and allows changes to occur without disrupting the success of communication. That replication is not, and can not be, perfect means that languages will change, regardless of functional benefits.

Nettle argues against the neutral-evolution of linguistic systems on three points:

1. Random changes would be non-directional and could be expected to cancel each other out, due to an *averaging effect*.

2. With a neutral model it is difficult to account for diversification without geographical isolation.

3. Structural correlations in many of the world's languages represents parallel evolution, showing that the path of linguistic diversification is not random.

Thus, Nettle proposes that in order for linguistic evolution to occur without geographical isolation additional mechanisms are required. Nettle argues that the social functions of language are *required* for the emergence of linguistic diversity, a view shared by Dunbar:

> "… dialects arose as an attempt to control the depredations of those who would exploit peoples natural cooperativeness", Dunbar (1996, page 169).

Nettle's third point does not relate to this contention, and is outwith the scope of this current work. It does not represent an unsurpassable problem, however. For example, linguistic diffusion seems to indicate that significant randomness exists in language change, despite the apparent regularity of the results. Constraints, innate or otherwise, also impact upon linguistic evolution and may account for many structural correlations (Kirby 1999).

Nettle's first and second points both rely on the equal distribution of individuals, with a uniform likelihood of any one individual interacting with any other. As recognised by Cavalli-Sforza and Feldman (1978), in any group the amount of influence exerted on any one individual by any one of the others will vary according to a number of factors. This reduces the effect of averaging, and increases the potential for sub-populations to vary from the mean. The different social networks within groups reduces the need for geographical isolation to produce linguistic diversity. The importance of social networks as an influence on language change is emphasised by Milroy and Milroy (1993).

Further, the averaging effect itself is questionable. For example, for random variation in the formant frequencies of phonemes it may not be the case that such variation will 'cancel out', or that the average values will be learned. Phonemic and articulatory constraints (see Lindblom, 1998, and de Boer, 1999) may prevent the 'cancelling out', and the learned phonemes may be only *tolerably* close to those heard. With discrete forms, given two different forms of a linguistic feature a learner does not choose just one to learn, but learns both. One may be preferred, but both may be used in varying amounts. As well as applying to the lexicon, it has been proposed that language learners learn multiple grammars, so as to cope with the variation in grammars in use around them (Kroch, 1989).

Computational models of linguistic diversity

We now review and criticise the computational model built by Nettle. Then we describe our model, and review others which add strength to the arguments for the neutral evolution of linguistic diversity.

First, in an earlier paper, Nettle and Dunbar (1997) built a model which demonstrates that language diversity can operate as a linguistic marker that serves an important social function. In the case of the model, by excluding non-cooperative individuals from participating in social exchanges. Nettle concludes that its use as a social marker explains the reason *why* linguistic diversity exists, but this model is merely a demonstration that linguistic diversity *can* serve a social function. That the utility of dialect diversity is what leads to its emergence is not proven.

The model in Nettle (1999, Chapter 3) attempts to demonstrate how neutral evolution of language cannot lead to diversity, using a computational model in which learners learn phonemes from adults in the same group. There are, however, some questionable characteristics of the model. Language learners in this model learn the average formant frequency values used by the adults in the population, with slight perturbation due to noise. All children in a group learn simultaneously from a single 'snapshot' of the adult phonological systems. This learning process explicitly limits the possible variation within a group, and the phonetic systems evolve very slowly.

The model tries to show that inter-group diversity cannot emerge through random change alone. A number of 'family' groups exist in the model, and where there is no interaction between groups, inter-group diversity emerges. When individuals migrate between groups, given the explicit averaging and the exceedingly slow rate of change, it is not too surprising that the migration destroys inter-group diversity. Nettle proceeds to show that if learners only learn from selected role-models – a form of social selection – inter-group diversity will again emerge. We shall see, however, that such selection is not required to demonstrate the evolution of linguistic diversity.

In an attempt to show the effect that the geographical or social organisation of language users can have on linguistic diversity, we previously developed a model in which the language learners are spatially distributed along a line (Livingstone and Fyfe 1999). The artificial neural network based agents are implemented similarly to those in Livingstone and Fyfe (2000), but are homogenous in structure. Learners are taught only by individuals in the parent generation within a local neighbourhood.

From random initial conditions, the resultant languages are such that considerable diversity exists globally, but local clusters of agents learn similar languages with low diversity. A dialect continuum is observed in the model – all individuals are capable of successful communication with others in the near vicinity. The likelihood of successful communication decreases with the distance between any pair of communicating of agents. As new generations learn, there is continued fluctuation and change in the learned languages. Thus, without geographical isolation, or any means of functional or social selection, diversity and change are still observed.

One weakness of this model is that starting from non-random initial conditions, where all agents in the first generation have the same language, no changes to the linguistic system take place and diversity does not emerge. However, by adding a small amount of noise to the model it is seen that diversity does emerge and the results are soon indistinguishable from those of random starting conditions. We use noise to represent errors in linguistic transmission and random variations in an individuals own use of language, factors where it is difficult to determine a realistic figure, but believe the value used (below 1%) to be reasonable, perhaps even conservative.

Other computational models, investigating language change in spatially homogenous populations, have shown similar results. In Steels and Kaplan (1998) random variation, caused by stochastic errors introduced into language games between pairs of agents, is able to create competition between linguistic forms in the agent languages, resulting in change over time. Stoness and Dircks (1999) present a similar model, one which does not rely on random noise to maintain competition between forms.

The assertion that some social function is necessary for linguistic diversity and language change to occur is not supported by these models. Change and diversity is caused simply by the (imperfect) replication of language over many interactions between individuals.

A modified-neutral theory of linguistic evolution

We have argued that no functional or adaptive benefits are required to create linguistic diversity, and that diversity should arise naturally from the imperfect transmission of language from users to learners. This represents a neutral theory of linguistic evolution.

We have reviewed some of the objections to a neutral theory, and shown them to be inconclusive. Accordingly, social or linguistic functions are seen to be unnecessary for the emergence of diversity. What then is the role of social and personal motivation in language change? To say that adaptive benefits are not required for the evolution of diversity is not to say that such benefits do not exist, or that they do not influence the evolution of languages. Indeed, classic studies such as that of language change in Martha's Vineyard (Labov, 1972) show that social factors do exert a strong influence.

Accepting that language changes are influenced by social pressures on language users, we can question why language users adapt their language according to such pressures. Is there something remarkable in the human ability to determine significant social information simply from accent and dialect, without regard to the content of the speech?

Rather than claim that it is the usefulness of such abilities that led to the evolution of linguistic diversity, we argue that the reverse is more likely – that the existing linguistic diversity may have led to the development of such abilities. Humans have had many millennia to adapt to living in societies with linguistic diversity present. That people are able to modify language according to situation means that linguistic change is unlikely to be truly neutral, as social pressures influence which changes will succeed in a population. This does not negate the neutral evolution of language, it adds to it.

We conclude that the neutral evolution of languages is unavoidable and remains a factor in language change, but is not the only cause of change. Yet neither social nor linguistic function are required to create linguistic diversity – geographical spread and imperfect transmission alone are sufficient to account for this.

References

de Boer, B. (2000). Emergence of sound systems through self-organisation. The Evolutionary Emergence of Language. C. Knight, M. Studdert-Kennedy and J. R. Hurford, Eds., Cambridge University Press.

Cavalli-Sforza, L. L. and M. W. Feldman (1978). Towards a Theory of Cultural Evolution. Interdisciplinary Review 3: 99-107.

Dunbar, R. (1996). Grooming, gossip and the evolution of language, Faber & Faber.

Kirby, S. (1999). Function, Selection, and Innateness: The emergence of language universals. Oxford University Press.

Kroch, A. (1989). Reflexes of grammar in patterns of language change. Journal of Language Variation and Change 1: 199-244.

Labov, W. (1972). Sociolinguistic Patterns. Philadelphia, University of Pennsylvania Press.

Lass, R. (1997). Historical Linguistics and Language Change. Cambridge University Press.

Lindblom, B. (1998). Systemic constraints and adaptive change in the formation of sound structure. Approaches to the Evolution of Language. J. R. Hurford, M. Studdert-Kennedy and C. Knight, Eds., Cambridge University Press: 242-264.

Livingstone, D. and C. Fyfe (2000). Modelling Language-Physiology Coevolution. The Evolutionary Emergence of Language. C. Knight, M. Studdert-Kennedy and J. R. Hurford, Eds., Cambridge University Press.

Livingstone, D. and C. Fyfe (1999). Dialect in Learned Communication. AISB '99 Symposium on Imitation in Animals and Artefacts, Edinburgh, The Society for the Study of Artificial Intelligence and Simulation of Behaviour.

Milroy, J. and L. Milroy (1993). Mechanisms of change in urban dialects: the role of class, social network and gender. International Journal of Applied Linguistics 3(1).

Nettle, D. and R. I. M. Dunbar (1997). Social Markers and the Evolution of Reciprocal Exchange. Current Anthropology 38(1): 93-99.

Nettle, D. (1999). Linguistic Diversity, Oxford University Press.

Steels, L. and F. Kaplan (1998). Stochasticity as a Source of Innovation in Language Games. Artificial Life VI, C. Adami, R. Belew, H. Kitano, C. Taylor., Eds., MIT Press.

Stoness, S. and C. Dircks, (1999) Investigating Language Change: A Multi-Agent Neural-Network Based Simulation, Hahn, M. and Stoness, S., Eds., Proceedings of the Twenty-First Annual Meeting of the Cognitive Science Society, Vancouver, BC.

Saussure and catastrophe: A formal investigation into the evolution of word-combination

Chris LONG

University of Durham

ABSTRACT

We offer independent, mathematical, support for Bickerton's thesis that in the evolution of language its functor (or predicate) /argument structure (henceforth: “f/a structure”) emerged catastrophically (all in one go) rather than by piecemeal adaptation. Using Saussurean principles, we outline a formal proof that f/a structure is an emergent, necessary and unitary feature in the growth of any system of meaningful combinations of symbols. We also mention how the formal character of combinatorial evolution partially reinforces Bickerton's view that f/a gave expression to the thematic structure of co-operative skills which must already have been in existence.

The proof is part of a wider mathematico-philosophical study of how language, as a combinatorial system, might be understood as having evolved through *structural growth*, somewhat like a biological system (The inspiration came originally from Piaget, D'Arcy Thompson, and Rene Thom. But its formulation draws on Saussure, categorial logic, Kuratowski and upon Wittgenstein and Heidegger.)

Combination is treated here not as mere concatenation or juxtaposition but, following Saussure, as a richer activity which provides a mode of word *differentiation*, transcending phonology and generating categories. (In another register: as the construction of “ordered”, i.e. well-differentiated, pairs.) We identify a simple procedure for initiating and recursively expanding combination, and trace the growing organisation it yields in a formal model. This study reveals that starting from scratch, without prior linguistic abilities, there is only one kind of way in which a system of meaningful word combinations *could* grow (though this has many alternative varieties): namely, towards a structure familiar as that of natural language. Like biological growth, combinatorial evolution follows a distinctive structural envelope, or geometry, but one determined by its internal organisational requirements, which are semantically motivated, not by prior structures or genes.

This paper is confined to showing how f/a patterning emerges quite rapidly, at a second recursion, in the metamorphosis from proto-language to a semantically powerful system. A concrete illustration is supplied.

Our investigation also uncovers *organisational* reasons why language must be social. Any technique of combinatorial differentiation can be exerted in alternative, mutually exclusive, ways. It is *arbitrary* which is employed, and such an arbitrary selection can only be established by agreement in practice (Lewisian “conventions”). Consequently, language, a system of myriad, cumulative, densely nested, differentiations, could only be assembled, like a dance, among participants already skilled in the recognition and negotiation of co-operative practices. That is, in communities of just the kind which Bickerton, for separate reasons, envisages.

Two approaches to the evolution of speech: Neodarwinism and generative phonology

Peter F. MacNeilage & Barbara L. Davis

The University of Texas at Austin

ABSTRACT

The “Frame/Content” theory is a Neo-Darwinian account of the evolution of speech by means of descent with modification from pre-speech capacities. According to this theory, the Frame provided by mandibular oscillation (paired with phonation) results in the mouth close-open alternation associated with the syllable. This capacity may have evolved first, and then the capacity to program Frames, with internal segmental “Content” elements may have evolved later. Beyond mandibular oscillation itself, the presence of three patterns common to infants’ first words, words of modern languages, and a proposed proto-language corpus is taken as evidence of the structure of original word forms. (1) The tendency of labial consonants to co-occur with central vowels (e.g. “ba”) is taken as evidence of an evolutionary role of “Pure Frames” produced by mandibular oscillation alone; (2) The tendency for coronal consonants to co-occur with front vowels (e.g. “di”) and dorsal consonants to co-occur with back vowels (e.g. “go”) is evidence for a lingual inertial effect produced by a static non-resting tongue position throughout an utterance. (3) The favoring of a labial consonant - vowel - coronal consonant (LC) sequence in first words of infants (e.g. “bado” for “bottle”) is considered to be the result of a self organizational increase in serial output complexity made both by earlier hominids and infants in response to the demand to increase the size of the lexical message set. The first two patterns are considered characteristic of the initial “Frame” stage and the third pattern may initiate/have initiated the Frame/Content stage in infants and in early hominids.

Generative Phonology is suggested by many to provide an alternative approach to the evolution of speech. However, there is no evidence for the most basic claim from this perspective - namely that a genetic specification for the phonological component of speech evolved de novo as part of universal grammar. In addition, no evidence is available from the biological sciences to suggest that the genes contain sufficient information to play an underlying role in the production of skilled actions (such as speech, which occurs at the rate of 15 segments per second). As in other essentialist conceptions, no time domain is specified (either historically or in contemporary acquisition of speech production) by generative phonology. Consequently generative phonology has no ability to cope with the evident fact that speech has evolved from simpler initial forms to present forms historically, and develops from simpler forms to more complex forms in infants. In addition, there is presently nothing approaching a comprehensive set of abstract distinctive features - the basic elements of generative phonology - though it is often asserted that the set of features is finite. There is good reason to be skeptical that cross-language continua in values of even the most common phonetic parameters such as those associated with consonant place of articulation and the vowel space could ever be susceptible to discrete parcelling into features. In addition while the core of phonology is considered to be abstract, the currently favored set of distinctive features is defined in concrete (articulatory) terms, which results in difficulties in characterizing aspects of sound patterns that are perceptually motivated. Moreover, to the extent that such features are considered to “account for” or “explain” the existence of speech phenomena, they presumably do so in terms of phonetic substance, and therefore cannot be taken as evidence of an independent abstract level

of phonological form with explanatory value. The concept of markedness, considered to give a unitary “explanation” of the relative prominence of certain sounds and patterns, is defined in terms of a set of properties that are not always correlated (frequency of occurrence, complexity, time of acquisition by infants). Explanatory claims are based on circular reasoning; e.g. frequency reflects markedness, therefore markedness explains frequency. One specific problem for a supposedly unitary concept of markedness is that universal and therefore unmarked phenomena are sometimes different in infants and languages; e.g. a predominance of inter-syllabic repetition in infants versus a relative prohibition of inter-syllabic repetition in languages (the Obligatory Contour Principle). Specific differences between the Neo-Darwinian and Generative Phonological approaches will be illustrated by comparing their treatments of the basic consonant-vowel alternation of speech, the three intra-syllabic CV co-occurrence patterns and the labial-coronal sequence.

Reference

MacNeilage, P.F. and Davis, B.L. On the Origin of Internal Structure of Word Forms. Science (in press).

Grammaticalization and evolution of the grammatical structure

Christiane MARCHELLO-NIZIA

ENS Fontenay/St-Cloud , Institut Universitaire de France
UMR CNRS-ENS 8503 « Analyses de corpus linguistiques»
marchell@ccr.jussieu.fr

ABSTRACT

The term “grammaticalization” is generally used to designate changes which occur at the level of concrete units of language: morphemes or syntactic constructions. The goal is to see how a new unit enters into a given language's grammar within a pre-existent paradigm, how a noun becomes a preposition, or a verb a modal auxiliary, or again, how a certain word order has become fixed.

This term does have another definition, introduced at the start by A. Meillet (1912), and well illustrated in particular by J. Bybee (1984) as well as by J. Bybee and W. Pagliuca (1994). It deals with the appearance of new paradigms to mark distinctions that grammar had not expressed until then : as such, this research concerns itself for example with the progressive “grammaticalization” of modal, or temporal, distinctions.

In one case as in the other, the question *why* a new phenomenon appears isn't explicitly posed. We can, indeed, ask ourselves if the facts of linguistic change are aleatory and dispersed,or if, on the contrary, they , at least in some instances, motivate and relate to each other. In short, is it possible to trace a causality behind these processes - whatever the level might be?

We will defend the following hypothesis : that certain changes at least, although they appear unmotivated, are in fact governed by the emergence or development of distinctions which are situated at the level of grammatical semantics.

This kind of path of development was hypothesized on the basis of the analysis of a very important change which occurred in French language : that is to say, the radical transformation which affected the demonstrative system between the 13th and the 16th century, including in particular the appearance of a new paradigm (*ce / ces*), and the grammaticalization of the adverbs “*ci*” and “*là*” into suffixes (“*ceux-ci*” and “*ceux-là*”).

We proposed arguments for explaining that change by the extension of a syntactic hierarchised distinction between pronouns and determiners : the old forms were pronouns as well as determiners, the new forms are only determiners (*ce / ces*) or only pronouns (*celui-ci...*). A new distinction can indeed generate new units in the old paradigms which it contributes to destabilizing.

In order to test this hypothesis, we examined the changes which occurred at the same period in the grammar. All of them in the nominal system, lead to generalize the same morpho-syntactic distinction between two levels : such as the apparition of the determiner *chaque* alongside *chacun* (which till then was pronoun and determiner), of *quelqu'un* alongside *quelque*, etc..

Furthermore, phenomena of change which seem to have no link with these may in fact be linked to the same general movement.

This is the case with the replacement in French of the quantifying / intensifying adverb *moult* (which could modify all kind of words), by two terms, *très* which is used only with adjectives and adverbs, and *beaucoup* which modifies the noun and the verb. This phenomenon is to be related to the generalization of the distinction we mentioned above, between two morpho-syntaxtical levels.

A last point to conclude our analysis. As far as we can see, all these changes seem to be "unidirectional" - what is one of the most important charasteristics of grammaticalization.

Thus we propose, lastly, to specify the concept of "grammaticalization" by distinguishing three levels of analysis, which will enable us to create a place for the type of process at issue here. The term "grammaticalization" can designate either the dynamic process by which a new unit enters into grammar, or the constitution of a new paradigm "grammaticalizing" a distinction expressed until then by the lexicon. It may also designate a change situated at the level of the grammatical system, and integrating into it a distinction of abstract, purely grammatical order - as in the case studied here the distinction between the level of nouns, verbs and pronouns (heads), and that of determiners, adjectives and adverbs ; and the referent of this distinction is *in* the grammar itself.

The advantage of this conception is double.

On the one hand it enables us to explain why certain forms, otherwise perfectly "viable"- as in the case of *moult* -, suddenly disappear at a specific point.

But also and most of all, it permits us to shed light on fundamental movements of grammatical systems, and to see certain regularities that govern their evolution on a very long span of time : in this case, the emergence and unidirectional development of grammatical - and cognitive – distinction.

Evaluating the influence of language contact on lexical changes

Egidio MARSICO, Christophe COUPÉ,
François PELLEGRINO

Laboratoire Dynamique Du Langage – Université Lumière Lyon 2
Egidio.Marsico@ish-lyon.cnrs.fr, ccoupe@ens-lyon.fr,
Francois.Pellegrino@univ-lyon2.fr

Introduction

This paper focuses on modelization of language contact using language games in a community of spatially distributed agents.

According to Thomason and Kaufman (1988, p. 72), the intensity of contact between two populations has an impact on the amount of borrowing. They define this intensity has a function of three parameters, (1) duration of contact, (2) relative population sizes and (3) social settings.

In this paper, we will present results of experiments testing these parameters. We developped a special toolkit for our simulations, LEMMingS (Language Evolution Modeling & MonitorING System). This model is inspired from Babel$^{©}$ (Sony Computer Science Laboratory) in the way agents interact by playing naming games, (Steels and McIntyre, 1998).

LEMMinS overview

In the LEMMingS environment, we generate several populations of different sizes. Each population is socially stratified and has a special "prestige" value (this notion subsumed cultural prestige as well as military or economical). Aside from belonging to a special social class, an agent is provided with a "communicativity" value which allows it "to choose" his social network of interactions.

LEMMingS also provides population with a specific demographic rate, and allows drawing maps of agent, word and concept spreading. We can measure the coherence of the lexicon per population, the percentage of foreign words of an agent lexicon and the average number of words for a given concept.

Experiments and results

We ran several simulations of contact between two populations. It is a fact that for a given equal prestige (resp. a given equal size of population) the biggest (resp. the most prestigious) population will impose its lexicon to the other. The most interesting point to look at is when these parameters work in opposite direction.

We will show in our presentation the extent to which prestige and size may affect foreign lexicon acquisition both in terms of number of foreign words acquired and of speed of acquisition.

Examples

We present below two simulations illustrating acquistion of foreign lexicon.

Both figures show the percentage of Foreign Known words (FK) and Foreign Preferred words (FP) of a population. FK are borrowed words that compete with the native words an agent (and thus a population) has for a given meaning. FP are borrowed words that replace the native words.

Prestige range from 0 (strong) to 9 (weak). In both simulations, the number of meaning is equal to 10 and they are shared by the two populations.

Case study 1

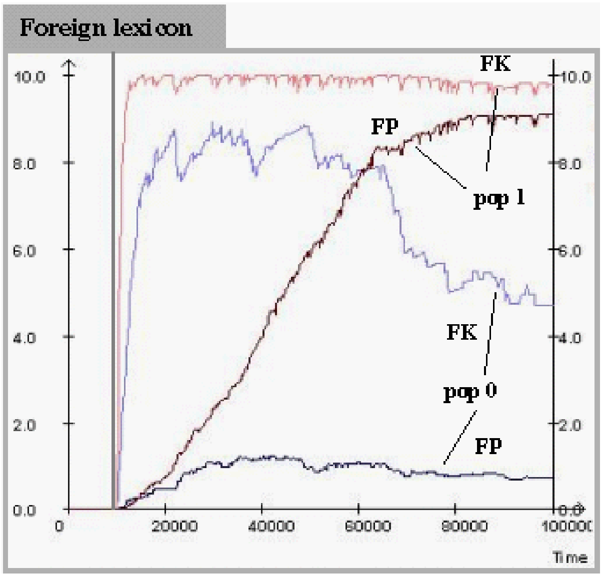
Population 0 : size = 20 ; prestige = 0

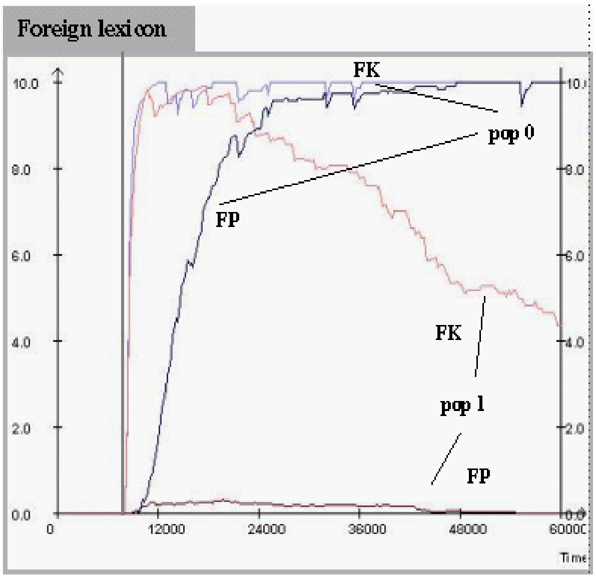
Population 1 : size = 40 ; prestige = 9

Case study 2

Population 0 : size = 20 ; prestige = 0

Population 1 : size = 40 ; prestige = 4





In this case, we can see that the smallest population (pop 0) which has the highest prestige imposes its lexicon on population 1.

Right after the contact (grey vertical line) both populations start knowing foreign words, but the main difference lies in the fact that pop 1 soon acquires preferred words (i.e. words that agents use at first place) whereas pop 0 don't.

Moreover, due to the renewal of agents, pop 0 starts losing its foreign known words as they did not manage to become preferred words (FK-pop 0 decreases). In the same time, pop 1 is renewing its lexicon with almost all the foreign known words becoming preferred.

Here, prestige balances the effect of size.

This graph shows a case where the difference in prestige is too small to counterbalance the effect of size, thus leading the smallest population (pop 0) to change its entire lexicon for the one of the biggest population.

What is really interesting here is the speed at which pop 1 imposes its lexicon.

In case study 1, prestige and size act like competing forces, thus it takes more time to replace a lexicon, whereas in case study 2, the smaller difference in prestige lets size operate alone and faster.

The evolution of language: A systemic functional exploration of phylogenetic phases

Christian M.I.M. Matthiessen

Department of Linguistics, Macquarie University
cmatthie@ling.mq.edu.au

While speculations about the origin and evolution of language were considered outside the realm of scientific enquiry in linguistics for many years, there have in recent years been a growing number of valuable contributions to this topic from a range of scholars, including now also linguists (e.g. Beaken, 1996; Deacon, 1997; W. Foley, 1996: Ch. 2; Hawkins & Gell-Mann, 1992; Hurford, Studdert-Kennedy & Knight, 1998; Jablonski & Aiello, 1998; Noble & Davidson, 1996; Ruhlen, 1994).

There is still a certain tendency to see the beginning of the evolution of language as *dependent* on conditions created by the evolution of the brain (to cope with general intelligence) and the vocal organs (resulting from bipedalism, a change in foraging behaviour, and so on), thus suggesting that language is a secondary phenomenon – perhaps even a property unique to a human "language faculty" that arose because of some mutation, as suggested by e.g. Bickerton (1990), thus creating a sharp divide between humans and other primates or indeed other animals in general. Such a line of interpretation seems to be made more plausible when there is a lack of a developmental model that could be used to explain an evolutionary sequence leading up to language: language comes to be seen only by reference to the full complexity of modern human languages (though Bickerton himself has a notion of proto-language).

Against this kind of view, a number of scholars have argued that language was a much more central factor in the evolution of humans (and as part of that, in the evolution of the human brain): see e.g. Deacon (1992, 1997), Edelman (1992) and Bickerton (1995: Ch. 2). According to this line of interpretation, language and the human brain *co-evolved*. This is certainly the view that makes sense from the point of view of a systemic functional understanding of language and the evolution of language: see Halliday (1995). However, the accounts offered in the literature can be further supported by a general theory of language that will allow us to interpret the co-evolution of language with the brain and other aspects of modern humans as a gradual but accelerating increase in complexity. I believe Halliday's (1975) account of the ontogenesis of language from protolanguage via a transitional phrase to (adult) language provides us with a very plausible model for how language might have evolved. Significantly, it does not force us to assume a sharp divide between (modern) humans and other primates – or indeed other mammals: Halliday's account of ontogenesis provides a very detailed model of the ontogenetic phases; it combines both continuity and transformation and it interprets language development as children actively learning how to mean – developing a meaning potential together with the people they interact with in a growing range of critical social contexts (cf. Trevarthen's, 1987, discussion of intersubjectivity in ontogenesis) – rather than as children "acquiring" a ready-made product.

In this paper, I will adopt a "cosmogenetic" perspective (cf. Layzer, 1990; Delsemme, 1998; Smith & Szathmáry, 1999) on the evolution of language. Language will be assumed to have evolved "emergently" as a higher-order semiotic system within an ordered hierarchy of systems of increasing complexity – physical, biological, social and semiotic systems (see Halliday & Matthiessen, 1999: Ch. 13). The evolution of language will be assumed to have been a gradual but accelerating process following general principles of increasing complexity in complex adaptive systems in general (cf. Steels, 1998) and building on earlier capacities shared with other primates (following the general evolutionary principle of new functions out of old structures); and the evolution of languages is further assumed to have been a process characterized by co-evolution with both biological systems (bipedalism, vocal apparatus, brain etc.; see Deacon, 1992, 1997; Edelman, 1992; Halliday, 1995) and social systems (division of labour, social hierarchy etc.). Through this co-evolution, there has been a gradual increase in complexity in all the systems involved as greater complexity in one creates the conditions for greater complexity in another. (I thus do not accept the arguments put forward against "gradualism" by e.g. Noble & Davidson, 1996.) In other words, the evolution of language has been an integral part of the evolution of "humanity". Importantly, the evolution of modern language has provided a new order of organization that has made human evolution possible without further brain expansion after the appearance of modern Homo sapiens around 100 K years ago.

In reviewing the evidence cited in the evolution of the human line – from the splitting off of Australopithecus, via Homo habilis, Homo erectus, archaic Homo sapiens to modern Homo sapiens, I will sort it according to the different orders of system – physical (changes in climate; changes in habitat etc.), biological (bipedalism; brain evolution; evolution of the vocal organs; state of newborn and prolonged infant dependency; etc.), social (sociality; prolonged infant dependency and social learning; recognition of individuals and group evolution; division of labour; geographical spread; settlement etc.) and semiotic (the appearance of art; evidence of burial; body adornment; but also implications from "deep" language history [cf. Nichols, 1992; Ruhlen, 1994] etc.) – while noting that the only direct evidence that remains exists at the order of physical systems (excluding backward projections based on DNA, language history etc.).

The notion of a gradual but eventually accelerating evolution of complexity fits well with Mithen's (1996) account of the evolution of the mind, with major transitions such as the evolution of a "general intelligence" in addition to previously separate specialized intelligences (cf. the discussion below of the transformation of linguistic macro-functions into simultaneous modes of meaning). However, the account that I will explore here is language-based, with the evolution of the meaning potential as the central motif (cf. Halliday & Matthiessen, 1999). The model for the phased evolution of complexity in the meaning potential is based on Halliday's (e.g. 1975) investigation of ontogenesis, expanded by subsequent research by Painter (1984), Torr (1997) and others. According to this model, it is possible to identify three non-discrete phases of development after a period when infant engage with semiotic instances and precursors (such as exchanges of attention – what Catherine Bateson has called proto-conversation) that have not yet crystallized into a semiotic system –

> *phase I: protolanguage [primary semiotic]* (bistratal [content/ expression]; and microfunctional [where function = use and is directly tied to context: regulatory, interactional, instrumental, personal at the onset]);
>
> *phase II: transition* (from bistratal to tristratal: the emergence of lexicogrammar (the system of wording: "morpho-syntax" and lexis) as a new stratum within content and of a phonological system out of the vocal "postures" of protolanguage; from an axial/stratal system to one with axis and stratification as separate dimensions of organization; and macrofunctional [where the microfunctions are generalized into two initially alternative but ultimately simultaneous macrofunctions: mathetic and pragmatic]); and
>
> *phase III: language [higher-order semiotic]* (fully tristratal [semantics/ lexicogrammar/ phonology – possibly expanded later by graphology at the level of expression and by grammatical metaphor in the relationship between semantics and lexicogrammar]; and

metafunctional [ideational (logical + experiential), interpersonal, textual – as simultaneous modes of meaning]).

This account embodies both continuity (the functional strands) and transformation (the metafunctional re-interpretation; the stratal fission of content with the emergence of lexicogrammar) and lexicogrammar is seen as arising as an emergent formation of a new stratum (cf. Luc, 1998: 399-402, on "level formation") rather than as the result of genetic preprogramming. This puts into perspective Bickerton's (1995: 68-69) view that the emergence of "syntax" as "a catastrophic event", interpretable in terms of Eldredge & Gould's theory of punctuated equilibrium. (It is important to note that in ontogenesis lexis and grammar develop together. We have to revise common view is that lexis develops first, followed by grammar: the early signs of protolanguage are neither words nor lexical items – they are protolinguistic signs. It is also important to note that at the level of the phonological system, articulation begins as a kind of prosody, developing out of protolinguistic postural vocalizations.)

Each ontogenetic phase has critical implications for both social and biological developments in the growing child. For example, while protolanguage goes with crawling and primary consciousness, language goes with walking and higher-order consciousness (Halliday, 1998). In addition, the account also has clear socio-cultural implications: each phase represents a distinct relationship between language and context, from the constrained phase of protolanguage where the microfunctions are directly tied to particular contexts of use in a one-to-one mapping to the very complex and open-ended relationship that characterizes phase III. This will make it possible at least to explore the question of what our ancestors would have been able to achieve with the linguistic potential characteristic of each particular phase.

The ontogenetic model can, I believe, serve as an explicit and detailed outline for exploring phylogenesis, with one central difference. (This is of course an analytical strategy; as far as language development is concerned, it is the other way around: just as in the biological development of an individual, ontogenetic strategies recapitulate phylogenetic ones – up to a point.) Children learn how to mean in interaction with their immediate caregivers who have a full-fledged linguistic meaning potential and possibly with older siblings. Thus when they have learned the basic principles through their own protolanguages and begin the transition into the mother tongue spoken around them, they have an existing model to draw on in building up their own meaning potentials. This was of course not the case in human evolution: there were no models; and all members of a given social group were at the same stage of language evolution. As a result, I believe the transitional phase must have extended over a very long period of time in phylogenesis even though it is fairly brief in ontogenesis. But this seems to be what the general picture of human evolution suggests – a long period of relative phylogenetic stasis for the duration of Homo erectus from around 1.8 m years ago until the appearance of archaic Homo sapiens some 400-200 K years ago. There would thus have been a gradual build-up in semiotic complexity during the transition, followed by a more rapid evolution once the break-through to language had taken place.

It seems plausible that some form of phase I – protolanguage – is very old indeed – predating the evolution of the hominid line. It would have been multifunctional from the beginning, evolving in regulatory, interactional, instrumental and personal contexts of use rather than being restricted to something like social grooming or bonding or linked to something like tool production and use (which can be achieved through non-linguistic apprenticeship; cf. Savage-Rumbaugh & Lewin, 1994: 246). It is very likely to be a semiotic potential shared with apes in the wild (as I would interpret accounts of the natural semiotic systems of modern chimpanzees [cf. Goodall, 1986], bonobos [Savage-Rumbaugh & Lewin, 1994: e.g. 106-107; 112-113, 119] and orang-utans [cf. Kaplan & Rogers, 1999]) – perhaps even one that goes back to the evolution of primates some 60 million years ago, correlating with R. Foley's (1997: 173-174) observation that "primates are the social order par excellence" and that "sociality is really part of the primate core adaptation". In human infants, protolanguages seem to develop to the size of 60-70 signs; but it seems quite plausible that such inventories would have been further extended during phylogenesis until the point where the protolinguistic system had to evolve a more complex form of organization to cope with the meaning-making pressures it had helped to create. Protolanguages of this kind would probably have been vocal-gestural combinations in expression (including facial expression of the kind we find with modern apes), just as with human infants, and would thus not have been subject to any of the requirements on vocal production characteristic of modern languages (as emphasized by Philip Lieberman); these requirements would only have emerged much later in the evolution of language. Such protolanguages would have enabled members of groups to share attention and togetherness, to exercise simple control, to issue warnings – that is, they would have provided these members with microfunctional modes of meaning situated within very specific types of context; but they would not have enabled them to engage in dialogue (as opposed to simple turn taking), to construe their experience (e.g. by naming plants and animals independently of the microfunctional act), to create and exchange information; and it would not have enabled them to mean more than one thing at the same time. These semiotic capacities only begin to become possible with the transitional phase.

In the paper, I will explore the question of when the transition might have started – relating it to observations about both biological and social changes (against the background of material changes); and I will explore the question of when the transitional phase may have been evolving into the modern linguistic phase. While it seems plausible that the transitional phase would have begun already with Homo habilis (2 to 2.5 million years ago) – and certainly with Homo erectus (around 1.8 million years ago), the particular challenge is to identify the possible period during which the transitional phase evolved into the modern linguistic phase: did this happen when archaic Homo sapiens appears (200-400 K years ago) or when anatomically modern Homo sapiens appears (very roughly somewhere around 100 K years ago); had it happened by the time people migrated to Australia (perhaps even 60 K years ago); what linguistic changes correlate with the "Upper Paleolithic revolution" when there is an apparent explosion of cultural artifacts (perhaps around 40-30 K years ago)? Based on evidence from existing languages and assumptions about branching and stock density, Nichols (1998: 138-139) estimates "the linguistic age of the world" to be around 132,000 years: this is the time it would take "to populate the traditionally inhabited world with language families at the density that can be reached when circumstances are favourable – that of New Guinea". This is one possible source of evidence, pushing the date back before the first traces of anatomically modern humans; but it does not extend further back into the long transitional phase of language evolution.

References

Beaken, Mike. 1996. *The making of language.* Edinburgh: Edinburgh University Press.

Bickerton, D. 1990. *Language and species.* Chicago: Chicago University Press.

Bickerton, D. 1995. *Language and human behaviour.* London: UCL Press.

Deacon, Terrence. 1992. Brain-language coevolution. In Hawkins & Gell-Mann (eds.). 49-83.

Deacon, Terrence. 1997. *The symbolic species: the co-evolution of language and the human brain.* Harmondsworth: Penguin Books.

Delsemme, Armand. 1998. *Our cosmic origins: from the Big Bang to the emergence of life and intelligence.* Cambridge: Cambridge University Press.

Edelman, G. 1992. *Bright air, brilliant fire: on the matter of the mind.* New York: Basic Books.

Foley, Robert. 1997. *Humans before humanity: an evolutionary perspective.* Oxford: Blackwell.

Foley, William A. 1997. *Anthropological linguistics: an introduction.* Oxford: Blackwell.

Goddall, Jane. 1986. *The chimpanzees of Gombe.* Cambridge, MA: Harvard University Press.

Halliday, M.A.K. 1975. *Learning how to mean.* London: Edward Arnold.

Halliday, M.A.K. 1995. On language in relation to the evolution of human consciousness. S. Allén (ed.), *Of thoughts and words: proceedings of Nobel Symposium 92 "The relation between language and mind", Stockholm 8-12 August 1994.* London: Imperial College Press; Singapore: World Scientific Publishing Co.

Halliday, M.A.K. 1998. Representing the child as a semiotic being (one who means). Paper presented to conference «Representing the child», Monash University, 2-3 October 1998.

Halliday, M.A.K. & Christian M.I.M. Matthiessen. 1999. *Construing experience through meaning: a language-based approach to cognition.* London: Cassell.

Hawkins, John A. & Murray Gell-Mann (eds.) 1992. *The evolution of human languages.* Menlo Park, CA: Addison-Wesley. (Santa Fe Institute, Studies in the Sciences of Complexity. Proceedings Volume XI.)

Hurford, James R., Michael Studdert-Kennedy & Chris Knight (eds.) 1998. *Approaches to the evolution of language: social and cognitive bases.* Cambridge: Cambridge University Press.

Jablonski, Nina G. & Leslie C. Aiello (eds.). 1998. *The origin and diversification of language.* Memoirs of the California Academy of Sciences, Number 24.

Kaplan, Gisela & Lesley J. Rogers. 1999. *The orang-utans.* St Leonards, NSW: Allen & Unwin.

Layzer, David. 1990. *Cosmogenesis: the growth of order in the universe.* New York & Oxford: Oxford University Press.

Mithen, Steven. 1996. *The prehistory of mind: a search for the origins of art, religion and science.* London: Thames and Hudson.

Nichols, Johanna. 1992. *Linguistic diversity in space and time.* Chicago: Chicago University Press.

Noble, William & Iain Davidson. 1996. *Human evolution, language and mind: a psychological and archaelogical inquiry.* Cambridge: Cambridge University Press.

Painter, Claire. 1984. *Into the mother tongue: a case study in early language development.* London: Pinter.

Ruhlen, Merritt. 1994. *On the origin of languages: studies in linguistic taxonomy.* Stanford, CA: Stanford University Press.

Savage-Rumbaugh, Sue & Roger Lewin. 1994. *Kanzi: the ape at the brink of the human mind.* New York: Wiley.

Smith, John Maynard & Eörs Szathmáry. 1999. *The origins of life: from the birth of life to the origin of language.* Oxford: Oxford University Press.

Steels, Luc. 1998. Synthesizing the origins of language and meaning using coevolution, self-organization and level formation. In Hurford, Studdert-Kennedy & Knight (eds.). 384-404.

Torr, Jane. 1997. *From child tongue to mother tongue: a case study of language development in the first two and a half years.* University of Nottingham: Department of English Studies. (Monographs in Systemic Linguistics 9.)

Trevarthen, Colwyn. 1987. Sharing making sense: intersubjectivity and the making of an infant's meaning. In R. Steele & T. Threadgold (eds), *Language topics: essays in honour of Michael Halliday.* Volume 1. Amsterdam: Benjamins.

Grunts: A gateway to vocal communication and language?

Lorraine McCune

Rutgers University

There are shared properties of communication systems currently observeable in extant species including humans and other primates that offer a window into communciation systems of common ancestors and suggest a mechanism for the emergence of systematic vocal communication. A 1996 study of the transition to reference in human infants (aged 9-16 months, McCune, et al., 1996) demonstrated that referential language was first observed in the month following first communicative use of grunt vocalizations. The autonomic grunt is an involvuntary vocalization resulting from a sequence of physiological events attendant upon metabolic demand for oxygen following effort, thermoregulation, and possibly other events. Briefly, reduced oxygen blood level sets in motion reflexive activation of the intercostal muscles and laryngeal closure, mediated by the vagus, the tenth cranial nerve. When respiratory rhythm demands inhalation, a sudden burst of air at the larynx leads to the grunt vocalization.

Communicative use of grunts in human infants demonstrated a voluntary non-autonomic production of this vocalization. Examination of developmentally earlier contexts of grunt use revealed that the infants first grunted under conditions of physiological effort, then during bouts of focussed attention, and finally in communicative contexts. Review of primate literature revealed that many species of nonhuman primates include communicative grunts in their repertoire, and some like vervet monkeys (Cheyney & Seyfarth, 1982), and chimpanzees (Goodall, 1986) use grunts with consistent acoustic properties in specific contexts. Developmental examination of the literature reveals that in both of these species grunts of effort have been observed prior to communicative grunt use, and that use of communicative grunts precedes appropriate use of other calls in the species vocal repertoire (Plooij, 1984; Seyfarth & Cheney, 1986).

Based on these observations I propose that infant primates' experience of their own autonomic grunts in conjunction with internal states of effort prompts their attention to this juxtaposition of an interior "mental" state and a vocalization. Following Tinbergen (1952) I suggest that the grunt vocalization, initially integrated with effortful activity, becomes a displacement activity associated with internally effortful or focussed states. Given the common occurrence of this vocalization among conspecifics, the induction of a matchng internal state in the communicative partner is likely and would facilitate joint understanding. Such matching internal states can then form a bridge to induction of the young into the existing adult communicative repertoire. In humans this is attested by the shift into referential language learning at the point when communicative grunts are observed, and by the child's subsequent request for names for a variety of objects encountered. McCune (1999) presents an analysis of comparative data in support of this thesis.

This developmental analysis exhibits the Darwinian (1872/1965) proposition regarding the derivation of purely expressive acts from those originally serving some basic function for the animal. He first suggested that a "useful" behavior accompanying a particular internal state might later recur under similar internal states, in which although in the later cases the behavior may serve no direct use, it nonetheless expresses information about the internal state of the animal. With respect to evolution of communication I propose that this developmental mechanism has operated across millennia as a gateway for new members of each primate species into the vocal communicative repertoire of that species, regardless of the simplicity or complexity of that repertoire. Thus the cradle of meaning is attributed always to infancy, to initial mother/infant communication. As adults of a community broaden or strengthen that repertoire, the development of laryngeal control in service of communication would continue to unlock the initial rungs of the communicative ladder for each new member. The continued use of communicative grunts by adult humans across cultures (e.g., “mm” and “uh huh” in English; Schlegoff, 1972) provides a comparison point with adult vervets and chimpanzees who rely upon such vocalizations for regulating much of their social life. I attach Figures 1-3 to demonstrate comparative steps in three primate species to their quite diverse repertoired, including the developmental timing of occurrenceof grunts.

References

Cheney, D. L., & Seyfarth, R. M. (1982). How vervet monkeys perceive their grunts: Field playback experiments. *Animal Behaviour*, *30*, 739-751.

Darwin, C. (1872/1965). *The expression of the emotions in man and animals*. Chicago: University o Chicago Press.

Goodall, J. (1986). *The Chimpanzees of Gombe*. Cambridge, Mass: Harvard University Press.

McCune, L. (1999). Children's transition to language: A human model for development of the vocal repertoire in extant and ancestral primate species? In B. J. King (Ed.), *Evolotion of language: What can evidence from nonhuman primates tell us?* Santa Fe: SAR Press.

McCune, L., Vihman, M. M., Roug-Hellichius, L., Delery, D. B., & Gogate, L. (1996). Grunt communication in human infants. *Journal of Comparative Psychology*, *110*, 27-37.

Plooij, F. X. (1984). *The Behavior of Free-living Chimpanzee Babies and Infants*. Norwood, NJ: Ablex.

Schegloff, E. (1972). Sequencing in conversational openings. In J. Gumperz & D. Hymes (Eds.), *Directions in Sociolinguistics* (p. 3). New York: Holt, Rinehart & Winston.

Seyfarth, R., & Cheney, D. (1986). Vocal development in vervet monkeys. *Animal Behaviour*, *34*, 1640-1658.

Tinbergen, N. (1952). "Derived" activities; their origin, causation, biological significance, origin, and emancipation during evolution. *The Quarterly Review of Biology*, *27*, 1-32.

Birth	*3 months*	*7-9 Months*	*9-11 months*	*11-15 Months*	*13- 28 Months*
Cry					
Grunt					
	Effort Grunt				
	Coo				
		Attention Grunt			
		Babble			
			Communicative		
			Gestures		Early Talkers:
			(Context-limited Words)		13-16 + Months:
				Communicative Grunt	Vocal Motor Schemes for 2 Consonants
				Representational	
				Meanings in	
				Pretend Play	Later Talkers 18-28 Months
				Language Comprehension	Vocal Motor Control In Words/ Sentences
					Referential Words

Table 1: *Steps Toward Language In Human Infants*

Birth	*1 to 3 Months*	*3-4 Months*	*6-7 Months*	*1-2 Years*
Scream (adult-like)				
Whrr (adult-like)				
Grunt – some features adult-like, but others significantly different				
	(Effort ?) Grunts			
	Often accompany movement			
	Alarm Calls begin. They are acoustically like the adults' but applied to broader categories of species	Attention (?) Grunts 60% Predict or notice movement (adult use)		
		Same response to all alarm types	Consistent adult-like response to each alarm type	Communicative (?) Grunts develop all adult acoustic properties and appropriate use
				Alarm calls develop appropriate use

Table 2: *Steps Toward Adult Communication in Vervet Infants*

Birth	*2 Months*	*3-5 months*	*7-9 months*	*12 months*
Effort Grunts accompany movement			Vocal data limited	Vocal data unavailable
Uh-Grunts and series are reactions to startling stimuli				
Whimper- Loss of ventral contact	Attention (?) Grunt- Uh-Grunt at sight or sound of other chimps	Mutual regulation in mother-infant play		
Scream - Distress	Whimper- Out of contact+ out of sight	Communicative Grunt - Uh-grunt and series to adult chimps		
Clings; mother supports	Clings ventrally without support			
		Clings on back	Whimper- Out of arm's reach	
	Interest in mother's face	Walks in quadruped	Follows familiar others	
	Grasps objects		Communicative Gestures such as begging	Gestures inviting activities , e.g., tickle
				Plays at nest-building

Table 3: *Steps Toward Adult Communication in Chimpanzee Infants*

Uniformitarian assumptions and language evolution research

Frederick J. NEWMEYER

University of Washington
fjn@u.washington.edu

ABSTRACT

Most linguistic research in this century has been guided by a key assumption, namely, that of ‘uniformitarianism’. In a nutshell, uniformitarianism embodies the idea that all languages are, and always have been, cut from the same mold. The assumption has both political and methodological implications (for discussion, see Newmeyer 1986), though in this paper I will focus only on the latter. If uniformitarianism is right, then the grammatical theorist and the researcher into the evolution of grammar are free to ignore a host of factors that might (*a priori*) be thought to be relevant to their tasks. Among these are sociocultural facts about the speakers of the languages under investigation and the historical periods in which the languages are or were spoken.

There are, however, two ways that uniformitarianism might be mistaken (for full discussion, see Newmeyer 1998: ch. 6). Let's call the first 'WEAK-NON-U'. In this scenario, the functional forces responsible for the observed properties of language (and the correlations among them) have remained constant throughout human history, but they are, so to speak, 'lopsided'. That is, they are propelling language in a particular overall direction as far as its distribution of typological features are concerned. The second I will call 'STRONG-NON-U'. In this scenario, the functional forces *themselves* have changed indeterminately throughout human history. Such could be the case, for example, if there is, contrary to the mainstream view, a non-accidental correlation between 'purely' grammatical features and aspects of culture, climate, and so on.

In this paper I will first review the evidence for both WEAK-NON-U and STRONG-NON-U and then discuss the implications of the fact that both might be wrong for studies of language evolution.

The idea of 'lopsided functional forces' and WEAK-NON-U is explicit in much of the work that posits that there has been a general mostly unidirectional 'drift' from OV order to VO order (Vennemann 1973; Newmeyer in press). In a rather different way, WEAK-NON-U is implicit in Nettle (1999), where it is suggested that typologically rare features are concentrated in languages with small numbers of speakers. The most extensive marshalling of evidence for STRONG-NON-U is found in Perkins (1992), a book that argues that less complex cultures tend to have more complex deixis systems.

The farther back we go in historical time (and the closer that we get to the 'event' that created true human language), the more plausible become both versions of non-uniformitarianism and the more dramatic their probable effects. Take parsing-dictated grammatical consequences, for example, such as the principle of subjacency or the statistical correlation between the order of grammatical relations and adpositionality. Would they have been manifest in 'early human language'? It is not obvious that they would have been. One can easily imagine that in the historical infancy of human language the influence of parsing would have been submerged by more pressing functional needs and that subordination would have been so rare that principles

such as subjacency could not have emerged (note that it is claimed that today's preliterate societies use fewer subordinate clauses than do literate ones: Mithun 1984; Kalmár 1985).

I go on to demonstrate that the great majority of published work in language origins and evolution presupposes both versions of uniformitarianism. Uniformitarianism is implicit, I would say, in the debate over the degree to which grammar is innate and therefore what a theory of biological evolution of language has to explain. For example, as I understand them, the computational simulations in work such as Kirby and Hurford (1997) take uniformitarianism for granted, as do more ‘catastrophic’ scenarios for the emergence of grammar such as Bickerton (1998) and Berwick (1998). The last part of the paper is a general discussion of the extent to which the central conclusions of such work might be maintained in the light of the probable incorrectness of a central assumption underlying it.

References

Berwick, Robert C. (1998). Language evolution and the Minimalist Program: The origins of syntax. Approaches to the evolution of language: Social and cognitive bases, ed. by James R. Hurford, Michael Studdert-Kennedy, and Chris Knight. Cambridge: Cambridge University Press, pp. 320-340.

Bickerton, Derek (1998). Catastrophic evolution: The case for a single step from protolanguage to full human language. Approaches to the evolution of language: Social and cognitive bases, ed. by James R. Hurford, Michael Studdert-Kennedy, and Chris Knight. Cambridge: Cambridge University Press, pp. 341-358.

Kalmár, Ivan (1985). Are there really no primitive languages? Literacy, language, and learning : The nature and consequences of reading and writing, ed. by David R. Olson, Nancy Torrance, and Angela Hildyard. Cmbridge: Cambridge University Press, .

Kirby, Simon and James Hurford (1997). Learning, culture, and evolution in the origin of linguistic constraints. Proceedings of the Fourth European Conference on Artiifical Life, ed. by Phil Husbands and Harvey Inman. Cambridge, MA: MIT Press, pp. 493-502.

Mithun, Marianne (1984). How to avoid subordination. Berkeley Linguistics Society 10, 493-523.

Nettle, Daniel (1999). Linguistic diversity. Oxford: Oxford University Press.

Newmeyer, Frederick J. (1986). The politics of linguistics. Chicago: University of Chicago Press.

Newmeyer, Frederick J. (1998). Language form and language function. Cambridge, MA: MIT Press.

Newmeyer, Frederick J. (in press). On reconstructing 'proto-world' word order. The emergence of language, ed. by Chris Knight, James Hurford, and Michael Studdert-Kennedy, .

Perkins, Revere D. (1992). Deixis, grammar, and culture. Typological Studies in Language 24. Amsterdam: John Benjamins.

Vennemann, Theo (1973). Explanation in syntax. Syntax and semantics, vol. 2 ed. by John Kimball. New York: Seminar Press, pp. 1-50.

Defining animal communication, and why it matters for understanding the evolution of language

Jason NOBLE

Center for Adaptive Behavior and Cognition
Max-Planck-Institut für Bildungsforschung
Lentzeallee 94, 14195 Berlin, Germany
noble@mpib-berlin.mpg.de

Animal communication and human language

Human language may or may not be continuous with animal communication. If there is a continuity, and language is an elaborated version of an earlier animal signalling system, then clearly language-evolution theorists should learn as much as they can about how animal communication evolves. If there is no continuity, and human language turns out to be an adaptation with a unique genesis and structure, then theories of language evolution will still be constrained by general results from evolutionary theory, such as Hamilton's (1964) findings on altruistic behaviour towards kin, or the possibility that the cost of a signal can ensure its reliability (Zahavi, 1975; Grafen, 1990). So ideas on animal communication will have at least some relevance to theories of language evolution.

However, animal communication is a broad subject. Animals influence each other's behaviour in many different ways, and pinning down exactly which kinds of influence that we wish to call communication can be troublesome. It is uncontroversial to say that vervet monkeys are communicating when one gives a leopard alarm and the others scramble for the safety of the trees. But is a camouflaged insect signalling to its predators? By running away, is an antelope signalling to a cheetah? In both cases the answer is yes under certain definitions of communication that have been adopted in the biological literature. Intuitions differ about how such borderline cases should be treated; mimicry and deception are two other notable problem areas.

This definitional problem is despite the fact that in ordinary language we have a clear idea of what we mean by communication, or at least an archetypal image: a sender imparts information to a receiver via some sort of signalling channel. For example, one person says truthfully to another, “It's raining outside.” This has been dubbed the conduit metaphor (Reddy, 1979; Lakoff & Johnson, 1980).

The ordinary language definition has not been sufficient for most biologists who have looked at communication. I will consider three types of definition from the biology literature: those phrased in terms of behavioural influence, those dealing with information transmission, and those invoking the intent to communicate. All three of these categories will ultimately be rejected as inadequate.

Communication as behavioural influence

An example of the first type of definition is Krebs and Dawkins (1984), who defined a “signal” as an action or structure which increases the fitness of an individual by altering the behaviour of other organisms. The main problem with such a definition is over-inclusiveness: one animal directly causing a behaviour in another, for instance by attacking it, would count as communication. So might the incidental transfer of information: it might well be to the advantage of bird A that bird B should notice that it (A) has found food, perhaps because the two are related. However, that would mean that A's behaviour in simply observing food and approaching it counts as a signal to B; this is certainly a counter-intuitive result.

Communication as information transfer

Hurd (1997) provides a good example of communication defined in informational terms: "Information is said to be received whenever an agent changes it's [sic] expectations about the consequences of an action, and communication has occurred whenever the action of one animal transmits information to another." One problem is that this is still over-inclusive, because incidental information transfer qualifies as communication. A deeper difficulty for such a definition is the problem of error: suppose we have a working hypothesis that a certain bark is used by vervet monkeys to mean "leopard approaching!". One day we observe the bark being made when a hyena approaches. We are then faced with a problem of indeterminate content, which we cannot evade by stipulating that the bark means leopard and that all other uses of it are mistakes, because that would be begging the question.

Intentional communication

Grice (1969) put forward the case – later developed by Bennett (1976) and Dennett (1987) – that considering intentionality allows us to pick out a special kind of communication that is genuinely worthy of the name. These authors rely on the intuition that there is a difference worth marking between a situation in which causal automatons exchange signals, and a communication system in which participants really mean what they say. Their argument is that real communication can be roughly equated with human speech acts, and must involve, at a minimum, third-order intentionality. Two problems exist for this definition: first, even if it could be demonstrated that an apparently communicative system involved mere zero-order tropisms, we may well still want to classify the system as communicative. Second, a definition of communication in terms of higher-order intentionality is founded on the dubious premise that a particular animal either unambiguously does or unambiguously does not possess such intentional capabilities.

Proper signalling

After consideration of the alternatives, I will defend an evolutionary-functional definition of communication based on the work of Millikan (1984, 1993). This definition stipulates that true communication, or "proper signalling", involves the production of a signal and the performance of a response that both have a history of selection in this context. We should expect to find proper signals when it is evolutionarily stable for two animals to coordinate their behaviour in an interaction, that is, when there is a mutual benefit in transmitting information. However, this does not mean that proper signals will only evolve in nakedly cooperative situations. Processes such as the handicap principle show that communication can be evolutionarily stable despite an apparent conflict of interests: poor-quality signallers at a handicap equilibrium do not honestly signal their low quality for any mystical reason, but because the excessive costs of exaggeration make it in their interests to do so. It is (apparently) the function of the peacock's tail to signal male quality just as much as it is the function of the bee dance to indicate the location of nectar.

The practical consequences of this definition will be explored. The most important implication is that those sections of the biological literature on animal communication that do not deal with cases of proper signalling will probably not be relevant to work on the evolution of human language.

References

Bennett, J. (1976). *Linguistic Behaviour*. Cambridge University Press.

Dennett, D. C. (1987). *The Intentional Stance*. MIT Press / Bradford Books, Cambridge, MA.

Grafen, A. (1990). Biological signals as handicaps. *Journal of Theoretical Biology, 144*, 517-546.

Grice, H. P. (1969). Utterer's meaning and intention. *Philosophical Review, 68*, 147-177.

Hamilton, W. D. (1964). The genetical evolution of social behaviour. *Journal of Theoretical Biology, 7*, 1-52.

Hurd, P. L. (1997). *Game Theoretical Perspectives on Conflict and Biological Communication*. Ph.D. thesis, Department of Zoology, Stockholm University, Sweden.

Krebs, J. R., & Dawkins, R. (1984). Animal signals: Mind reading and manipulation. In Krebs, J. R., & Davies, N. B. (Eds.), *Behavioural Ecology: An Evolutionary Approach* (Second edition)., pp. 380-402. Blackwell, Oxford.

Lakoff, G., & Johnson, M. (1980). *Metaphors We Live By*. University of Chicago Press.

Millikan, R. G. (1984). *Language, Thought, and Other Biological Categories*. MIT Press / Bradford Books, Cambridge, MA.

Millikan, R. G. (1993). *White Queen Psychology and Other Essays for Alice*. MIT Press / Bradford Books, Cambridge, MA.

Reddy, M. J. (1979). The conduit metaphor: A case of frame conflict in our language about language. In Ortony, A. (Ed.), *Metaphor and Thought*. Cambridge University Press.

Zahavi, A. (1975). Mate selection – a selection for a handicap. Journal of Theoretical Biology, 53, 205-214.

The irrelevance of the lowered larynx in modern Man for the development of speech

John J. OHALA

Phonology Laboratory, Department of Linguistics
University of California, Berkeley, CA 94720 USA
ohala@socrates.berkeley.edu

In speculation about the evolutionary development of speech much significance has been attached to the lowered larynx in humans. The lowered larynx, the argument runs, creates a pharyngeal cavity that is separate from the buccal cavity (the oral cavity bounded on the sides by the cheeks) and that this two-cavity configuration permits the production of more distinct vowels which, it is assumed, is essential or at least advantageous to speech communication. I offer the following arguments against this proposal:

1. Although the larynx is low in adult male humans (vis-à-vis the larynx position in closely related primate species, such as the chimpanzee), it is less low by a considerable amount in adult female humans. Normally in cases of dimorphism, the age group or sex that has more of a given “marked” feature supposedly serving a certain function should be better adapted for that function. For example, among sea lions severe limitation of resources leads to marked size dimorphism in males; the larger males, it has been documented, are better adapted to successfully compete for food, territory, and mates. But are human males with their lowered larynx better adapted for speech than females? No; it is just the opposite. In cases where males and females have equal access to education and health resources, females invariably outperform males on verbal tests. Similarly, males are more prone to speech disorders stemming from neurological conditions such as delayed acquisition of speech, stuttering, and autism by a ratio (in comparison to females) of 4 or 5 to 1.
2. Typically when there is sexual and age-related dimorphism, as is the case with the degree of larynx lowering in humans, the onset of a distinct development of the dimorphic feature is timed to that phase in the individual's life when the feature is needed. For example, some birds develop a so-called “egg tooth” shortly before hatching. This facilitates breaking the shell at the appropriate time. After hatching, this egg tooth gradually disappears. In human males the disproportionate lowering of the larynx (vis-à-vis females) occurs at puberty, long *after* the onset of speech. It is unlikely, then, that it is related to speech function.
3. The hypothesis that the larynx lowering is an adaptation to speech would have to treat as wholly coincidental the fact that there is another remarkable anatomical dimorphic development in adult males vocal apparatus that parallels the descent of the larynx, namely, the enlargement (elongation and increased mass) of the vocal cords. Equally, it would treat as coincidental yet another dimorphic development in males at puberty: the growth of facial hair. However, I present below a hypothesis that integrates all these dimorphic secondary sexual features.
4. Other species – that obviously do not have speech – also show sex- and age-determined enhancements of the vocal apparatus. These range from cranes to howler monkeys to elephant seals. For example, whooping cranes have tracheas that are about as long as the

> birds themselves. The “extra” length of the trachea is achieved by part of the trachea being coiled up inside the sternum (not unlike the convolutions of a French horn). This enhancement of the resonating tube of their vocal apparatus is assumed to contribute to the loud and far-carrying vocalizations that they are known for.

The function of the lowered larynx in humans, especially males, is probably to enhance threat displays rather than speech itself. Threat displays involve both visual and acoustic components. A cross-species analysis of agonistic vocalization in birds and mammals reveals what is called the 'frequency code': vocalizations in threats tend to have low fundamental frequency (F_0), whereas non-threat (submissive or affinitive) vocalizations have high F_0. Ohala (1984, 1994) suggested that the same code applies not only to the F_0 but also to the resonances shaping the vocalization, too. Predominantly low resonances should correlate with threat and high resonances with non-threat. Functionally this comes about since the natural frequencies of a sound producer are inversely related to its size. Low frequencies are associated with a larger – and thus more dangerous – vocalizer. This explains the temporal coincidence of the larynx lowering, vocal cord enlargement, and the growth of facial hair in male humans during puberty. (Facial hair, i.e., beards, in males presumably functions as a visual, implastic, component in threat displays: it makes the bearded one seem larger since his head subtends a larger angle in the visual field of the viewer.) With sexual maturity, males, by virtue of the common sex role they must assume, need to compete for resources. This would account for why males have these anatomical features enhancing threat displays more than females.

Whether the lower larynx, as an adaptation for threat displays, occurred before, during, or after the evolution of speech and language is uncertain. My point is that it is independent of and thus irrelevant to the evolution of speech.

References

Ohala, J. J. 1984. An ethological perspective on common cross-language utilization of F0 of voice. *Phonetica* 41.1-16.

Ohala, J. J. 1994. The frequency code underlies the sound symbolic use of voice pitch. In L. Hinton, J. Nichols, & J. J. Ohala (eds.), *Sound symbolism.* Cambridge: Cambridge University Press. 325-347.

Finite-state syntax in bengalese finch song: From birdsong to the origin of language

Kazuo Okanoya

Faculty of Letters, Chiba University and
PRESTO, Japan Science and Technology Corporation
Okanoya@cogsci.L.chiba-u.ac.jp

Summary

Male Bengalese finches sing complex songs that are expressed by finite-state automata. Females of the same species prefer syntactically complex songs to simple, linear songs. We hypothesized that song complexity arouse from sexual selection, because complex song could be a handicap trait. Behavioral experiments supported this notion. A finite state syntax thus could evolve through sexual selection, independently from the semantics. Conventional theories of language evolution state that the symbolic competence could somehow elaborate to produce syntax, but how this could be done is not specified. Our Bengalese finch studies inspire that the form and content of human language could have developed independently.

Key Words: Birdsong, Generative Syntax, Finite State-Syntax, Hierarchical Structure, Evolution

Song of the Bengalese finch

In passerine birds, song is a learned behavior used by males to attract females (Catchpole & Slater, 1996). Generally, a song is composed of several types of song notes and the order of these notes is stereotyped. However, unlike most songbirds studied to date, Bengalese finches (*Lonchura striata* var. *domestica*) sing non-deterministic songs that can be described by a finite-state syntax.

Bengalese finches critically depend upon auditory feedback when singing (Okanoya & Yamaguchi, 1997). When deafened, temporal organizations of the song immediately deteriorate. This is another special aspect of the Bengalese finch song, but critical dependence to auditory feedback and song complexity may be closely related: Bengalese finches are probably listening their own songs in real time to produce immediate varieties.

Each male Bengalese finch sings a unique song with unique set of song elements and individual specific finite-state syntax. Two to five song elements are organized into a “chunk.” Several chunks are further organized into phrases by passing through a finite-state syntax. Several phrases are, in turn, arranged based on the finite-state syntax. A Bengalese finch can produce infinite varieties by taking different paths on the finite-state automaton. Why do Bengalese finches sing such complex songs? We shall try to answer to this question from four perspectives as suggested by Tinbergen (1963): its mechanism, development, function, and evolution.

Brain hierarchy and behavioral hierarchy

The forebrain song control system in songbirds consists of a set of discrete nuclei including NIf, HVc, and RA (Margoliash, 1997). We hypothesized that each of these nuclei might be responsible for the control of the hierarchical structure of the Bengalese finch song.

When NIf was bilaterally lesioned, the song lost phrase level variability: that is, the complex song changed into simple, stereotyped ordering of linear syntax (Hoshino & Okanoya, submitted). When HVc was partially lesioned, a particular state transition disappeared but all song elements were preserved (Uno & Okanoya, in preparation). Thus HVc is responsible for chunk level variability. When RA was partially lesioned, a certain song note dropped from the song but overall finite-state syntax remained unchanged (Hirata & Okanoya, in preparation). RA, thus, is responsible to produce each note. Taken together, we found that the finite-state syntax is expressed in hierarchically organized brain nuclei.

Functions of the finite-state song syntax : A handicap process

Next, to answer from the functional point of view, we examined reproductive behavior of females when stimulated with complex or simple song syntax. Number of nesting material carried each day when stimulated with complex or simple song syntax was counted and compared. Female finches carried more nesting materials and had higher levels of estrogen when stimulated with the complex song.

We thus showed that a song with complex syntax could effectively stimulate the reproductive system of females. From this result we postulate that a complex song patterning should be more attractive to female birds, and therefore the song syntax in Bengalese finches may have evolved through sexual selection (Darwin, 1871).

Why do females prefer song complexity? An ability to sing a complex song may be an honest signal that can advertise the potency of the singer (Zahavi and Zahavi, 1997). Since singing a complex song may require 1) higher level of testosterone, 2) more cognitive load, 3) and more brain spaces. We tested one of the above hypotheses by perturbating an ongoing song by a flashlight. Birds with lower song linearity (i.e., more complex song) were less prone to stop singing (Okanoya & Nakamura, in preparation). The data support that singing a complex song probably requires more cognitive load and make the animal less careful about potentially dangerous situations.

Evolution of the finite-state song syntax

Bengalese finches are domesticated strain of the wild white-backed munia. White-backed munias were imported in Japan about 250 years ago and domesticated since then. Aviculturists selected white-backed munias based on their parental abilities, but never by their songs. Nevertheless, we found that the complex song syntax was absent in the ancestor species (Honda & Okanoya, 1999). The finite-state song syntax in Bengalese finches must have evolved during 250 yeas.

We hypothesized that the sensory-exploitation process (Ryan, et al., 1990) worked upon the evolution of complex song syntax in Bengalese finches. That is, females of the ancestor species had a preference for complex songs, although males of the ancestor species were not able to sing complex songs because of constraints in the nature including predation risk, foraging cost.

We tested this hypothesis again by examining reproductive behavior of the ancestor species. A simple, linear song of a white-backed munia male was spliced into song notes and a finite-state syntax from a Bengalese finch was applied on the song notes of the munias. We thus composed a new song that had the phonology of the white-backed munia but had the syntax of the Bengalese finch. We played either the original simple munia song, or the newly composed hybrid song. The string-carrying behavior was much more active when stimulated with the hybrid song.

Based on these results, we discuss the evolution of complex behavior and associated changes in the brain. Bengalese finches developed finite-state syntax presumably through the process of sexual selection. However, each token in the Bengalese finch song does not have any associated semantics, and the finite-state syntax in Bengalese finch song does not produce any meanings. Thus, a finite state syntax could evolve without meaning.

Independent evolution of form and content

Most authors arguing the origin of language assume that proto-language that had symbolic contents but no syntactical structure evolved first, then the proto-language somehow evolved into true language with the syntax (i. e., Bickerton, 1995; Pinker, 1994). In this scheme, ability for state-dependent signaling could evolve into an ability to represent symbols through natural selection (i.e., Munn, 1986). Then, the ability to develop symbolic representation could evolve into the competence to produce a context free grammar through social elaboration. However, how the proto-language acquired the syntax is always the hardest question to answer when considering the origin of language (but see Calvin & Bickerton, 2000). Through the analyses of Bengalese finch songs, we reached to a hypothesis that a syntactical behavior could evolve without a need of semantics.

We propose that the symbolic aspect and the syntactical aspect of human language could have evolved independently. As in Bengalese finch songs, rudiment of syntax might be evolved thought sexual selection in humans; sexual display between males and females might be a pre-adaptation to the syntax.

This could happen if an animal that use courtship display with a serial pattern generator evolved into an animal that use a finite-state automaton for courtship display through sexual selection. To produce complex sequences of behavior, a random number generator would be more appropriate. However, the nervous system would rather develop a finite-state machine than a random number generator. Using the vehicle of sexual display as a rudiment of syntax, symbolized tokens were presumably arranged to form a primitive, but true language.

When the animal had already established a competence for the symbolic representations, an elaboration between the syntactical module and the symbolic module by social and sexual selection pressures would reach to a context free grammar (Pereira & Wright, 1998). This could happen by making an external indexing system to bootstrap the finite-state syntax.

Acknowledgements

Research supported by PRESTO, Japan Science and Technology Corporation.

References

Bickerton, D. 1995 *Language and human behavior*. Seattle: University of Washington Press.

Catchpole, C. K. & Slater, P. J. B. (1996). *Bird Song: biological themes and variations.* Cambridge: Cambridge University Press.

Calvin, W. H. & Bickerton, D. 2000 *Lingua ex machina: Reconciling Darwin and Chomsky with human brain.* Cambridge: MIT Press.

Darwin, C. 1871 *The descent of man and selection in relation to sex*. London: John Murray.

Hoshino, T. & Okanoya, K. Lesion of a higher-order song control nucleus disrupts phrase-level complexity in Bengalese finches. Submitted

Honda, E. & Okanoya, K. 1999 Acoustical and syntactical comparisons between songs of the white-backed munia (*Lonchura striata*) and its domesticated strain, the Bengalese finch (*Lonchura striata* var. *domestica*). *Zoological Science, 16*, 319-326.

Margoliash, D. 1997 Distributed Time-Domain Representations in the Birdsong System. *Neuron, 19*,963-966

Munn, C. A. 1986 Birds that "cry wolf" *Nature, 306*, 583-584.

Okanoya, K. & Yamaguchi, A. 1997 Adult Bengalese Finches (*Lonchura striata var. domestica*) Require Real-Time Auditory Feedback to Produce Normal Song Syntax. *Journal of Neurobiology*, *33*, 343-356.

Pereira, F. C. N. and Wright, R. N. 1998 Finite-state approximation of phrase-strucutre grammers. In: E. Roche & Y. Schabes (Eds.), *Finite-state language processing*, Cambridge: MIT Press, pp149-174.

Pinker, S. 1994 *The language instinct*. New York: William Morrow and Company.

Ryan, M., Fox, J. H., Wilczynski, W. & Rand, A. S. (1990). Sexual selection for sensory exploitation in the frog Physalaemus pustulosus. Nature, *343*, 66-67.

Tinbergen, N. (1963) On aims and methods of Ethology. *Z. f. Tierpsycholgie*., *20*, 410-433.

Zahavi, A. & Zahavi, A. 1997 *The Handicap Principle*. Princeton: Princeton University Press.

Immediate checkability as a criterion for the establishment of the very first words

Ernest PALFREY

Coxboro' Cottage SL6 9HR UK

ABSTRACT

Early hominids would no doubt use deception to benefit themselves. But used in the earliest speech, deception would tend to cause this *very first speech* to self-destruct. In the psychological or social sphere, there is a further problem: there are inherent difficulties in checking the truth or accuracy of psychological/social 'words' used to represent, say, attitudes or beliefs in the mind of the speaker. Between them, these difficulties would make it unlikely that the very earliest use of 'words' was in this area: there is no ready way to check the truth or the accuracy of a psychological/social statement. True today, it was no doubt even more true in the earliest days of speech.

But there *is* an area where 'words' can be easily checked: the area of the material world. The reasons are that, firstly, we generally want to know only the surface apparent qualities ... not, for instance, an inner, mental life. Secondly, this information is usually readily apparent and easily checked.

What has turned us away from such a plain and straightforward view is, I believe, the idea of increasing social intelligence (a well-established fact in the primate line) culminating in speech. But for the very first 'words', the difficulty of checking the truth or accuracy of any 'words' used for psychological/social meanings … this presents a real hurdle.

Two consequences follow from the ready checkability of material things. The first is that any 'social use' origin for the very first 'words' is not really possible … *if there is this easier alternative route.* Secondly, because vision would be the preferred sense for immediate checking, nocturnal animals –however social – would not have developed speech; and the move by early hominids from woodland to savannah would have made immediate checkability of material objects and their names easier … while bringing no obvious improvements to checkability in the psychological/social area.

A significant natural background exists in that *feedback checking* is a biological universal for all muscle use: in every activity, there is simultaneous feedback to provide corrections to achieve a goal. Speech is a muscle activity, involving the mouth, throat and so forth. It is suggested that speech checking could not have been avoided: it would have been automatically applied, not only to the production of the correct sound heard, but also to the appropriateness of the 'word' used.

If the first use of 'words' was to describe the material world, there was a ready bridge to the psychological, because hominids were made (like us) of material: arms and legs, hands and fingers, a face…eyes and ears, nose and mouth, material objects which existed and moved. These and their activity would have psychological meaning: for example, a contorted or relaxed face, glaring or smiling eyes, clenched fists or open hands, and so forth.

The final argument is to point out that speech is fiendishly flexible. Like computers, it can be used in SO many ways. And like computers, present use … and *most* speech today is used to convey psychological/social meaning … is no certain guide to very first use.

The conclusion is that the very first use of ‘words’ was to label objects and their movement in the material world, and not to label psychological/social meanings. Speech had to learn to walk … before it could run, in the essentially ambivalent world of the psychological/social.

Cognitive, neurobiological, and evolutionary links between language and memory

Michael D. PATTERSON & Benjamin Martin BLY

Rutgers University - Newark, Department of Psychology
patterson@psychology.rutgers.edu
ben@psychology.rutgers.edu

ABSTRACT

There have been many attempts to account for the neuroevolution of human language, focused on topics ranging from the adaptation of motor processing systems in the basal ganglia, to the hypothesized development of a brain region for theory of mind (for a review, see Patterson and Bly, 1999). Some of these theories focused on the development of a specific area in the brain dedicated for language specific processing. We propose that language, while it is an elaborate and specific system of behaviors, is woven inextricably into another central component of cognition: memory, and therefore the neural substrates of language are intimately connected to those of memory. We investigate the relationship between language and memory, drawing on three sources of evidence: the link between the development and use of memory subsystems during language acquisition, the link between memory processes and language in adults, and a comparison of human and nonhuman primate memory and language abilities. Recent theories suggest that memory can no longer be thought of as a repository for information separate from other processing resources, or as a single location in the brain (Miyake and Prah, 1999). Recent research has also revealed that language processing can no longer be considered limited to a few specialized areas such as Broca's and Wernicke's area (for review, Neville et al., 1997). How separate are memory processes used for language from memory processes used for other cognitive functions? Studies in our lab and other labs (Caplan and Waters, 1999) will be reviewed, from the viewpoint of several different working memory models. Both behavioral and neuroanatomical comparisons of humans and phylogenetically related animals will be examined. Deacon (1997), has suggested a link between memory and language evolution. Only humans seem capable both of acquiring a large lexicon of more than a few hundred items and of learning syntactic rules (Bates and Goodman, 1997). We propose that a rearrangement of memory subsystems may have been responsible for the appearance of both of these capacities in humans.

References

Bates, E., & Goodman, J. (1997). On the inseparability of grammar and the lexicon: Evidence from acquisition, aphasia and real-time processing. In G. Altmann (Ed.), Special issue on the lexicon, *Language and Cognitive Processes, 12*(5/6), 507-586.

Caplan, D., & Waters, G. S. (1999). Verbal Working Memory and Comprehension. *Behavioral and Brain Sciences,22*(1): 77-126.

Deacon, T. (1997). The symbolic species: The coevolution of language and the brain. W.W. Norton: New York.

Miyake, A., and Prah (Eds.) (1999) . Models of Working Memory: Mechanisms of active maintenance and executive control. Cambridge University Press: New York.

Neville, Helen J; Bavelier, Daphne (1998). Neural organization and plasticity of language. *Current Opinion in Neurobiology*. Vol 8(2). 254-258.

Patterson MD, Bly BM (1999) The Brain Basis of Syntactic Processes: Architecture, Ontogeny, and Phylogeny. In Bly BM, Rumelhart DE (eds.) Handbook of Perception and Cognition: Cognitive Science. Academic Press, San Diego, CA.

The emergence of phonology in a population of artificial agents in a phonetic naming game

Emmanuelle Perrone

Laboratoire Dynamique du Langage (CNRS-UMR 5596)
Université Lumière, Lyon 2.
Emmanuelle-Perrone@ish-lyon.cnrs.fr

Introduction

Various recent studies have approached phonology as a complex adaptive system (Lindblom, 1998, De Boer, 1997). In this work we study the emergence of phonology in a population of artificial agents communicating by sounds to define which are the basic rules in interaction that lead to the emergence of the categorisation of phonology.

The LEMMingS toolkit

The tool used for modelling in this experience is a multi-agent system because it allows the study of emerging language phenomena in a population of artificial communicating agents.

The phonetic naming game is a module of the LEMMingS (Language Evolution Monitoring and Modelling System) platform developed at the Laboratory Dynamique Du Langage. It defines a new type of interactions inspired from the imitation game (De Boer, 1999) and the naming game (Steels, 1997) and allows to study phonological aspects of the origins of language in a population of agents exchanging monosyllabic words.

The phonetic naming game

At the beginning of the game, the agents have empty memories. In the interactions, they will have to name the objects of their virtual environment and invent a phonology for their lexicon.

The structure of the agents is composed of two modules and two memories :

- The production module is an acoustic tubes dynamic model based on the Distinctive Region Model of Carré (1995) : a sensibility function allows to define dynamically which are the acoustic modification produced by the gestures of the vocal tract.
- The perception module : the agents segment the acoustic signal into three formants coordinates (values of the 3 first formants of the vowel. The agents are able to retrieve the articulation of a sound from acoustic values with an inverse mapping based on the sensibility function of the DRM model. A threshold is given to the perception module to define the maximal acoustic distance tolerated between two sounds.
- The repertoire of sounds is composed of all the productions stored by the agent. The "phonemes" are represented in the phonetic memory by associations between articulatory gestures and the formants values of the sounds.
- The associative memory stores the lexicon. Words are formed by a sequence of three "phonemes" and are represented as an association by a meaning and a topic.

The game

Two agents of the population randomly chosen play the game. The speaker chooses a topic and sends the associated name to the hearer. The imitator looks in his phonetic memory for the closest representative sounds corresponding to the received sequence. If he has the 3 sounds, he looks if an association corresponds to the topic pointed by the speaker and compares the association to the one sent by the initiator. If the agent has no similar association, he creates a new word for the topic. The sounds of the sequence lacking from the phonetic memory are reproduced with an inverse mapping procedure. The association is then sent to the initiator.

The speaker compares the sounds sequence to the one he has emitted and sends a signal of failure or success to the receiver. At the end of the game the success and use of the sounds and of the associations are updated. If during the game the hearer has stored a new sound, he merges it with the closest sound of his repertoire that is in the merging area.

First results and discussion

These first results show a good convergence : the agents have similar vocalic systems although they present inter-individual differences.

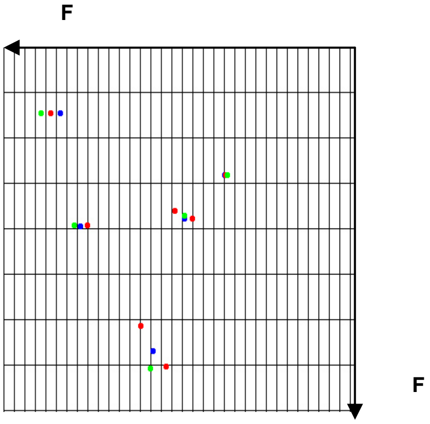


Figure 1 : vocalic system for 3 agents, 40 concepts.

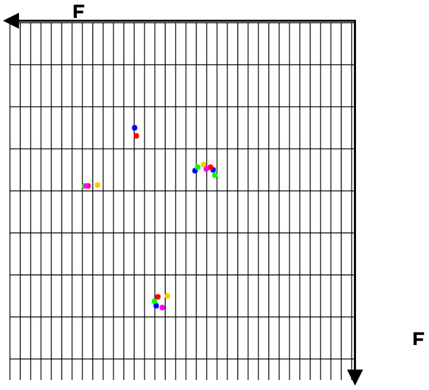


Figure 2 : vocalic system for 5 agents, 10 concepts

Nevertheless, the vocalic systems resulting of the communications in a population of 5 agents for 10 concepts don't show a similarity to any of the phonetic systems encountered in the world's languages.

The agents vocalic systems developed should show a greater realism with the implementation of the coarticulation model producing consonants and a more fine-grained perception module that can take into account syllables transitions.

The game offers a good way to study the lexicon pressure on the phoneme creation and the influence of the population size and the categorisation procedure on the emergence of common vocalic systems in a population of communicating individuals (Lindblom, Mac Neilage and Studdert-Kennedy, 1984).

Although this experience is an oversimplification of the communication in human populations, it offers a good tool to study the minimal conditions for the emergence phonology.

From icon to symbol:
An important transition in the evolution of language

Ullin PLACE

ABSTRACT

In constructing a speculative scenario for the evolution of language (Place forthcoming), I take it as axiomatic that before human beings developed the ability to communicate vocally, they communicated with one another by means of some form of sign language whose primary function was to organize the cooperative social activities ivolved in hunting ang foraging. In such a language the signs wold be entirely iconic, that is to say they would function as signs entirely by virtue of a resemblance between the sign and the object or movement it stands for. There would be no selective pressure to develop symbolic signs where the connection between the sign and what it stands for is purely arbitrary. However, studies of sign languages in general (Frischberg 1979) and of the homesigning developed by isolated deaf individuals in particular (Kuschel 1973) show that although the incidence of arbitrary symbols is much reduced as compared with vocal language, they are very much more common than would be predicted on this hypothesis. The explanation of this finding seems to be that AFTER the development of vocal language when the selective pressure for the development of symbolic signs was much greater due to the limited possibilities of representing objects and movements iconically, a genetic mutation occurred the effect of which was to give human beings a strong preference for communication by means of arbtirary symbols. The evidence for this mutation comes froma quite different source, from experimental studies that have been conducted in recent years within the behavior analytic paradigm on the phenomenon known as "stimulus equivalence" (Sidman 1971; 1986; 1990; Sidman and Tailby 1982). Associated with this is the work that has been done at the University of Wales, Bangor, (Dugdale and Lowe 1990; Horne and Lowe 1996) on the link between the emergence of stimulus equivalence responding in the child of 2-3 years and the so-called "naming explosion" observed by developmental psychologists which appears to occur at the point in development at which language "takes off" in the human child and fails to do so in its closest primate relatives.

References

Dugdale, N. and Lowe, C. F. (1990) "Naming and stimulus equivalence". In D. E. Blackman and H. Lejeune (Eds.) *Behaviour Analysis in Theory and Practice: Contributions and Controversies*. Hove and London: Lawrence Erlbaum Associates. pp. 115-138.

Frischberg, (1979) In Klima and Bellugi

Horne, P. and Lowe, C.F. (1996) "On the origins of naming and other symbolic behavior". *Journal of the Experimental Analysis of Behavior*, 65

Kuschel, R. (1973)

Place, U. T. (forthcoming) "The role of the hand in the evolution of language". *Psycoloquy*,

Sidman, M. (1971). "Reading and auditory-visual equivalences". *Journal of Speech and Hearing Research*,14,5-13.

Sidman, M.(1986). “Functional analysis of, emergent verbal classes”. In T. Thompson, N. Dugdale and M. D. Zeiler (Eds), *Analysis and Integration of Behavioral Units*. Hillsdale, N.J.: Lawrence Erlbaum Associates Inc., pp. 213-245.

Sidman, M. (1990) “Equivalence relations: where do they come from?” In D. E. Blackman and H. Lejeune (Eds.) *Behaviour Analysis in Theory and Practice: Contributions and Controversies*. Hove and London: Lawrence Erlbaum Associates. pp. 93-114.

Sidman, M.& Tailby, W.(1982). “Conditional discrimination vs. matching to sample: An expansion of the testing paradigm”. *Journal of the Experimental Analysis of Behavior*,37, 5-22.

Languages and genes: Modes of transmission observed through the analysis of male-specific and female-specific genes.

Estella S. POLONI*, Nicolas RAY*†,
Stefan SCHNEIDER* & André LANGANEY§*

** Laboratoire de Génétique et Biométrie*
Département d'Anthropologie et Ecologie, Université de Genève, Geneva;
† United Nations Environment Programme,
Global Resource Information Database (UNEP/GRID), Geneva; and
§ Laboratoire d'Anthropologie Biologique, Musée de l'Homme, Paris.

ABSTRACT

The human genome is composed of about hundred thousand different genes. Because of stochastic processes occurring during the transmission of genes from one generation to the other, each of these genes has its own more or less independent history. However, as shown in the eighties with the pionneering studies of Luca Cavalli-Sforza, Robert Sokal and their collegues, the affinities of human populations based on their genetic variability generally correspond with the classification of the languages spoken by these populations into linguistic families. A strong correlation of the processes of genetic and linguistic differentiation is thus suggested (see Cavalli-Sforza et al. (1988). *Proc Natl Acad Sci USA 85*: 6002-6006, and Chen et al. (1995). *Hum Biol 67*: 595-612).

There are two components of the genome whose story will also depend on the particular history of one sex: these are the non-recombining Y chromosome, strictly paternally transmitted, and the mitochondrial genome (mtDNA) strictly maternally inherited. Contrasting the genetic variability of these two genomic components has been shown to be usefull to infer resemblances or differences in the history of males and females (for a general review, see e.g. Owens and King (1999). *Science 286*: 451-453).

In a previous study, we compared the genetic and linguistic affinities of a small set of human population samples, as seen through the analysis of a male-specific genetic system (on the Y chromosome) and a female-specific one (on mitochondrial DNA). We observed that language classification was more correlated to the genetic affinities of populations based on the Y chromosome than on mitochondrial DNA. This result suggests that, to the end of reproducing, women may have moved more than men through linguistic borders, and doing so, women may have preferentially transmitted the language of the fathers to their progeny than their own language. We resumed this general pattern by suggesting the hypothesis that, on the long term, language is paternally transmitted (Poloni et al. (1997). *Am J Hum Genet 61*: 1015-1035).

However, the results of this previous analysis were based on a rather small set of population samples. Moreover, the process of human populations genetic differentiation is also correlated with the geographic distance that separates the populations. Language classification and geographic distance can thus be considered as two predictor variables of the genetic affinities among populations, but these two predictor variables are not independent since distinct linguistic families usually also extend on distinct geographic areas. In our previous work we

estimated the proportion of populations genetic differentiation that is explained by language classification independently from geography. However, geographic distances were computed very roughly, so that they didn't represent very realistic migration routes between populations.

Thus our aim, here, is to re-examine the relation between genetics, linguistics and geography in males and females through a new analysis scheme, that is: (1) by extending our previous analysis to more population samples; (2) by testing alternative linguistic classifications; (3) by using a new procedure that takes into account different environmental variables such as topography, hydrography and vegetation to model more realistic migration routes between populations.

Incremental simulations of the emergence of grammar: Towards complex sentence-meaning mappings

Andrei POPESCU-BELIS & John BATALI

University of California at San Diego
Department of Cognitive Science
9500 Gilman Drive, La Jolla, CA 92093-0515
{apopescu, batali}@cogsci.ucsd.edu

Introduction

Experiments with societies of communicating agents have shown that various communication conventions can emerge in order to express the structure of situations in an environment (e.g., Batali 1998, Steels 1997). However, it is often unclear how much implicit knowledge is initially given to the agents, or may come from the way meaning itself is encoded. In this study, we analyze two experimental models from the point of view of built-in knowledge and emergent capacities. The first one proves the emergence of syntax-specific capacities in agents that initially possess only semantic knowledge; the second model incorporates part of these capacities as initial knowledge. This gives its agents similar capacities in a shorter time, thus opening the way to more complex semantic structures, in the conceptual graphs formalism.

We defend an incremental paradigm for building models: to simulate increasingly expressive communication codes, it is possible to avoid evolving them from scratch, and start at a level reached by previous experiments. If a certain level of complexity has been attained, then it may not be useful to go again through all levels in a further simulation, but only through the new ones. This paradigm provides gain in simulation time that may prove crucial, and a better control over the intermediate states.

Semantic representations

The experiments described here presuppose that the agents are able to represent situations in a conceptual form, their purpose being an agreement on mappings between these forms and messages (sequences of letters). One of the most basic features of language, though not intrinsic to communication codes, is the division of this mapping procedure in several stages. Thus, messages are segmented into lexical units; most of these have a proper meaning (lexical semantics); and their composition yields a complex, non-additive meaning (propositional semantics). A computational simulation of the emergence of language has to account for these levels – and possibly also for other related phenomena (discourse, pragmatics, conceptual system, etc).

The first experiment (Batali, *in press*) uses formula sets, i.e. conjunctions of <feature, referent(s)> formulae. The referents are numbers designating the participants in the situation. The unary features represent characteristics of a referent, while the binary features represent relations between two referents. For instance, the formula set *{(goose 1) (sang 1) (noticed 1 2) (snake 2)}* can be glossed as "A goose that sang noticed a snake". This formalism is equivalent to a small subset of the first order predicate logic. It also provides a straightforward representation of the referents and could be linked to an agent's perceptual device (cf. experiment of Steels and Kaplan, *in press*). However, it has to be extended to represent more complex semantic aspects, as well as referring status (the previous example could also be glossed as "*The* goose that sang noticed *the* snake").

In an attempt to use a better known and more expandable representation, the second experiment uses conceptual graphs (CG), i.e. concept nodes and relations between them (Sowa 1992). Three relation types match the argument slots of the formulae. Hence, the previous situation is represented as:

[NOTICED] → (AGENT) → [GOOSE: #1] → (ATTRIBUTE) → [SINGING]
→ (PATIENT) → [SNAKE: #2]

For now, we use simplified CGs without the referent numbers: [GOOSE] stands for [GOOSE: #1], "the goose". These CGs are equivalent to formula sets. The full CGs, to be used later, can represent the referring status using notations in the concept nodes (for "a snake", "the snake", "some snakes", "three snakes", "all snakes").

Emergence of fundamental syntactic properties

In the first experiment, the agents are given situations (formula sets), and for each dialog the sender, chosen at random, has to produce a sequence of letters representing the situation. Therefore, it either creates a new *exemplar* (meaning-to-string mapping) or uses combination and/or substitution on the existing ones. If the receiver is in learning mode, it uses both the situation and the message to update its own exemplar set. In trial mode, it has to infer the situation described by the received message. If it succeeds, the overall communicative accuracy or success increases.

The built-in *parsing/generation mechanism* makes an agent capable of replacing part of an exemplar with another exemplar (substitution). Thus, strings and situations get gradually broken up in their irreducible constituents. Conversely, exemplars are also put together to fit new situations (combination). However, without a built-in *learning device*, an agent would only store unstructured mappings from meanings to strings. This retribution device makes the creation of such exemplars more costly than the reuse of elementary exemplars through substitutions and combinations. Only frequently used exemplars are reinforced, provided they lead to a correct interpretation.

Incremental design: Use of the emergent grammars as starting points

The first experiment starts with a "complexive" use of language: holistic mappings from strings to situations, without intermediate levels. Its results prove that in each population, a set of exemplars acting as grammatical rules emerges, each one corresponding to a situation model. In general, their internal structure is partitioned: the formulae are grouped according to the referents. Conversely, there is also an agreement on a lexicon, or strings associated to singleton formulae, which are substituted on the complex exemplars. Sometimes, strings playing only grammatical roles appear.

Any of the states emerging in low-knowledge conditions may be used at a starting point for new experiments. They authorize us to consider separately the study of lexical and of propositional semantics. Furthermore, word segmentation and word understanding can also be separated (figure 1). This built-in knowledge allows for faster convergence, easier implementation, and use of more complex semantic descriptions.

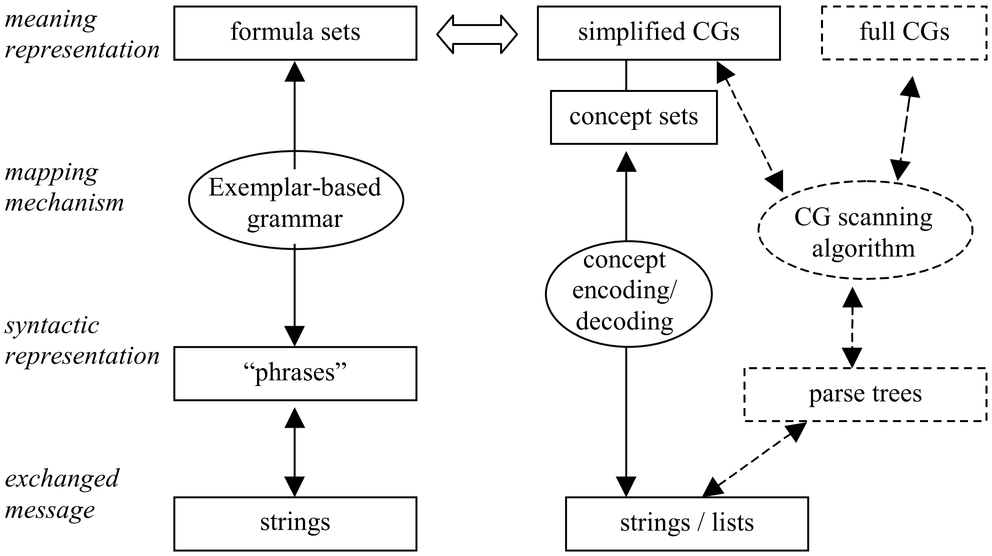


Figure 1. *Data structures and conversion mechanisms in the two experiments.*

Emergence of word-concept associations

In a first series of trials, we focused on lexical semantics conventions: the agents had to map a message to an unstructured set of concepts constituting the situation. General agreement ("convergence") on words has been observed in a variety of situations, using the same dialog protocol as in the first experiment. An agent's *conservatism* (in [0, 1]) is one of the main parameters governing convergence time. Other parameters include the number of words an agent is allowed to guess in a learning dialog (at least one), and the size of the situations that the agents observe (e.g., random size between one and five concepts). Their effects are undergoing extensive theoretical and experimental study. Here, figure 2 shows the difference in average convergence time between populations that use pre-segmented vs. concatenated messages. Both options converge, but the first is faster, and the difference increases with population size.

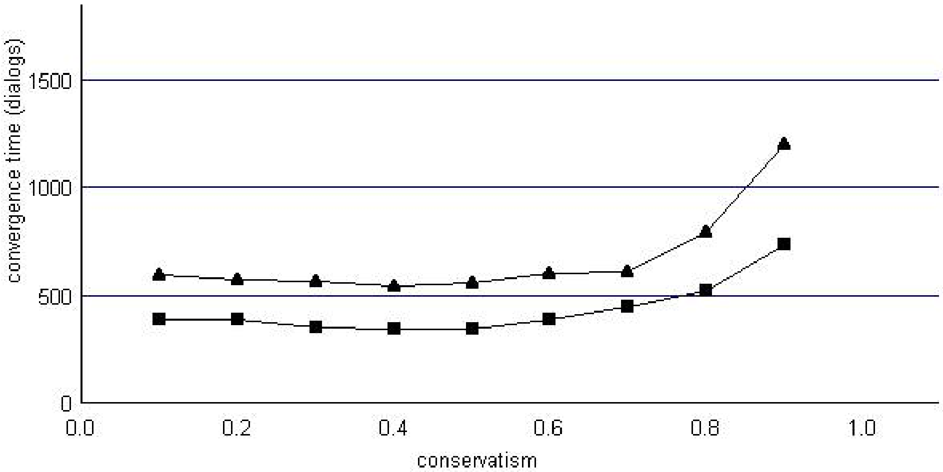


Figure 2. *Average convergence times depend on the agents' lexical conservatism. The lower curve is for separate words, the upper for concatenated ones (15 concepts, 5 agents, 1 to 4 concept situations, agents allowed to guess 1 or 2 words).*

Towards the emergence of syntactic conventions

Once words are mapped to concepts, the structure of situations is represented using a lexicalized tree grammar adapted to the simplified CGs (Popescu-Belis 1999, Allexandre and Popescu-Belis 1998). Each concept has an associated elementary tree – branch order being a syntactic parameter – and these are combined using substitutions and adjunctions. Despite a huge number of parameter combinations, this TAG-like grammar allows for much less variation than the exemplar-based one.

To understand a message, the agents use a two-phase inference mechanism. The receiver first maps the words to the situation's concepts, then performs a trial and error comparison between the message it could generate for these concepts. These modules are under implementation, but based on previous experiments, we expect an agreement on branch orders. Of particular interest are the measures of the convergence rate as a function of the number of concepts, and the comparison with the first experiment.

Conclusion

The grammars of the two experiments both fulfill analog tasks in individual agents (figure 1). However, while the first experiment shows how grammatical conventions emerge in low-knowledge populations, the second makes use of this experimental proof to seek faster convergence and a more open conceptual formalism. An exemplar-based grammar for conceptual graphs is possible, but would be long to converge, and its final shape could not be well controlled. On the contrary, faster convergence and a more constrained form allow us to consider environments that are more complex.

References

Allexandre C. & Popescu-Belis A. (1998) – "Emergence of Grammatical Conventions in an Agent Population Using a Simplified Tree Adjoining Grammar", *Proceedings of ICMAS'98*, Paris, p.383-384.

Batali J. (1998) – "Computational Simulations of the Emergence of Grammar", in Hurford J.R., Studdert-Kennedy M. & Knight C. (eds.), *Approaches to the Evolution of Language*, Cambridge, p.405-426.

Batali J. (*in press*) – "The Negotiation and Acquisition of Recursive Grammars as a Result of Competition Among Exemplars", in Briscoe T. (ed.), *Linguistic Evolution Through Language Acquisition: Formal and Computational Models*, Cambridge.

Popescu-Belis A. (1999) – *Modélisation multi-agent des échanges langagiers : application au problème de la référence et son évaluation*, Ph.D. thesis, Université de Paris XI.

Sowa J. (1992) – "Conceptual Graphs Summary", in *Conceptual Structures: Current Research and Practice*, Melksham.

Steels L. (1997) – "The Origins of Syntax in Visually Grounded Robotic Agents", *Proceedings of IJCAI'97*, Nagoya, Japan, vol. 2/2, p.1632-1641.

Steels L. & Kaplan F. (in press) – "Bootstrapping Grounded Word Semantics", in Briscoe T. (ed.), *Linguistic Evolution Through Language Acquisition: Formal and Computational Models*, Cambridge.

Primate language experiments: Instruction, translation and the evolutionary origins of language

Gregory RADICK

Department of History and Philosophy of Science
University of Cambridge
gmr22@hermes.cam.ac.uk

ABSTRACT

This paper will deal with some of the historical and conceptual issues surrounding primate language experiments. I shall focus in particular on a contrast between two different sorts of experiment: *instructional experiments*, which explore the abilities of primates to learn human-like language; and *translational experiments*, which explore the abilities of primates to communicate with one another in the wild. In the twentieth century, instructional experiments dominated, with little headway made in translational experiments until the 1970s. Why this asymmetry? There are two obvious explanations: first, that it was just easier to instruct in laboratories than to translate in the field; and second, that the recording technology needed for field translation was not available for a long time. I shall argue that the largely forgotten career of the pioneer primate language researcher R. L. Garner (1848–1920) shows both explanations to be inadequate. Garner in the 1890s was using the Edison cylinder phonograph for purposes of translation. On the basis of these experiments he claimed to discover a "simian tongue", different only in degree from human tongues. Furthermore, in 1892 he travelled to the French Congo to carry out phonographic experiments among wild gorillas and chimpanzees; and it was while there that he first attempted to teach a young chimpanzee to speak. Both sorts of primate language research thus began in their modern forms at roughly the same moment, receiving a great deal of popular and scientific attention. Both involved considerable difficulties. So why did instruction flourish and translation fizzle? And what does the dominance of instruction reveal about primate language research and its significance in the twentieth century? In offering some preliminary answers to these questions, I shall be arguing more generally for the value of a clearer view of the historical origins of the commitments – theoretical, methodological, institutional – that shape present debate on language and evolution.

Constraints on communities with indigenous sign languages: Clues to the dynamics of language origins

Sonia Ragir

Associate Professor of Anthropology
College of Staten Island, CUNY
RAGIR@ulster.net

Creole and signed languages inevitably appear when some critical number of language-ready children and/or adults engages in social activities over an extended period in schools, on plantations, on whaling ships, and in the African commercial centers for the slave-trade. New language-speakers utilize semantic content drawn from the local content and form of activities around them. Current descriptions of the growth of signed languages in deaf communities suggest that the emergence of new languages, both signed and spoken, is a function of group-cohesion (Kegl et al. 1998; McWhorter 1997). The emergence of a Creole from a plantation or trade Pidgin would appear to be quite different from the emergence of language in early human populations, where only a minority of individuals might have been “language-ready.” However, the indigenous signed systems among the deaf in hearing-speaking populations are natural experiments in language genesis, and they provide important insights into the constraints on forming a body of linguistic practices before languages became ubiquitous in hominid populations.

If new languages can form in deaf-signing communities, then we might expect the same capacity to exist in hearing-speaking communities. The fact that grammars unfold and acquire formal marking as new members of a community learn a sign system suggests that self-organizing properties of aggregate activities may be at least as important for the evolution of language as the innate properties of mind (Batali 1998). Steele (1998) hypothesized that, as in other biological communities, cooperative interaction in hominid groups was self-organizing. Computer simulations confirmed that, goal-oriented communicative behavior increased in its shared content and complexity over time, and that hierarchically embedded levels of semantic structure emerged spontaneously. Both Steele (1998) and Batali (1998) suggested that an innate language-specific structure or acquisition devise was not necessary to account for the origins of syntax, rather syntax originated spontaneously – an embedded heirarchy of compositional levels – based on the pressure to convey more information with limited resources of time, memory and processing power.

Deacon (1997) argued that:

> The source of information that is used to “grow” a language lies neither in the corpus of texts and corrections presented to the child, nor in the child’s brain to begin with. It is highly distributed across myriad interactions between children’s learning and the evolution of the language community.... The mechanisms driving language change at the socio-cultural level are also responsible for everyday language learning. (Ibid.:115)

These self-organizing properties of language and mind suggest rapid episodes of language emergence in language-ready populations rather than the gradual selection for language.

Perhaps the best known historical example of an indigenous sign language community is that of Martha’s Vineyard, whose members contributed substantially to the formation of American Sign Language during the 19th century (Groce 1985). Because a congenital deafness which affected roughly 4% of the total population was concentrated in isolated upland farming settlements, every large upland family contained one or two deaf members. No good description of the sign language was made prior to its virtual disappearance at the end of the 19th century. However, it was used by both hearing and deaf inhabitants of Martha's Vineyard for at least two hundred years, and the social context in which it developed and flourished has been carefully reconstructed. No economic, political, or social prejudices separated deaf from hearing citizens on Martha's Vineyard; the majority of the deaf married, and the number of deaf-deaf marriages was substantially lower than in New England. Intermarriage among the islanders kept the proportion of deaf in the population high, and for over two hundred years, the three upland farming communities, each consisting of 100-300 hundred persons, included at least ten deaf and many hearing children all learning sign from a prior generation of native signers.

Unlike the 18th and 19th century community on Martha's Vineyard, there were no indications of indigenous sign languages or congenital deafness in Kegl et al.'s (1998; Kegl & McWhorter 1997) survey of the northwest province of Nicaragua. Prior to 1980, when the Sandanista government established the first schools for the education of the deaf, most of the deaf children and adults were linguistically isolated within their own families and used home signs. The home signs were more discrete and conventional than mimetic gesture but tended not to combine multiple discrete signs in a relational way that resembled grammatical structures. Only a small enclave of four deaf adult siblings had developed what Kegl and Iwata (1989) described as a "signed pidgin." The shape and meaning of their signs were stable and combined in telegraphic phrases with little internal structure; the enclave was very isolated and the system static. Schaller and Sacks (Sacks 1999) documented a group of five deaf Mexican siblings who had developed a similar elaboration of home sign that they used to communicate with their Spanish-speaking father and each other. Even after two decades of use, the siblings had not developed a stable phrase structure, nor had they systematized the meanings of many of their gestures. These studies have demonstrated neither grammatical structure nor a stable lexicon will spontaneously emerge in the absence of the *right kind of social setting*, but that within an appropriate setting the input to individuals need not be a fully formed language to begin language genesis.

Based on her experience in Nicaragua, Kegl (1998) proposed that a critical number of about ten children was required to generate a language. I should like to add the hypothesis that a 'language-ready' community needs not only a critical mass but also historical continuity in order to generate a language. McWhorter (1997) documented the growth of semantic marking of grammatical categories in trade languages or "pidgins" in well-established historical trade systems, even though no children contributed to the formation of these "languages." His evidence directly contradicts Bickerton's (1990, 1998) hypothesis that spontaneous grammatical structuring can only appear with the first generation of children to learn the pidgin as a native language. McWhorter's careful historical documentation of the expansion of grammatical structures in Pidgins – in the centers for shipping slaves to the new world, on plantations or in mines using migrant male workers, and on whaling ships – gives us the best insight into the self-organizing properties of language from the perspective of a community rather than that of an individual.

Senghas (1995, 1997) described the development of grammatical structures in Nicaraguan Sign Language that emerged over a period of twenty years. After a few years, while the signed peer-group jargon of the school yard was still the major source of semantic and grammatical information, each new class learned to sign from the children who were already familiar with the school community's systematic, but limited, ordering of semantic roles and conventionalized signs. Over a decade of school admissions, the language pool became lexically enriched; grammatical marking of verb arguments, spacial inflection, size and shape specifiers, and person agreement began to appear in young (under 6.6 years) and medium aged signers (6.7-10 years) who entered during the second decade of the school (Kegl et al. 1999:221). Year after year, the youngest signers streamlined gesture-shape and movement, and formalized emergent sets of grammatical markers, while the older children appeared to contribute much of the new lexicon (Senghas 1995). The most proficient signers were the school community's youngest and *newest* members. The founding cohort and individuals older than ten when they first entered the school never attained the proficiency of these young signers.

A review of earlier studies of signed pidgins in Papua New Guinea (Kendon 1980abc), on Providence Island (Washabaugh 1980ab, 1986), and in Guatamala (Shuman 1980ab) confirms the hypothesis that lexical and grammatical complexity increases only in an open language pool – with the entrance of new cohorts of signers. The single cohort signed pidgins never developed a grammar, although some of the practitioners attained stable and fluent gesture-shapes and movements. In all cases of limited pidgins used between hearing and deaf signers, the number of signers was relatively small, few if any new members entered the community, and the sign system disappeared after the death of the last deaf-signer. These group-constraints on the emergence of indigenous sign languages – the contextual variables that affect whether or not a community develops a formal system with grammar – have profound implications for various theories about language origins.

Rather than evolve, languages and grammars appear self-organizing within a community of language-ready individuals. I propose that there are at least two interacting levels: (1) the self-organization of the central and peripheral nervous systems during ontogeny in response to the way that the child participates in a community (Deacon 1997; Edelman 1987; Elman et al. 1996; Karmiloff-Smith 1992; Kauffman 1993; Oyama 1985), and (2) the self-organization of the exchange of information constrained by the medium and intentional content of the representational system (Bates & Goodman 1999; Bates & Elman 1996; Chomsky 1988; Hemelrijk 1999; Parrish & Edelstein-Keshet 1999; Steels 1998). The comparison between indigenous sign language communities in which the number of speakers and the lack of historical depth constrain semantic and grammatical complexity can give us insight into the formation of linguistic communities and the dynamics of language emergence in prehistory.

Of course the deaf are easily identified in a speaking community, which makes their association with other deaf individuals relatively straightforward. As far as we know, language-readiness in a population without language does not provide language-ready children or adults with identifiable characteristics, and a critical mass would have to be achieved more or less by chance. How are groups of 50 to 150 hominids – the most likely size of hominid hunter-gather bands (Dunbar 1996, 1998) – to secure the interaction of a critical mass of "language-ready" individuals necessary for generating standardized, peer-group jargons or pidgins that can become grammatically structured as new individuals entered into the pool of language-users. The limited linguistic systematicity found in closed-groups of signers that were smaller than the presumed "critical mass" (Kegle 1998; Shaller and Sacks 1999) suggests that open, dynamic systems of communicative practices are fundamental to the sign languages, and that a language-ready brain is likely to be essential for both pidgin and language genesis. If language can be generated only after a dynamic community of language-ready individuals exists, then the brain cannot evolve language capacity through the *selection for* individuals with grammatically complex speech or *for* the successive stages of grammatical complexity differentiating linguistic practices of groups (Batali 1998).

Three assumptions often underlie discussions of language origins: (1) that natural selection for fundamental cognitive functions and for language shaped the evolution of language; (2) that the syntax and semantics of the earliest languages differed from modern languages because of species-level differences in brain structure or organization; and (3) that language emerged once, like a species, and then diverged from this original primitive state. There are strong indications that these three assumptions are not true. The brain and languages appear to be self-organizing, and new languages grow in semantic and syntactic complexity so long as the interacting pool of speakers can attain a critical mass and remain open to new members either through births or adolescent dispersal. Although general encephalization and global increases in cortical connectivity occurred during human evolution, neurologists and psychologists have found no unique, language-specific cortical sub-structures to account for language syntax or semantics (Calvin 1996; Donald 1998; Jerison 1995; Edelman 1987). Furthermore, since deaf communities lacking linguistic models have repeatedly organized new sign languages, there is reason to believe that large-brained hominid communities could have generated vocal languages independent of each other.

Finally, studies suggest that language formation takes place over a historical rather an evolutionary time span, and that ontogenetic development, rather than phylogenetic selection, is the central formative process affecting the neurological specialization underlying language. A common criticism of language origin hypotheses has been that they are too speculative – that there is no way of disproving a hypothesis by observing the proposed processes of language formation in nature, in the laboratory, and in archaeological and fossils remains. However, excellent descriptions of communication systems among linguistically isolated deaf individuals in speaking communities exist. When we understand why signed languages emerge spontaneously over historical time in some communities with deaf individuals and not in others, then we can begin to understand the constraints on language genesis in hominid evolution.

References

Batali J. 1998. Computational simulations of the emergence of grammar. In J.R.Hurford, M. Studdert-Kennedy, C. Knight (eds), Approaches to the Evolution of Language, pp.3405-426. Cambridge: Cambridge University Press.

Bates E. & Elman J.L. 1996. Learning rediscovered. Science 274:1849-1850.

Bates E. & Goodman J. C. 1999. On the emergence of grammar from the lexicon. In B. MacWhinney (ed), The emergence of language, pp.29-80. Mahwah NJ: Lawrence Erlbaum.

Bickerton D. 1990. Language and Species. Chicago: University of Chicago Press.

Bickerton D. 1998. Catastrophic evolution: the case for a single step from protolanguage to full human language. In J.R.Hurford, M. Studdert-Kennedy, C. Knight (eds), Approaches to the Evolution of Language, pp.341-358. Cambridge: Cambridge University Press.

Calvin W.H. 1996b. The Cerebral Code. Cambridge: MIT Press.

Chomsky N. 1988. Language and the problems of knowledge: the Managua lectures. Cambridge, Mass.: MIT press.

Deacon T.W. 1997. The symbolic species: The co-evolution of language and the brain. New York: Norton.

Donald M. 1998. Mimesis and the executive suite: missing links in language evolution. In J Hurford, M. Studdert-Kennedy, C Knight (eds), Approaches to the Evolution of Language, pp.44-67. Cambridge: Cambridge University Press.

Dunbar R. 1996. Grooming, Gossip and the Evolution of Language. London: Farber & Farber.

Dunbar R. 1998. The social brain hypothesis. Evolutionary Anthropology 6(5):178-190.

Edelman G. 1987. Neural Darwinism: The Theory of Neuronal Group Selection. New York: Basic Books.

Elman J.L., Bates E.A., Johnson M.H., Karmiloff-Smith A., Parisi D. & Plunkett K. 1996. Rethinking Innateness: A Connectionist Perspective on Development. Cambridge: MIT Press.

Groce, N.E. 1985. Everyone here spoke sign language: Hereditary deafness on Martha's Vineyard. Cambridge MA: Harvard University Press.

Hemelrijk C.K. 1999. An individual-orientated model of the emergence of despotic and egalitarian societies. Proc. R. Soc. Lond. B 266:361-369.

Jerison H. 1995. Evolution of prefrontal cortex. In N. A. Krasnegor, G. R. Lyon, P. S. Goldman-Rakic (eds), Development of the Prefrontal Cortex: Evolution, Neurobiology, and Behavior, pp.9-26. Baltimore: Paul H. Brooks.

Karmiloff-Smith A. 1992. Beyond Modularity: A Developmental Perspective on Cognitive Science. Cambridge: MIT Press.

Kauffman S.A. 1993. The Origins of Order: Self-Organization and Selection in Evolution. New York: Oxford University Press.

Kegl J. 1998. Language emergence in a language-ready brain. Presented at the 2nd International Conference on The Evolution of Language. April 6-9, 1998, The University of East London, U. K.

Kegl J. & Iwata G. 1989. Lenguaje de Signos Nicaraguense: A pidgin sheds light on the "Creole?" ASL. In R. Carlson, S. DeLancey, S. Gildea, D. Payne, and A. Saxena (eds), Proceedings of the Fourth Meetings of the Pacific Linguistics Conference, pp. 266-294. Eugene, Oregon: Department of Linguistics, University of Oregon.

Kegl J. & McWhortner J. 1997. Perspectives on an emerging language. In E. Clark (ed), Proceedings of the Stanford Child Language Research Form, pp. 15-36. Palo Alto: Center for the Study of Language and Information.

Kegl J., Senghas A., Coppola M 1999. Creation through contact: Sign language emergence and sign language change in Nicaragua. In M. DeGraff (ed), Comparative Grammatical Change: The Intersection of Language Acquisistion, Creole Genesis, and Diachronic Syntax, pp.179-237. Cambridge, MA: MIT Press.

Kendon, A. 1980a. A description of a deaf-mute sign language from the Enga Provice of Papua New Guinea with some comparative discussion. Part I: the formational properties of Enga signs. Semiotica 31(1/2): 1-34.

Kendon A. 1980b. The description of a deaf-mute sign language from the Enga Province of Papua New Guinea with some comparative discussion. Part II: the semiotic functioning of Enga signs. Semiotica 32 (1/2):81-117,

Kendon A. 1980c. A description of a deaf-mute sign langauge from the Enga Provicne of Papua NewGuinea with some comparative discussion. Part III: aspects of utterance construction. Semiotica 32 (3/4):245-313.

Koch C. & Laurent G. 1999. Complexity and the nervous system. Science 284:96-98.

McWhorter J.H. 1997. Toward a new model of Creole genesis. New York: Peter Lang.

Oyama S. 1985. The Ontogeny of Information: Developmental Systems and Evolution. Cambridge: Cambridge University Press.

Parrish JK & Edelstein-Keshet L. 1999. Complexity, pattern, and evolutionary trade-offs in animal aggregation. Science 284:99-101.

Sacks O. (1999) "In search of Lucy Doe." Oliver Sacks Mind Traveler video series. Presented at the American Museum of Natural History, May 17, 1999.

Senghas A. 1997. Do children generate and change language? A new signed langage emerges at a Nicaraguan school for deaf children. A paper read at the Perception and Cognition Neuropsychology Seminar, New York University, February 6, 1997.

Shuman M. K. 1980a. Culture and deafness in a Mayan Indian Village. Psychiatry: Journal for the study of interpersonal processes 43(4):359-370.

Shuman M.K. 1980b A preliminary account of sign language use by the deaf in a Yucatec Maya village. Language Sciences 2:144-73.

Steels L. 1998. Synthesizing the origins of language and meaning using coevolution, self-organization and level formation. In J.R.Hurford, M. Studdert-Kennedy, C. Knight (eds), Approaches to the Evolution of Language, pp.384-404. Cambridge: Cambridge University Press.

Washabaugh W. 1980a. The organization and use of Providence Island sign language. Sign Language Studies 26: 65-92.

Washabaugh, W. 1980b. The Manu-facturing of a language. Semiotica 29 (1/2): 1-37.

Washabaugh W. 1986 Five Fingers for Survival. Ann Arbor: Karoma Publishers Inc.

From calls to words: Bridging the divide.

Eric A Salzen

Psychology Department, Aberdeen University
Aberdeen, Scotland.
e.salzen@abdn.ac.uk

Speculations on the evolution of speech have rejected an origin from animal calls and vocalizations because of an apparently unbridgeable divide between *Vocal signals* which are *innate, conative, emotional, involuntary, & high intensity* and *Speech symbols* which are *learned, cognitive, rational, voluntary, & low intensity*. Furthermore the *Neural systems* of emotional *calls* are *sub-cortical, limbic* and *right hemisphere* dominant whereas those of *speech* are *cortical* and *left hemisphere* dominant. This divide can be bridged in *Phylogeny, Ontogeny* and in *Neurology* using existing studies that seem to have been overlooked in the controversy on the animal origins of speech.

In the case of *phylogeny*, an ethological analysis of social signals in animals (Morris, 1956) shows them to be patterns of intention movements due to frustration and conflict (thwarting). These movements and the vocalizations produced by these movements have referential significance because of their nature, orientation and releasing stimuli. Similar emotional conflict and thwarting signalling in humans induces co-operative social responses that may end the thwarting (Salzen, 1991). The evolutionary process of "*ritualisation*" (Blest, 1961) involving "*typical intensity* " and "*typical form* " (Morris, 1957) accounts for animal signals and calls with repeated elements and constant form. Such calls are comparable with the *consonant/vowel* or *c/v/c* or *c/v/c/v* of the first utterances and words of babies (cf. Lewis, 1936). According to some studies of comparative philology (Diamond, 1959; Swadesh, 1971) the first words of early man may have been of the same form and were verbs used as imperatives to get other people to do things (Diamond, 1959; Revesz, 1956). Words signifying the use of an implement or tool would be *verb-nouns* and lead to a naming lexicon. As in the case of animal signals where repetition can indicate intensity, in some modern languages repetition of the *cv* elements is indeed used to indicate plurality, intensity, size, and adverbs (Farb, 1974). In this way further variation of the elements can extend the lexicon and allow the development of a syntax.

In the case of *ontogeny*, a rarely cited study (Lewis, 1936) preempted the ethological use of intention movements in signalling by showing how the first speech sounds of infants arise from anticipatory and residual feeding/sucking actions in hunger and satiation. It showed how these sounds are single or duplicated labials or labial-dentals with "a" vowels i.e. *cvcv* and form the majority of infant first words in European languages. Similar words occur in many other languages (Wundt, 1924) and have similar meanings (cf. Clark & Clark, 1977). These first words have whole sentence meanings (holophrases) such as "*do x with y* " (cf. Greenfield & Smith, 1976) and they also show the syllabic repetition essential for the development of a lexicon. The same infant study (Lewis, 1936) also suggested that these first speech sounds give rise to *babbling* because they reinstate their associated motivational and motor states which then reinforce the emitted behaviour. In this way babbling provides the bridge between *involuntary* and *voluntary* vocalizations. Babbling becomes involved in first words when adults pick up and use the sounds made in babbling by speaking in "*motherese* " with simplified consonants and "*cvcv* " repetitions (*mama, papa*). Subsequently when the infant makes these particular babbles again it is reinstating the parental reinforcement sounds and comfort states. This provides a sensori-motor rehearsal that facilitates the developing voluntary motor control of the speech apparatus and verbal *mimicry*, which is undoubtedly a major means of speech and language learning.

In the case of *neurology*, another less well-known work (Jurgens, 1992) describes a *brainstem articulatory coordination* system, a *forebrain motor control* system, and a *limbic motivational initiation* system for phonation in primates. Humans, unlike other primates, appear to have a *direct connection* between the *primary motor cortex* and the *brainstem nucleus ambiguus* which contains the laryngeal motoneurons (Kuypers, 1958), and this may be how humans obtain exquisite *voluntary control* of articulation. Control of sequential articulations and hence of syntactic language develops along with voluntary control of sequential arm-hand movements during reaching and grasping in the neonate (cf. Young, 1977). In both cases this occurs primarily in the *left hemisphere*, which has long been implicated in *sequential single stimulus-response processing* (cf. Levy, 1969) and in the *lateralisation* of both manipulation and speech. If the present view that speech arose from the effects of thwarted action states on the autonomic nervous system is correct, then this system may still be involved in the *instigation* of speech. One might expect, therefore, that lesions of higher autonomic centres would affect the impulse to vocalize or speak rather than give specific speech defects. Human patients with parasagittal tumours near the anterior cingulate gyrus, a structure through which emotive motivational information enters the cortical system from the brainstem, may show poverty of speech. In one significant case (Damasio & Van Hoesen, 1983) the patient could repeat speech and with subsequent partial recovery reported that she had simply had no will or desire to speak. This is consistent with a thwarting theory of social signalling and its role in the origin and evolution of speech.

References

Blest,A.D. (1961) The concept of ritualisation. In Thorpe,W.H. & Zangwill,O.L. (Eds.) *Current problems in animal behaviour* p.102-124. London: Cambridge Univ. Press.

Clark,H.H. & Clark,E.V. (1977) *Psychology and language.* N.Y.: Harcourt Brace Jovanovich.

Damasio,A.R. & Van Hoesen. (1983) Emotional disturbances associated with focal lesions of the limbic frontal lobe. In Heilman,K.M. & Satz,P. (Eds.) *Neuropsychology of human emotion* p.85-110. N.Y.: Guilford Press.

Diamond,A.S. (1959) *The history and origin of language.* London: Methuen.

Farb,P. (1974) *Word play.* London: Jonathan Cape.

Greenfield,P.M. & Smith,J.H. (1976) *The structure of communication in early language development.* N.Y.: Academic Press.

Jurgens,U. (1992) On the neurobiology of vocal communication. In Papousek,H., Jurgens,U. & Papousek,M. (Eds.) *Nonverbal vocal communication: comparative and developmental approaches.* p.31-42. Cambridge: Cambridge Univ. Press.

Kuypers,H.G. (1958) Corticobulbar connections to the pons and lower brainstem in man. *Brain, 81,* 364-388.

Levy,J. (1969) Possible basis for lateral specialization of the human brain. *Nature,* Lond., *224,* 614-615.

Lewis,M.M. (1936) *Infant speech.* London: Kegan Paul, Trench, Trubner.

Morris,D. (1956) The feather postures of birds and the problem of the origin of social signals. *Behaviour, 9,* 75-113.

Morris,D. (1957) “Typical intensity” and its relation to the problem of ritualization. *Behaviour, 11,* 1-12.

Revesz,G. (1956) *The origins and prehistory of language.* London: Longmans.

Salzen,E.A. (1991) On the nature of emotion. *International Journal of Comparative Psychology, 5,* 47-88.

Swadesh,M. (1971) *The origin and diversification of language.* Chicago: Aldine Atherton.

Wundt,W. (1924) *Völkerpsychologie.* 4th Ed.

Young,G. (1977) Manual specialization in infancy: implications for lateralization of brain function. In Segalowitz,S.J. & Gruber,F.A. (Eds.) *Language development and neurological theory.* p.289-311. New York: Academic Press.

Language is not an instinct

Geoffrey SAMPSON

ABSTRACT

There is cultural as well as biological evolution; but the "evolution of language" is frequently discussed in terms that suggest that the complex forms of present-day languages are the outcome of a biological evolutionary process. This point of view has been powerfully advocated by Steven Pinker's popular and very influential book *The language instinct*, together with several other relatively popular expositions of linguistic nativism during the past decade. These works in turn rely heavily on arguments originally put forward by Noam Chomsky in the 1960s and 1970s in less popular but again very influential writings (though Pinker is more sophisticated than Chomsky about biological evolution, and quotes some additional categories of evidence for the claim that language is an instinct).

My *Educating Eve* (revised paperback edition, Cassell, 1999) urges that human language is in all important respects a purely cultural rather than biological phenomenon: there is no good reason to believe in a "language instinct" in Pinker's sense. I look at each strand of argumentation which appears to me to have been influential in winning converts to linguistic nativism since the 1960s, and show that in every case the argument either rests on factually false premises, or is logically fallacious (or, in some cases, both). Thus, the various defences of nativism put forward in Chomsky's writings can be classified as variants of seven distinct empirical arguments (together with a number of apparently harmless rhetorical moves which on closer examination involve question-begging). These writings inspired independent contemporary research by e.g. Philip Liberman, and B. Berlin & P. Kay, which appeared to reinforce Chomsky's case. But none of Chomsky's or his contemporaries' empirical arguments succeed. Writers of the 1990s have cited further empirical evidence, for instance Pinker discusses the family of alleged "language mutants" discovered by Myrna Gopnik, but these new categories of data also fail to support the consequences drawn from them by nativist writers.

My Paris paper will outline the nature of the case for linguistic nativism, and will illustrate its emptiness by refuting as large a sample of its various strands as can be covered in the time (alluding briefly to the fact that strands which have to be omitted from the oral presentation are also covered in the book). The multi-strandedness of the nativists' case is one of the features which has helped it acquire more converts than it deserves: someone who looks at any one of the arguments and notices something problematic about it is usually aware that this is only one of many arguments which have been put forward, so that the isolated problem may not seem a fatal flaw. But in reality every one of the arguments is fatally flawed.

Human upper airway muscles are uniquely specialized by containing slow tonic muscle, a distinct class of muscle that may have evolved for speech.

Ira SANDERS

Box 1189, The Grabscheid Voice Center
Mount Sinai School of Medicine
New York, NY 10029
I_Sanders@smtplink.mssm.edu

ABSTRACT

It is proposed that in comparison to other mammals, humans have added a highly specialized motor system to the upper airway muscles; that this system is capable of performing unique actions with great precision; and that this system has evolved for speech production.

Speech is perhaps the defining characteristic of the human species. At present the nature of the protolanguage that bridged the gap between humans and primates is unknown. Most investigators agree that the most important evolutionary changes occurred in the mechanism for producing speech rather then in perceiving it. It is well known that the human tongue and pharynx have evolved unique morphologies to allow for the articulation of speech. In contrast, the vocal folds, the actual site of sound production, are externally similar to those of other mammals and have been relatively ignored. Almost completely unstudied is the internal specializations within these muscular organs. Based on the study of nearly 100 human larynges and a smaller sample of other upper airway muscles we have found a type of muscle tissue called slow tonic. In humans, slow tonic muscle fibers (STMF) have only been found in the extra ocular muscles, the most precisely controlled of the skeletal muscles. Although extremely rare in mammalian muscles STMF are common in amphibians and birds, and most of what is known about their anatomy and physiology comes from studies in these species. One important aspect of STMF biomechanics is that they do not contract with a twitch like other skeletal muscle fibers, instead they shorten in a graded and precisely controlled manner. Therefore the neural control of STMF is fundamentally different from other muscles in the body. The muscle tissue containing STMF in the human upper airway has so many unusual specializations that it may be considered to be a fourth class of muscle, in addition to smooth, cardiac and skeletal muscle. The exact locations where STMF are concentrated suggest that their function is directly related to speech production. For example, in the vocal folds most STMF concentrate in an area just underneath the vibrating edge of the vocal fold. In this area the STMF appear to be arranged to control the shape and tension within the vibrating vocal fold, and by extension, the quality of the sounds produced. The same area in the vocal folds of other mammals contains only soft tissue, including those of non-human primates. Very recently we have also found STMF to be present in the human tongue and pharynx. In these regions it is concentrated just beneath the mucosa, suggesting that it participates in shaping the pharyngeal walls and dorsal surface of the tongue, critical actions seen during speech articulation.

This presentation will review comparative aspects of vocalization in rats and bats (two species that use both audible and ultrasonic vocalization), carnivores (dogs and cats), and non-human primates. Then the recent anatomical studies of the human larynx, tongue and pharynx will be presented. The significance of the unique human anatomy for the biomechanics of vocal fold vibration, sound production, and speech articulation will be discussed.

Supported by grant DC 01764 from the National Institute on Deafness and Other Communication Disorders

Evolution of grammar: The grammaticization of nicaraguan sign language by sequential iterations of native acquisition

Ann SENGHAS

Department of Psychology, Barnard College
as1038@columbia.edu

ABSTRACT

Early in life, children are more adept at learning language than they will be as adults. This ability is often ascribed to biologically-endowed language-learning capacities available only during an early sensitive period. Such a period would presumably have evolved to be of sufficient duration to enable children to acquire language efficiently in a normal social environment that includes a fully-developed natural language.

In such an environment the grammar acquired closely resembles that of the input language. In those rare situations in which a child's linguistic environment is impoverished, the language-learning capacities are confronted with a more demanding task. Here, it might be expected that the impoverished input would be augmented by the child's capacities to produce a richer output. However, no such child has generated a full language. Perhaps the time required to generate one exceeds the individual's sensitive period. If so, sequential cohorts of child learners might eventually converge on a fully-developed language.

The recent emergence of Nicaraguan Sign Language enables us to consider such a case. The cohort of deaf Nicaraguans entering school in the late 1970s began developing the first community-wide sign language known in Nicaragua. A second cohort of children, entering school ten years later, learned this system from their older peers. Augmentation of the language at that time is still evident in differences between the two cohorts today.

One difference between the cohorts is evident in their expression of semantic roles (Senghas, et al., 1997). The first cohort consistently uses a simple, systematic word order that effectively makes the agent/patient distinction. The second cohort has modified this system, often producing sentences in which semantic roles are no longer unambiguously derivable from the word order. However, certain spatial constructions may be emerging to fill the gap morphologically. In both production (Senghas et al., 1997) and comprehension (Senghas, 1999), the side to which a sign is produced was found to be a meaningful, contrastive feature for second-cohort signers only. This change means that in cases where the context provides minimal cues, second-cohort signers actually misinterpret the assignment of semantic roles in the sentences produced by first-cohort signers. Thus, the second cohort has reanalyzed spatial contrasts to form a stricter, narrower system in which certain referents are specified more precisely.

Additionally, these spatial mappings do not appear to be derived directly from those used to describe the physical locations of objects. The sign-space to physical-space mappings used by even the second cohort in locative descriptions is much more variable than that used for semantic roles. Only in the case of semantic roles has the younger cohort converged on a shared construction community-wide.

These findings indicate that Nicaraguan Sign Language has undergone significant grammatical developments as it has acquired a second cohort of native signers. Some of these changes, such as the rules concerning common word orders, involve a broadening, in which new constructions are allowed. Others, such as the emergence of spatial devices, involve increasing specificity. Because both kinds of processes play a role in language learning and language change, both are now evident in language genesis.

The singing origin theory of speech.

Dr. John R. Skoyles

6 Denning Road, London, NW3 1SU
skoyles@bigfoot.com

Primeval man, or rather some early progenitor of man, probably first used his voice in producing true musical cadences, that is in singing, Charles Darwin (1871, p. 133).

Language originated as play, and the organs of speech were first trained in this singing sport of idle hours, Otto Jespersen (1922, p. 433).

Introduction

Since the eighteenth century and people such as Rousseau, Diderot, Rameau and Condillac (Downing, 1995), song has been linked to the origins of speech. Increasingly, the idea has been proposed in the twentieth century (Jespersen, 1922; Marler, 1970; Geist, 1978; Richman, 1993; Vaneechoutte & Skoyles, 1998). Highly developed thoracic respiratory control underlies our ability on a single out breath to create multiple strings of vocalisations accurately timed and synchronised with complex vocal tract movements. This ability is notably absent in our close relatives such as chimps (Provine, 1996, p. 40). Thoracic breathing control has been shown by Ann MacLarnon and Gwen Hewitt (1999) to have originated between 1,500,000 and 100,000 years ago. Amongst untrained singers, the respiratory adjustments used in singing are similar to those used in normal or loud speaking (Hixon & collaborators, 1987, p. 361). Those exploring the evolution of breath control and related vocal changes link them to speech (for example: Kay, Cartmill & Balow, 1998; MacLarnon & Hewitt, 1999). However, these changes could have evolved first to enable singing, and only then by the addition of vocabulary and syntax become used for speech. Here I argue for this latter theory.

Song *vs*. speech

Song and speech reflect respectively two different types of communicable information: (1) identifiable repetitions such as rhythm, melody, stressing and intonation; and (2) syntactically structured sequences of symbolic tokens. Song information can be used to: (a) create and display identity; (b) synchronise relationships between singers and their listeners; and (c) provide a 'carrying' structure to enable the transmission of higher levels of information. Due to (a) and (b), song tends to be used to create and maintain social attachments such as pair-bonds (breeding birds often bond through duetting), and group-bonds (for example, in birds that engage in chorus singing), Moreover, due to (a), it can create a recognised link – 'ownership' – between a singer and a resource (such as a territory). These functions are primarily ‘limbic’ concerned with inducing behaviour and emotion in other animals.

Speech information comes from the capacity of symbols when combined syntactically to describe things and narrate events. Such information is referential rather than generative of a relationship between a speaker and their listeners (though what is described can at a referent level involve them). While what is described might cause behaviour and emotion, they are not caused directly by the symbols but the message they encode – speech is primarily ‘cortical’. Because they communicate different types of information, song and speech can easily be blended together as in poetry, chant and pray.

Parsimony

Logical arguments about precedence underlie important areas of science, for example, the conjecture that RNA life arose before life based upon proteins and DNA (Freeland, Knight & Landweber, 1999). Similar precedence arguments apply to why human song preceded speech. For humans to sing requires: (i) the capacity to produce and learn repetitive patterns, and (ii) the thoracic control of expirations to enable long sequences of different tones and articulations made upon a single out breath. However, speech requires at least two additional components: syntax and word vocabulary. The latter consists itself of two components: the ability to link semantics to word pronunciations (both in perception and production), and the ability to acquire words and their meaning from their presence in the talk of others. An asymmetry thus exists: while the components needed for song can independently precede those needed for speech (you can sing without words and syntax), those for speech cannot independently precede those needed for song (speech needs the breath control required for song). The evolution of biological structures, moreover, goes through stages whereby inherited modifications become increasingly complex by additions – feathers, for example, preceded their use in flight by initially being evolved to provide thermal insulation then only became structurally adapted (elongation etc) as wing feathers. Song has many functional advantages (see below), and is a form of communication that easily mixes with speech (chants, prayers, poetry). Thus, it is a natural proto-stage which could initially arise and then further develop by elaboration into speech.

Evolution

Song has evolved many times in diverse species including crickets, birds (on many independent occasions), sea-mammals (dolphins, porpoises and sperm whales), and in all monogamous primates with stable territories (indris, titis, tarsiers and gibbons) (Haimoff, 1986). In contrast, the use of words and syntax for symbolic-based communication has evolved only once. One reason for the evolution of song in diverse animals is bonding of groups (Bown, Farabaugh & Veltman, 1988) and breeding pairs (Diamond & Terborgh, 1968; Thorpe & North, 1966). Even the function of song to defend territory is arguable one of bonding – though in this case between an individual and a resource. Humans (ignoring vocal communication) are distinct from our closest apes, the chimpanzee, in three prominent respects that link to our capacity to bond. First, we maintain lifelong attachments with dispersed offspring – indeed we are the only primate that does this (Rodseth, Wrangham, Harrigan & Smuts, 1991). As a result, all humans exist within complex social networks built around group and kin attachments. Second, human parents bond either monogamously or polygamously. Third, we have grossly enlarged brains – the development of which was made possible by the resources provided by bonded parents. As much as song bonds diverse animals, it also bonds humans: for example, in rituals observed by anthropologists (Bowra, 1962; Blacking, 1973), and in such familiar activities as marching and work songs, football stadium chants, National Anthems, camp-fire songs and hymns. As Ellen Dissanayake (1992, p. 119) puts it: 'by means of music a supra individual state is created in which singer and listener can exist together, joined in a “common consciousness”, a common pattern of thought, attitude and emotion'. Thus, human evolution had in song a faculty: (i) that is not only used by many diverse animals (including all territorial monogamous primates) to create and maintain parental and social bonds, but (ii) that is used by modern people for related purposes, and (iii) that would have enabled the parental and complex social bonds needed for brain expansion. Moreover, if hominids sung to maintain social bonds, such vocalisations would have been available for natural selection to modify by the addition of words and syntax into speech.

Development

Abilities linked to song, precede and aid those of speech. Children treat vocalisations both of themselves and others as a kind of 'song' (Papousek & Papousek, 1981). Newborns identify prosody sufficiently well (due to hearing low-pass filtered maternal vocalisations in the womb) to recognise the language that surrounds them from foreign ones (Mehler *et al.*, 1988). 'Motherese' strongly emphases prosody (Trehub, Trainor & Unyk, 1993); in addition songs – lullabies – are an important part of prelinguistic communication (Trehub, Trainor & Unyk, 1993). Not surprisingly, babbling by eight month olds contains intonation patterns of the surrounding language (de Boysson-Bardies, Sagart & Durand, 1984). The perception of intonation is used (together with distribution regularities in speech sounds) to segment out word boundaries within speech and so enable words to be identified and acquired (Sansavini, 1997). Intonation similarly segments phrase boundaries and thus aids the acquisition of syntax (Morgan & Demuth, 1996). Thus, while song could evolve before speech, speech could not have on developmental grounds been acquired without the earlier existence of song.

Vocal tract evolution

In addition to breath control, human evolution adapted the vocal tract. Usually this is assumed to enable it to create the wide range of speech sounds found in human languages. An alternative explanation is that the vocal tract evolved to provide humans with a wide variety of musical sounds. Supporting this is the fact that intelligible speech can be produced using only a small part of the vocal tract: for example, Arandic languages of Central Australia use only two distinct vowels, both central non-high ones (Maddieson, 1998). This is enigmatic since natural selection would have adapted the vocal tract to produce no more than the minimum range of vocal sounds needed for intelligible speech. The most parsimonious explanation is that vocal tract was shaped by a factor other than speech (such as singing) that required a much more extensive range, and then was only secondarily adapted for speech. Interestingly, the ability to produce musical rhythm and initiate singing is left hemispheric like the ability to speak suggesting the use of overlapping motor control circuits in both (Borchgrevink, 1991).

While on their own, none of the above four areas of arguments is conclusive, together as a group they strongly support Darwin and Jespersen's proposal that in human evolution the capacity to sing preceded the capacity to speak. Indeed, as Darwin (1872, p 476) once observed, ‘It can hardly be supposed that a false theory could explain, in so satisfactory a manner … the several large classes of facts above specified.

References

Blacking, J. (1973). *How musical is man?* Seatle: University of Washington Press.

Borchgrevink, H. M. (1991). Prosody, musical rhythm, tone pitch and response initiation during amytal hemisphere anaesthesia. In *Music, language, speech and brain*, (eds. J. Sundberg, L. Nord, & R. Carlson) pp. 327-343, Macmillan.

Bowra, C. M. (1962). *Primitive song*. London: Nicolson & Weidenfeld.

Brown, E. D., Farabaugh, S. M. & Veltman, C. J. (1988). Song sharing in a group-living songbird, the Australian magpie. Part 1. Vocal sharing within and among social groups. *Behaviour, 104*, 1-28.

Darwin, C. (1871). *Descent of man*. London: Murray.

Darwin, C. (1872). *Origin of Species*. 6th. ed., London: Murray.

de Boysson-Bardies, B., Sagart, L. & Durand, C. (1984). Discernible differences in the babbling of infants according to target language. *Journal of Child Language, 11*, 1-15.

Diamond, J. M. & Terborgh, J. W. (1968). Dual singing by New Guinea birds. *Auk, 85*, 62-82.

Dissanayake, E. (1992). *Homo aestheticus*. New York: Free Press.

Dowing, T. A. (1995). *Music and the origins of language*. CUP.

Freeland, S. J., Knight, R. D. & Landweber, L. F. (1999). Do proteins predate DNA? *Science, 286*, 690-692.

Geist, V. (1978). *Life Strategies, Human Evolution, Environmental Design*. New York: Springer Verlag.

Haimoff, E. H. (1986). Convergence in the duetting of monogamous old world primates. *Journal of Human Evolution*, 15, 51-59.

Hixon, J. & collaborators. (1987). *Respiration function in speech and song*. London: Taylor and Francis.

Jespersen, O. (1922). *Language*. London: Allen.

Kay, R. F., Cartmill, M. & Balow, M. (1998). The hypoglossal canal and the origin of human vocal behavior. *Proceedings of the National Academy of Sciences, USA, 95*, 5417-5419.

MacLarnon AM & Hewitt GP (1999). The evolution of human speech: the role of enhanced breathing control. *American Journal of Physical Anthropology, 109*, 341-363

Maddieson, I. (1999). Vowel systems and language origins. Abstracts, Proceedings of the Evolution of Language, 2nd International Conference.

Marler, P. (1970). Birdsong and speech development. *American Scientist, 58*, 667-673.

Mehler, J., Jusczyk, P., Lambretz, G., Halsted, N., Bertoncini, J. & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition, 29*, 143-178.

Morgan, J. L. & Demuth, K. (1996). *Signal to syntax*. Mahwah, NJ. Erlbaum.

Papousek, M. & Papousek, H. (1981). Musical elements in he infant's vocalization. *Advances in Infancy, 1*, 163-218.

Provine, R. R. (1998). Laughter. *American Scientist, 84*, 38-45.

Richman, B. (1993). On the evolution of speech: singing as the middle term. *Current Anthropology, 34*, 721-722.

Sansavini, A. (1997). Neonatal perception of the rhythmical structure of speech. *Early Development and Parenting, 6*, 3-13.

Thorpe, W. H. & North, M. E. (1966). Vocal imitation in the tropical Bou-bou shrike as a means of establishing and maintaining social bonds. *Iris, 108*, 432-435.

Trehub, S. E., Trainor, L. J. & Unyk, A. M. (1993). Music and speech processing in the first year of life. *Advances in Child Development, 24*, 1-35.

Vaneechoutte, M. & Skoyles, J. R. (1998). The memetic origin of language: modern humans as musical primates. *Journal of Memetics, 2*. http://www.cpm.mmu.ac.uk/jom-emit/1998/vol2/vaneechoutte_m&skoyles_jr.html

Learners are losers: Natural selection and learning in the evolution of communication

Kenneth SMITH

Language Evolution and Computation Research Unit
Department of Theoretical and Applied Linguistics
University of Edinburgh
kenny@ling.ed.ac.uk

The notion of the innateness of some part of human language competence is central to the most influential linguistic theories of modern times (Chomsky, 1965, 1980, 1981, 1987). Pinker & Bloom (1990) outline what could be considered to be an orthodox explanation for the origins of this innate, language specific mental organ. They argue that the human language faculty has been shaped by natural selection, assisted by positive interactions between natural selection and learning such as the Baldwin effect (Baldwin, 1896; Hinton & Nowlan, 1987).

Computational simulations of the emergence of communication systems, including language-like syntactically structured communication systems, fall into three main groups: those which suggest natural selection alone is capable of developing and refining innate communication systems (eg Werner & Dyer, 1991; Oliphant, 1996; MacLennan & Burghardt, 1994; Ackley & Littman, 1994; Levin, 1995; Cangelosi & Parisi, 1996; Bullock, 1997; Werner & Todd, 1997; de Bourcier & Wheeler, 1997; Di Paolo, 1997; Noble, 1998), those which suggest that repeated learning interactions alone are capable of developing and refining entirely learned communication systems (eg Oliphant, in press; Kirby, in press a, in press b; Hurford, in press; Batali, 1998, in press; Hutchins & Hazelhurst, 1995; Steels & Vogt, 1997), and those which suggest that natural selection and learning interact to develop communication systems which are part innate and part learned (eg Batali, 1994; Briscoe, 1997; Kirby & Hurford, 1997).

This paper explores the interactions between natural selection and learning in the evolution of simple communication, or signalling, systems (Lewis, 1969). Such systems consist of a set of meaning-form pairs, where both the meanings to be communicated and the communicative forms used to communicate those meanings are unstructured. Optimal communication among a population of individuals using such a communication system requires that the entire population use mutually intelligible communication systems and that these communication systems make an unambiguous mapping between meanings and forms (ie the population's communication systems must be free of synonymy and homonymy).

A simulated population of agents using genetically encoded communication systems was evolved, with breeding based on ability to communicate with other members of the population. No learning took place in this simulation. As suggested by similar computational simulations (eg Werner & Dyer, 1991; Oliphant, 1996; MacLennan & Burghardt, 1994; Ackley & Littman, 1994; Levin, 1995; Cangelosi & Parisi, 1996; Bullock, 1997; Werner & Todd, 1997; de Bourcier & Wheeler, 1997; Di Paolo, 1997; Noble, 1998), an innate, optimal communication system rapidly emerged in the simulated population.

Cultural transmission was then added to this model. Mature individuals were selected for breeding according to communicative success, as before, but their offspring learn a communication system based on the communicative behaviour of the mature population. An individual's genes now function as a starting point for, and constraint on, learning, rather than fully determining an individual's communication system. Previous work on the interaction between natural selection and learning in the evolution of communication (eg Batali, 1994; Briscoe, 1997; Kirby & Hurford, 1997) suggests that, under these circumstances, communication systems which are part innate and part learned should emerge.

This was found not to be the case. Such a simulated population was incapable of developing an optimal communication system from random behaviour. Furthermore, such a population was incapable of maintaining an optimal communication system over time, even if this optimal communication system was both encoded in the genes of the population and learnable by observation.

The combination of selection for communicative success and learning was incapable of developing an optimal communication system due to an overly-plastic phenotype. Learning overrides the genetically encoded communication system of individuals in the population. This has two effects. Firstly, natural selection is effectively disabled - due to learning in the phenotype, there is no selectional pressure for genes encoding optimal communication systems. This phenomenon is known as shielding (Ackley & Littman, 1991). Secondly, suboptimal communication systems, which form the vast majority of the communication systems used by the initial random population, are preserved - learning alone is incapable of developing optimal communication systems.

The combination of selection for communicative success and learning was also incapable of preserving optimal communication systems, even if those communication systems were encoded in the genomes of all members of the initial population. This was due to shielding of genetic information and cultural transmission. The absence of selectional pressure on the population's genes allows those genes to drift through genetic space due to random mutations. Eventually, one individual will inherit genes which are so bad that those genes do influence the learning process, preventing that individual from acquiring the optimal communication system of the population. While this individual will be weeded out by natural selection, its communicative behaviour will be observed by some language learners. These language learners run the risk of failing to acquire the optimal communication system, due to the noise introduced by the suboptimal communicator. If these individuals learn a suboptimal communication system they will also be weeded out by natural selection, but not before their communicative behaviour influences the communicative behaviour of yet more learners. Cultural transmission allows suboptimal communication systems to spread through the population like a virus, until the whole population communicates using a suboptimal system.

These simulations suggest two conclusions on the relative importance of natural selection and cultural transmission in the evolution of the human language faculty. Firstly, extreme plasticity in the phenotype of a communicative agent may disable natural selection, leading to behaviour which is determined by the agent's learning apparatus, rather than a combination of natural selection and learning. Secondly, if optimal learned communication systems are to emerge in a population, the agents in that population must be capable of selective acquisition or production of communicative forms. This linguistic selection at the individual level drives linguistic evolution at the population level, pushing the communication system of the population towards optimal systems of the type seen in Oliphant (in press), Kirby (in press a, in press b), Hurford (in press), Batali (1998, in press), Hutchins & Hazelhurst (1995) and Steels & Vogt (1997).

References

Ackley, D. and M. Littman (1991). Interactions between learning and evolution. In C. Langton, C. Taylor, J. Farmer and S. Rasmussen (eds.), *Artificial Life 2,* pages 487-509. Redwood City, CA: Addison-Wesley

Ackley, D. and M. Littman (1994). Altruism in the evolution of communication. In R. Brooks and P. Maes (eds.), *Artificial Life 4: Proceedings of the Fourth International Workshop on the Synthesis and Simulation of Living Systems*, pages 40-48. Redwood City, CA: Addison-Wesley.

Baldwin, J.M. (1896). A new factor in evolution. *American Naturalist 30*, 441-451.

Batali, J. (1994). Innate biases and critical periods: Combining evolution and learning in the acquisition of syntax. In R. Brooks and P. Maes (eds.), *Artificial Life 4: Proceedings of the Fourth International Workshop on the Synthesis and Simulation of Living Systems*, pages 160-171. Redwood City, CA: Addison-Wesley.

Batali, J. (1998). Computational simulations of the emergence of grammar. In J.R. Hurford, M. Studdert-Kennedy and C. Knight (eds.), *Approaches to the Evolution of Language: Social and Cognitive Bases.* Cambridge: Cambridge University Press.

Batali, J. (in press). The negotiation and acquisition of recursive grammars as a result of competition among exemplars. In E. Briscoe (ed.), *Linguistic Evolution through Language Acquisition: Formal and Computational Models.* Cambridge: Cambridge University Press.

Briscoe, E. (1997). Language acquisition: the bioprogram hypothesis and the Baldwin Effect. MS, Computer Laboratory, University of Cambridge.

Bullock, S. (1997). An exploration of signalling behaviour by both analytic and simulation means for both discrete and continuous models. In P. Husbands and I. Harvey (eds.), *Fourth European Conference on Artificial Life*, pages 454-463. Cambridge, MA: MIT Press.

Cangelosi, A. and D. Parisi (1996). The emergence of a `language' in an evolving population of neural networks. Technical Report NSAL-96-004, Institute of Psychology, National Research Council, Rome.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, N. (1980). *Rules and Representations*. London: Basil Blackwell.

Chomsky, N. (1981). *Government and Binding*. Dordrecht: Foris.

Chomsky, N. (1987). *Knowledge of Language: Its Nature, Origin and Use*. Dordrecht: Foris.

de Bourcier, P. and M. Wheeler (1997). The truth is out there: The evolution of reliability in aggressive communication systems. In P. Husbands and I. Harvey (eds.), *Fourth European Conference on Artificial Life*, pages 444-453. Cambridge, MA: MIT Press.

Di Paolo, E. (1997). An investigation into the evolution of communication. *Adaptive Behaviour 6*, 285-324.

Hinton, G. and Nowlan, S. (1987). How learning can guide evolution. *Complex Systems 1*, 495-502.

Hurford, J. (in press). Expression/Induction Models of Language Evolution: Dimensions and Issues. In E. Briscoe (ed*.), Linguistic Evolution through Language Acquisition: Formal and Computational Models*. Cambridge: Cambridge University Press.

Hutchins, E. and B. Hazelhurst (1995). How to invent a lexicon: The development of shared symbols in interaction. In N. Gilbert and R. Conte (eds.), *Artificial Societies: the computer simulation of social life*. London: UCL Press.

Kirby, S. (in press a). Syntax without natural selection: How compositionality emerges from vocabulary in a population of learners. In C. Knight, M. Studdert-Kennedy and J. Hurford (eds*.), The Evolutionary Emergence of Language*. Cambridge: Cambridge University Press.

Kirby, S. (in press b). Learning, bottlenecks and the evolution of recursive syntax. To appear in E. Briscoe (ed.*), Linguistic Evolution through Language Acquisition: Formal and Computational Models*. Cambridge: Cambridge University Press.

Kirby, S. and J. Hurford (1997). Learning, culture and evolution in the origin of linguistic constraints. In P. Husbands and I. Harvey (eds.), *Fourth European Conference on Artificial Life*, pages 493-502. Cambridge, MA: MIT Press.

Levin, M. (1995). The evolution of understanding: a genetic algorithm model of the evolution of communication. *BioSystems 36*, 167-178.

Lewis, D. (1969). *Convention: A Philosophical Study*. Cambridge, MA: Harvard University Press.

MacLennan, B. and G. Burghardt (1994). Synthetic ethology and the evolution of cooperative communication. *Adaptive Behaviour 2*, 161-187.

Noble, J. (1998). Evolved signals: Expensive hype vs. conspiratorial whispers. In C. Adami, R. Belew, H. Kitano and C. Taylor (eds.), *Artificial Life 6: Proceedings of the Sixth International Conference on Artificial Life*. Cambridge, MA: MIT Press.

Oliphant, M. (1996). The dilemma of saussurean communication. *BioSystems 37*, 31-38.

Oliphant, M. (in press). The learning barrier: Moving from innate to learned systems of communication. To appear in *Adaptive Behaviour*.

Pinker, S. and Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Sciences 13*, 707-784.

Steels, L. and P. Vogt (1997). Grounding adaptive language games in robotic agents. In P. Husbands and I. Harvey (eds.), *Fourth European Conference on Artificial Life*, pages 474-482. Cambridge, MA: MIT Press.

Werner, G. and M. Dyer (1991). Evolution of communication in artificial organisms. In C. Langton, C. Taylor, J. Farmer and S. Rasmussen (eds.), *Artificial Life2*, pages 659-687. Redwood City, CA: Addison-Wesley.

Werner, G. and P. Todd (1997). Too many love songs: Sexual selection and the evolution of communication. In P. Husbands and I. Harvey (eds.), *Fourth European Conference on Artificial Life*, pages 434-443. Cambridge, MA: MIT Press.

Crucial factors in the origins of word-meaning

L. STEELS(1,2), F. KAPLAN(1), A. MCINTYRE (1) & J. VAN LOOVEREN(2)
(1) Sony Computer Science Laboratory - Paris
(2) VUB Artificial Intelligence Laboratory – Brussels
steels@arti.vub.ac.be, kaplan@csl.sony.fr, angus@csl.sony.fr,
joris@arti.vub.ac.be

We have been conducting large-scale public experiments with artificial robotic agents to explore what the necessary and sufficient prerequisites are for word-meaning to evolve autonomously through a self-organised process (see [Steels 97] for an overview of this new research field). The experiments employ an open-ended set of visually grounded autonomous robotic agents which play language games with each other about scenes containing geometrical objects before them (details about the cognitive architecture of the agents can be found in [Steels and Kaplan 1999]). The robots are located in different places in the world (Paris, Brussels, Tokyo, Antwerpen, Lausanne, San Jose, etc.) and are connected through the Internet. Agents are created by human users and can teleport between the different locations. Using a web-page (http://talking-heads.csl.sony.fr/) anyone can follow the experiment and interact with the agents to explore human influence on the emerging artificial language.

Our first `Talking Heads' experiment has been running for 4 months during the summer of 1999 and shows the validity of the mechanisms that were used for the agent architecture and of the interaction patterns and group dynamics of the agents. A shared lexicon and its underlying ontology emerged after a few days, enabling successful communication by the agents about the scenes before them. In total, 400,000 grounded games have been played. The population of agents reached 1500, increasing steadily over a period of 4 months. Despite the many perturbations due to grounding, intermittent technical failures, continuous influx of agent populations, and unpredictable human interaction, the lexicon was maintained throughout the period. A total of 8000 words and 500 concepts have been created, with a core vocabulary consisting of 200 basic words expressing concepts like up, down, left, right, green, red, large, small, etc.

The goal of this paper is to identify the factors that we found to be crucial for the success of the experiment. These can be grouped into two subsets: factors relating to the individual architecture of the agents and factors relating to the group dynamics and the environments encountered.

Internal factors

Agents must be able to engage in coordinated interactions. This means that they must be able to have shared goals and a willingness to cooperate. To enable a coordinated interaction, each agent must be able to follow a script of actions in agreement with a shared protocol, and have a way to see whether the goal of the interaction has been satisfied. In our experiment, we simply assumed this capability and explicitly programmed into each agent the scripts achieving the desired cooperative interaction. Emergence of cooperation is not addressed in this research.

Agents must have parallel non-verbal ways to achieve the goals of verbal interactions. The goal that we have chosen for the interaction between the agents is to draw attention through verbal means to an object in a visually perceived reality. There are of course many other things humans do with language but this is surely one of them and a prerequisite for more sophisticated verbal exchanges. We have found that it is crucial that the agents have a non-verbal way to achieve the same goal: by pointing, gaze following, grasping, etc. This alternative way must be sufficiently reliable, at least initially when the system is bootstrapping from scratch. Once the language system is in place however, external behavioral feedback is less crucial or may be absent altogether.

Agents must have ways to conceptualise reality and to form these conceptualisations, constrained by the ontology underlying the emerging lexicon and the types of situations they encounter. Obviously, conceptualisation precedes verbalisation. Words (even proper names) express categories as opposed to names of specific situations, but the repertoire of concepts need not and cannot be fixed in advance. There are in principle many equally effective possible ways to conceptualise reality. So there must be a concept acquisition process, for which we found important constraints:

First of all, the concept formation processes of the agents must be based on similar sensory channels and result in similar structures (even though there are still many possible solutions). We have incorporated this constraint at present by giving each agent the same low level sensory apparatus and by assuming binary discrimination trees for the agent's conceptual repertoires. Conceptualisation schemes based on randomly structured discrimination trees, prototypes, or inductive neural networks, are adequate for finding a distinctive conceptualisation but they result in larger differences between the repertoires of the agents and it is therefore more difficult to get coherence in the group. The strength of this constraint needs to be further explored.

Second, the conceptualisation for a particular game must itself also be constrained (even if there is a more or less shared repertoire) so that the agents have a reasonable chance to guess the conceptualisation that a speaker may have used. We have achieved this by using saliency: sensory differences that stand out more will be preferred for conceptualising the scene, thus reducing the search space for the meaning of unknown words.

Agents must have ways to recognise signals and reproduce them. This is quite obvious, because otherwise words would be confused all the time. In our experiment we have simply given the agents the capability to recognise or reproduce each others' signals perfectly. Other work has been going in our group to study how a repertoire of signals may itself become shared by an imitation game and the impact of errors in recognition or reproduction.

Agents must have the ability to discover what are the strongest associations (between words and meanings) in the group. The associative memory of an agent must be two-way (from words to meanings and meanings to words), must handle multiple competing associations (one word–many meanings, one meaning–many words), and must keep track of a score that represents how well the association has been doing based on their own past experience. When a decision must be made (which word to use, which meaning to prefer), there is an internal competition between different associations and the one with the highest score wins. All this can be achieved with a quite general associative memory mechanism. This mechanism could be used for other tasks, such as associating physical locations with sources of food. We can therefore assume this to be a standard part of the neural machinery of humans.

It is perhaps important also to point which factors we did NOT incorporate:

a. No theory of mind. There is a widespread belief that verbal communication requires a strong theory of mind of the other agents before verbal interactions are possible. In our experiment, this is not the case, even though for more sophisticated language games (such as for phrases like “I believe that you know the name of this woman.”) it is obviously required. However to get going, it is sufficient that agents follow specific protocols of interaction. They do not need to know why these protocols are successful. (Just like a child does not need explicit knowledge of theories of physics to throw a ball but just has to acquire the appropriate behaviors compatible with these laws.)

b. No prior ontology. There is also a widespread assumption that concepts (particularly the perceptually grounded concepts that are the focus of our experiment) need to be shared prior to and independent of language. For some cognitive researchers this implies that they are innate. For others, it suggests that they are acquired through a universal inductive mechanism that yields the same concepts for all agents. We do not assume a prior ontology in our experiments and in fact believe this to be impossible given the adaptive nature of verbal communication. Instead we have set up a strong interaction between language acquisition and concept formation: The ontology develops in a selectionist fashion under pressure from the language and concepts which have no success in verbal interaction are not encouraged.

c. No telepathy. We have not assumed that agents have a way of knowing what meaning the speaker transmits independently of language. Although non-verbal communication, similarity of sensors, shared history of past experiences, saliency, etc. help to restrict the set of possible meanings, the hearer can only guess what the speaker meant. Neither have we assumed that agents have exactly the same perception. Usually raw perception and consequently derived sensory features are different. Equal perception is of course an unrealistic assumption for embodied agents because each agent sees the scene from a different point of view.

External factors

The group dynamics must exhibit self-organisation so as to reach ontological and lexical coherence. We have achieved this in two ways. The agents have been made sensitive to the statistical spreading of word-meaning pairs in the population of individuals with which they interact by individually maintaining word-meaning scores (cf internal factor 5). But equally important is that they then use the associations with the highest scores because they will give the most success in the game. This creates a positive feedback loop between use and success. The more success a word-meaning pair has, the higher the score in each agent, and the more this word-meaning pair will be used in the future.

Second, a structural coupling has been established in each agent between the ontology and the lexicon. Each are independently developing processes but the lexicon gives feedback to the ontology on whether the conceptual repertoire is adequate for verbal interaction, and the ontology gives feedback to the lexicon proposing various possible conceptualisations. We see in the experiment that through this structural coupling agents settle on a shared ontology which is adapted to the environments they encounter, and do so without the need for innateness or a universal inductive process.

There must be sufficient group stability to enable a sufficient set of encounters between agents. We have found (in simulations) that if there is a too rapid in- and outflux of agents, a lexicon will collapse because there is not enough time for new members to acquire the conventions (so they build their own) and older members leave too quickly so that there is no memory in the population of the existing conventions. The exact critical levels of the fluxes depend on the size of the population. In the experiment, critical levels were not reached despite quite large changes in the population of active language users.

Another related constraint is that initial group size should not be too large so that there are enough encounters between the same individuals. Once a lexicon is in place however, there can be an almost unbounded increase in the population. The base lexicon in the experiment was created by a group of about 20 agents and then spread to the rest of the population, which eventually reached 1500 agents.

We have noted that sublexicons form when there is geographical separation, causing less opportunity for interaction, and that phenomena familiar in studies of language contact start to appear when the role of geographical separation is diminished.

There must be sufficient environmental stability and different degrees of complexity. The environments encountered by the agents and perceivable by the agents through their sensory apparatus must have certain invariant structural properties so that concepts can form and word-meaning pairs can settle. This does not mean that the environment needs to be closed (indeed it should not be if we want to be realistic), nor even that the sensory space should be closed (new sensory routines surely develop in the child even after she has acquired the first words).

We found that if the agents encounter only complex scenes, they cannot settle on a successful repertoire or at least they have much more difficulty due to unstable concepts. So there must be scenes, at least initially, which can be handled by making simple distinctions (such as between left and right). The learning environments of children exhibit this kind of graded complexity as well, partly because many sensory capabilities are initially not available thus simplifying the world.

For the external factors, we can cancel out other prerequisites. Based on our experience, they need not and often can not be assumed if we want to have a realistic model:

a. No global view nor central control. A central puzzle in the origins of language is how a population of distributed autonomous agents can reach coherence without a central controlling organism nor without access to a global view by the individual agents. A model should never introduce such a global control point. We have abundantly shown that self-organisation is perfectly adequate to explain language coherence without this.

b. No total coherence. It is often assumed that all individuals have exactly the same linguistic competence and that deviations are only due to performance errors. We have shown that this is not necessary. The conceptualisations and lexicons of the individual agents in the experiment were NEVER exactly the same. They had different degrees of knowledge and there were unavoidable individual differences due to the absence of a global view. The experiment shows that communicative success can nevertheless be reached without this absolute coherence. For example, words can often be maintained in a polysemous state without causing confusion in a series of environments, while synonyms are tolerated because agents can understand words that they themselves might not necessarily choose to use.

Conclusion

This paper is an attempt to show how experiments based on software simulations or robotic set ups, like the Talking Heads experiment, can play an important role in the debate on the origin and evolution of Human languages. In a field where “real” experimentation is not possible, this type of experiments enables to compare hypothesis and test through models which factors are crucial and which are contingent to achieve a communication system. It will be exciting to see now what we need to add to see the emergence and complexification of grammar.

References

[Steels 97] Steels, L. (1997) The synthetic modeling of language origins. Evolution of communication journal, 1 (1): 1-34.

[Steels and Kaplan 1999] Steels, L. and Kaplan, F. (1999) Situated grounded word semantics. In Proceedings of IJCAI 99, p.862-867.

Imitation, Segments, Words, and the Self

Michael STUDDERT-KENNEDY

Haskins Laboratories
New Haven, Connecticut, USA
msk@haskins.yale.edu

ABSTRACT

Words as units of symbolic currency in a syntax-free protolanguage are often taken for granted. Yet any vocal lexicon appreciably larger than the standard primate acoustic repertoire of 30-40 holistic signals demands engagement of the discrete combinatorial processes made possible by (and impossible without) a particulate vocal machinery, that is, by the six functionally independent articulators of the vocal tract (lips, tongue blade, tongue body, tongue root, velum, larynx). Without such a segmented machinery, few words; with few words, no syntax or higher level semantics. How then did a particulate vocal tract morphology and its associated neural controls evolve?

The proposed answer begins with the work of de Boer (London conference, 1998: Knight, et al., 2000) who showed, in a series of computer simulation studies, that the dispersion of vowels across the acoustic space of the vocal tract (in seemingly adaptive arrays that maximize perceptual distinctiveness and appear in every language) arises automatically over the course of a few thousand imitative exchanges between speaker-hearer agents, equipped with a modern vocal tract and a capacity to replicate the speech sounds of others. Evidently, adaptive dispersion arises directly from modern vocal tract structure and function, under pressure for vowel system growth and accurate imitation. In some sense, then, adaptive phonological systems are intrinsic to the vocal tract itself, built into it, as it were, by evolutionary history. From this assumption, it will be argued that vocal imitation and a particulate vocal machinery coevolved in an interactive spiral between the increasingly precise imitation necessary for communicative exchange, and an increasingly sharply segmented vocal machinery, a process analogous to the progressive differentiation of vocal control in modern children.

For insight into the mechanism of imitation, supporting its evolutionary emergence, we turn to two strands of recent research: "mirror neurons" in macaque cortex and facial imitation in infants. Rizzolatti & Arbib (Trends in Neuroscience, 1998, 21 , 188-194) found neurons, in an area of macaque cortex arguably homologous with Broca's area, that fire not only when a monkey grasps or manipulates food, but also when it sees a human experimenter do the same. Mirror neurons are of particular interest for vocal imitation because they seem to be organized not only by function grasping, manipulating, eating, and so on but also somatotopically. The likely somatotopic representation of the articulators, necessary for imitation in order to determine which articulators should act, is evident in children's early words. New tabulations of data on consonants suggest that errors on place of articulation are rare; errors tend to be of gestural amplitude and timing rather than of place. Somatotopic organization is also evident in infant facial imitation where errors of "organ identification" (i.e. which body part to move) are again rare. Thus, vocal and facial imitation (both unique among primates to humans) evidently draw on morphological correspondences between imitator and model.

Finally, it is precisely from these correspondences and from children's occasional errors in identifying them that we may infer a role for imitation in the emergence of the self, as distinct

from a conspecific other (Meltzoff & Moore (1998) in S Braten (ed.) Intersubjective Communication and Emotion in Early Ontogeny. Cambridge University Press). In both vocal and facial imitation, children recognize their errors without external feedback: if they make a mistake, they persist in correcting it by closer and closer approximations to the target. From this we may infer that imitative errors grant the child (as perhaps also the early hominid) a representation of both the similarities and the differences between the acts of itself and of a conspecific. Emerging recognition of words as shared phonetic forms may thus contribute to the emergence of words as expressions of shared meaning and intention.

Semantic-driven emergence of syntax : The principle of compositionality upside-down

Isabelle TELLIER

LIFL and Université Charles de Gaulle-lille3 (UFR IDIST)
59 653 Villeneuve d'Ascq Cedex, France
tellier@univ-lille3.fr
http://www.grappa.univ-lille3.fr/~tellier/

Introduction

In Chomskyan views of linguistics, syntax is a preeminent independent level whose knowledge is mainly innate. Semantics is then usually seen as a secondary structure deduced from the syntactic analysis. The Principle of Compositionality is, in this context, a precise way of specifying the passage from syntax to semantics.

But from an emergence perspective, this conception is very awkward. As a matter of fact, the ability to communicate meanings is of far higher priority than the ability to build a formal grammar, so *semantics must precede syntax*.

We still assume a computational point of view on syntax and semantics, but we propose to reinterpret the Principle of Compositionality to show that *syntax can derive from semantics*. The idea, inspired by recent results in the domain of Machine Learning, is to prove that a grammar can be completely specified by the description of the way semantic items (roughly corresponding with word meanings) are combined into global meanings. We thus provide theoretical arguments to avoid strong hypotheses about the innateness of syntax.

the principle of compositionality

Intuitive formulation

The Principle of Compositionality, mainly known by linguists and logicians, allows to characterize the connection between the syntax and semantics of natural languages. It is usually (and seemingly wrongly) attributed to Frege ([Janssen 97]). Its contemporary version states that : "the meaning of a compound expression is a function of the meaning of its parts and of the syntactic rules by which they are combined" ([Partee 90]). It has been the basis of several formal theories in computational linguistics, among which the best known may be Montague's semantics ([Montague 74], [Dowty 81]).

If the "parts" mentioned in the definition are assimilated with morphemes (or, to simplify, with words), and the "compound expressions" with phrases (which is the usual interpretation), this formulation implies that :

* words have individual meanings ;

* the meaning of a phrase (and thus of a sentence) *only* depends of the meaning of its words and of its syntactic structure.

The Principle of Compositionality has strong psychological justifications, as it "can explain how a human being can *understand* sentences never heard before" ([Janssen 97]).

Formal definition

The previous definition can be specified in a more formal way, inspired by [Montague 74] and [Janssen 97]. Two mappings need to be defined :

* a mapping associating each *word* with a *meaning*, i.e. a *semantic item* ;

* a mapping associating each *syntactic rule* (or, more generally, each *class* of syntactic rules) with a *semantic composition* (i.e. a function applying on semantic expressions and producing other semantic expressions).

Figure 1 shows on an abstract example how these mappings are combined to define a structure-preserving correspondence between syntactic and semantic trees. On the left tree, indexes g_1 and g_2 denote two classes of syntactic rules and on the right one, h_1 and h_2 denote the corresponding semantic compositions. If the grammar and both mappings are known, then the global meaning $h_2(h_1(\text{meaning}_1, \text{meaning}_2), \text{meaning}_3)$ of the sentence "word_1 word_2 word_3" can be automatically computed.

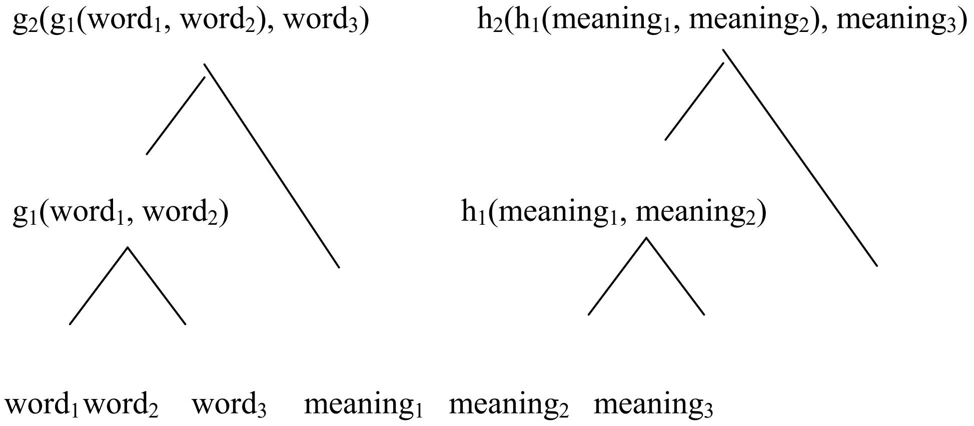


Figure 1 : *application of the Principle of Compositionality on an abstract example*

In cases of syntactic ambiguities, each different syntactic tree is associated with a different semantic tree whose global meaning may be different. Lexical ambiguities, due to polysemic words, can be handled by considering as many *copies* of the polysemic words as they have different meanings. Both mappings are in this case *bijective*.

Although this classical formal version of the Principle of Compositionality is only used as a one-sided way (*from syntax to semantics*), note that it is stated as a *tree isomorphism*. Thus, nothing prevents us from going *against* the usual stream.

Turning upside-down the principle of compositionality

Theoretical statements

We believe that the emergence of natural languages was mainly motivated by the need to convey not only atomic invariable meanings but also complex functional combinations of these semantic items. We will now show that syntactic structures can be considered as a *direct consequence* of such a need to combine meanings.

For this, we first have to turn upside-down the Principle of Compositionality. This means that what is now supposed to be given is a set of semantic items and a set of semantic compositions noted $\{h_i\}_{1 \leq i \leq n}$. We then consider two new mappings :

* a mapping associating each *semantic item* with a *symbol*, which can also be called, in Saussure's terms, a *signifier* ;

* a mapping associating each *semantic composition* h_i with *a class index* noted g_i, for any $1 \leq i \leq n$.

These mappings are again bijective : they are the reversed versions of the previous two mappings. The result of applying this reversed version of the Principle of Compositionality on a semantic combination is a tree structure whose leaves are symbols and whose nodes are indexed by the members of a finite set of elements. This resulting tree is *not*, in most cases, a full syntactic tree, but a simplified version of it, where usual non terminal symbols are replaced by class indexes. The interesting point is that this structure exactly coincide with a recently emerged notion in the domain of Grammatical Inference : the notion of *Structural Example*.

In the domain of Grammatical Inference, subpart of Machine Learning, the purpose is to identify a formal grammar from sentences it generates. Various theoretical results tend to prove that strings of words are not informative enough to specify a unique formal grammar ([Gold 67], [Valiant 84]). But Structural Examples, i.e. parenthesized strings of words with eventual class indexes, allow to achieve this goal : there exist algorithms able to converge towards the description of the unique formal grammar compatible with a set of Structural Examples ([Sakakibara 90 & 92], [Kanazawa 96]).

A detailed example

Our new version of the Principle of Compositionality applies *from semantics to syntax*. For sake of simplicity, let meaning representations be expressed by logical formulas. We note *John'* and *Mary'* two logical individual constants and run_1' and $love_2'$ two logical predicates of arity respectively 1 and 2. By convention, we suppose that the first argument of a two-place predicate coincide with its direct object and the second one with its grammatical subject. We suppose that the admitted semantic compositions h_1 and h_2 are *oriented functional applications* defined as follows :

* for any couple of semantic expressions a and b, $h_1(a, b)=a(b)$;

* for any couple of semantic expressions a and b, $h_2(a, b)=b(a)$.

The logical proposition : $run_1'(John')$, denoting the fact that “John runs”, can be obtained from the logical items and from the semantic compositions h_1 and h_2 in two ways : $run_1'(John') = h_1(run_1', John') = h_2(John', run_1')$. By the upside-down Principle of Compositionality, those two ways are respectively associated with two Structural Examples : g_1(run, John) and g_2(John, run), where “run” and “John” are the signifiers respectively associated with the semantic items *run'* and *John'*. Of course, the first structure will give rise to a grammar where verbs precede their grammatical subject and the second one to a grammar where grammatical subjects are uttered first.

Similarly, $love_2'(Mary')(John')$, expressing the fact that “John loves Mary” can be obtained in six various ways, each one corresponding with a possible ordering of a subject S, a verb V and a direct object O :

* $love_2'(Mary')(John') = h_2(John', h_1(love', Mary'))$: SVO

* $love_2'(Mary')(John') = h_2(John', h_2(Mary', love'))$: SOV

* $love_2'(Mary')(John') = h_1(h_2(Mary', love'), John'))$: OVS

* $love_2'(Mary')(John') = h_1(h_1(love', Mary'), John')$: VOS

For the last two possible orderings, because of our notational convention, we need the semantic item $\lambda x \lambda y.love_2'(y)(x)$, where lambda abstractions allow to invert the order of the arguments of the predicate, instead of $love_2'$:

* $love_2'(Mary')(John') = h_2(Mary', h_2(John', \lambda x \lambda y.love_2'(y)(x)))$: OSV
* $love_2'(Mary')(John') = h_1(h_1(\lambda x \lambda y.love_2'(y)(x), John'), Mary')$: VSO

The last two constructions are less frequently found in natural languages than the others. Each of these compositions specifies a unique Structural Example.

Combining semantic items in regular ways means that, if the semantic composition chosen to express “John runs” is $h_1(run_1', John')$, then the semantic composition chosen to express “John loves Mary” should be one where the predicate also precedes its first argument, and where the semantic composition h_1 is used to combine both items.

It has been proved in [Kanazawa 98] that huge sub-classes of context-free grammars are identifiable (in the sense of [Gold 67]) from Structural Examples built on the model of this example (i.e. based on a two-classes partition of the set of syntactic rules).

Conclusion

This work suggests a scenario for the emergence of syntax. The first step is the association of *symbols* (or signifiers) with semantic items. Computational simulations of this process have already been proposed ([Siskind 97]). The second step is the intention to communicate *combined meanings* built from the semantic items. The language of semantic representation is then supposed to be first acquired (or innate), but its syntax is much simpler than the one of natural languages. If these combined meanings are obtained in regular ways, then the definition of these combinations is equivalent, using our reversed version of the Principle of Compositionality, with the specification of a set of Structural Examples. In a last step, this sample of Structural Examples naturally leads to the description of a unique formal grammar. The only innate structures supposed are the semantic compositions : in our example, two very general functional applications are enough to explain the various possible orderings of phrases in natural languages.

Each natural language then appears as the result of choices made *at the semantic level* and reflected at the syntactic one through the upside-down Principle of Compositionality. Of course, other parameters than the order of phrases should be considered to distinguish one language from another one, and many other features could not be detailed here, but the upside-down Principle of Compositionality seems an ideal underlying mechanism allowing to connect semantic combinations with syntactic structures.

References

[Dowty 81] : D. R. Dowty, R. E. Wall, S. Peters, *Introduction to Montague Semantics*, Reidel, Dordrecht, 1989.

[Gold 67] : M. Gold, "Language Identification in the Limit", *Information and Control* 10, P447-474, 1967.

[Janssen 97] : T. M. V. Janssen, "Compositionality", in : *Handbook of Logic and Language*, Elsevier, Amsterdam and MIT Press, Cambridge, J. Van Benthem & A. ter Meulen (Eds), p417-473, 1997.

[Kanzawa 98] : M. Kanazawa, "Learnable classes of Categorial Grammars", CSLI Publications and FoLLI, *Studies in Logic, Language and Information*, 1998.

[Montague 74] : R. Montague, *Formal Philosophy; Selected papers of Richard Montague*, Yale University Press, New Haven, 1974.

[Partee 90] : B. Partee, A. ter Meulen, R. E. Wall, *Mathematical methods in Linguistics*, in "Studies in Linguistics and Philosophy" n¡30, Kluwer, Dordrecht, 1990.

[Sakakibara 90] : Y. Sakakibara, "Learning context-free grammars from structural data in polynomial time", *Theoretical Computer Science* 76, p223-242, 1990.

[Sakakibara 92] : Y. Sakakibara, "Efficient Learning of Context-Free Grammars from Positive Structural Examples", *Information and Computation* 97, p23-60, 1992.

[Siskind 97] : J. M. Siskind : "Learning Word Meanings", in *Computational Approaches to Language Acquisition*, M. Brent (Ed), MIT Press, 1997.

[Valiant 84] : L. G. Valiant, "A theory of the learnable", *Communication of the ACM*, p1134-1142, 1984.

Serial expertise and the evolution of language

H. S. TERRACE

Columbia University

ABSTRACT

Language and thought are inextricably entwined. Until the latter half of the 20th century there was little reason to ask whether language and thought could be dissociated, that is, whether thinking could occur without language? That question was raised and answered affirmatively by experiments on animal cognition. The results of such experiments have obvious implications for the evolution of language. To the extent that our non-human ancestors could think without language we can assess the contribution of language to pre-existing cognitive processes.

With that goal in mind, I would like to summarize some experiments I've performed on the serial skills of monkeys. Serial skills play a crucial role in intelligent behavior, in particular, during the development of various types of cognitive expertise. In one experiment, monkeys learned, by trial and error, to execute lists composed of 7 arbitrarily selected photographs. The lists were trained with all of the items displayed simultaneously throughout each trial. The location of list items was varied from trial to trial (typically on a touch-sensitive video monitor) to insure that the subjects could not learn a list as a fixed-motor pattern. The task was to respond to each item in a particular order, regardless of its spatial position. To execute a simultaneous chain, subjects had to generate internal cues to define its position in the sequence (that is, think about which item to respond to next) before making each response. Although the probability of executing the entire list correctly by chance was 1/5040, each monkey readily acquired four 7-item lists and showed evidence of increased serial expertise as they learned successive lists.

The following thought experiment illustrates the serial skills needed to execute a 7-item sequence. Imagine trying to enter your personal identification number (PIN) at a cash machine, say *9-2-1-5-8-4-7,* on which the positions of the numbers were changed each time you tried to obtain cash. You could not enter that (or any) PIN by executing a sequence of distinctive motor movements, i.e., first pressing the button in the lower right corner of the number pad to enter *9*, then the button in the upper middle position to enter *2*, and so on. Instead, you would have to search for each number and mentally keep track of your position in the sequence as you pressed different buttons. With the exception that 7 different photographs were used instead of 7 different numerals, this is precisely the problem that the monkeys had to solve when they were trained on 7-item lists. As difficult as that may seem, it would be far more difficult to discover the correct PIN by trial and error. To do so, you would have to recall the consequences of any of the 36 types of logical errors you could make while attempting to produce the required sequence. With the exception that different photographs were used instead of different numerals, this is precisely the problem that monkeys had to solve at the start of training on each of the 4 lists on which they were trained.

Three types of evidence of serial expertise were obtained. (1) There was a progressive decrease in the number of sessions needed to satisfy the accuracy criterion on each list. Subjects needed, on average, 31.5, 17.5, 13 and 12.25 sessions to master Lists 1, 2, 3 and 4, respectively (ranges: 21-55, 11-25, 11-19, and 7-17, respectively). (2) Because food reward was provided only after responding to all 7 items in the correct order, subjects had to determine the correct order in which to respond to the first 6 items of each list solely on the basis of the secondary

reinforcement they received following correct responses and errors. (3) Subjects became progressively more proficient at remembering the consequences of correct responses and errors at each position of the sequence. Indeed, solely on the basis of secondary reinforcement they almost succeeded in identifying the first 2 items of their 4th list with the minimum number of logical guesses.

This is the first study to show the development of serial expertise in an animal. Within the framework of human memory systems, the expertise observed in this experiment seems closer to that characterized as declarative knowledge ("knowing that" C comes before E, but after B, etc.) than that characterized as procedural knowledge ("knowing how" to execute a fixed motor sequence). Given that our subjects' serial expertise cannot be characterized linguistically, it is unclear how a non-human primate represents declarative knowledge. It is clear, however, that procedural knowledge cannot suffice as an explanation of our subjects' behavior during the execution of a list because they had to generate a different motor sequence on each trial. Another feature of subjects' serial expertise that falls outside the realm of procedural knowledge is their ability to ascertain an item's ordinal position with almost maximal logical efficiency.

The sequences that the 4 subjects of this study learned are by far the most difficult lists mastered by non-human primates, including those trained in experiments on the linguistic abilities of ape. It is, however, doubtful that their performance reflects the upper limit of their serial capacity. The ease with which each subject learned 7-item lists and the steady decrease in the number of sessions needed to master new lists suggests that they could learn them in fewer than 13 sessions and that they could also master longer lists. Their performance also shows that the serial expertise for producing arbitrary sequences evolved long before the evolution of language in hominids and that it is possible to investigate mechanisms for planning and executing complex sequences in a non-human primate without the complication of language.

Social cognition and the origins of linguistic communication

Michael TOMASELLO

ABSTRACT

Human communication is most clearly distinguished from that of other primate species by its use of (a) symbols, and (b) grammar. Both of these unique features are products of a unique form of human social cognition and its deployment in certain forms of social interaction. Linguistic symbols are actually constituted by two or more humans with the species-unique form of social cognition attempting to communicate with one another - i.e., by the human form of intersubjectivity. Grammar results from the historical process of grammaticalization in which speakers using loose discourse sequences of symbols create from them more tightly organized syntactic constructions - often creating in the process grammatical items that serve as the structuring elements of those constructions. Grammaticalization derives from certain features of human intersubjectivity, social interaction, and information processing at work in repetitive - and therefore to some degree predictable - communicative situations.

Minimally biased learners and the emergence of compositional language

Bradley TONKES[*] **& Janet WILES**[**]

[*] *Department of Computer Science and Electrical Engineering*
University of Queensland, 4072, Australia
btonkes@csee.uq.edu.au

[**] *Department of Computer Science and Electrical Engineering,*
School of Psychology
University of Queensland, 4072, Australia
janetw@csee.uq.edu.au

ABSTRACT

With advances in computational techniques, it has become possible to model aspects of the origin and subsequent evolution of languages. One of the more intriguing aspects of language evolution is the emergence of syntax. While many species have evolved forms of signalling systems, humans are seemingly alone in the use of recursive compositional structures that result in ``essentially infinite'' languages. The complexity of human languages raises the question of how human infants acquire language, particularly since the infant learner can never observe the entire language. Many linguists argue for a strong form of innate linguistic endowment that provides the human infant with knowledge about the universal structures of human languages. An alternative view emphasises the point that languages themselves can act as complex adaptive systems, which evolve to their human ``hosts''.

Kirby (1999a, 1999b) gives a compelling demonstration of syntactic structures emerging from a population of agents. Despite the absence of phylogenetic adaptation in the population, a compositional language emerges as a result of the dynamics of language acquisition. Kirby proposes that individuals' languages can be described in terms of *replicators* (corresponding to grammatical rules or rule-sets) that compete for survival, with the more general replicators having a greater chance of being learned. Thus, languages evolve towards forms that consist of broad, compositional rules where every utterance provides an opportunity to learn the general rule.

Although there is no phylogenetic adaptation in the course of Kirby's simulations, the model incorporates phylogenetic adaptation implicitly in the design of the individuals' language learning mechanisms. That is, the starting point of the simulations is a population of individuals who are innately endowed with a particular learning mechanism. It seems to us that the chosen induction algorithm is highly biased towards language-like, compositional structures. Although Kirby highlights the importance of languages as adaptive systems that adapt to their human hosts, inherent in his choice of learning algorithm is a strong form of language-specific learning bias. Kirby's simulations demonstrate a set of biases that are sufficient for the emergence of compositional syntax. In this study we wish to consider their necessity. That is, are strong domain-specific learning biases required for compositional language to emerge?

In previous work (Tonkes, Blair and Wiles, 2000) we have considered communication between a pair of recurrent neural networks which provide a more general-purpose learning

mechanism than that used by Kirby (though like any learning algorithm, some degree of bias is inherent). The two networks try to communicate a ``concept" over a symbolic channel. The *encoder* network sends a sequence of symbols for each concept, which the *decoder* network receives and processes back into a concept. Using this framework, we have shown how a language can evolve to mediate opposing biases between encoder and decoder, and how language evolution can facilitate learning by adapting towards the forms that exploit the weak biases of a general purpose learner.

In the present study, we investigate the extent to which domain-specific biases are required to replicate Kirby's (1999a) results. We use the same basic design, but substitute our learner (recurrent neural network) and task. The initial population consists of recurrent encoder networks with random weights arranged in a ring. Each of these networks map concepts in the unit interval to sequences of symbols. Following Kirby, in each generation one network is replaced by a new individual which is subsequently trained on the language of its neighbours. As new individuals are introduced to the population, the overall language of the population gradually changes so that only the learnable forms persist.

Although populations converge to languages that are easily learnable, the evolved languages are degenerate - the networks use the same message for each meaning. Attempts to enforce language expressivity result in the non-convergence of the population. Consequently, we investigate the constraints that are necessary in a learning mechanism for the emergence of language that is both expressive, and consistent across a population.

References

Kirby, S. (1999a). Syntax without natural selection: How compositionality emerges from vocabulary in a population of learners. In Knight, C., Hurford, J., and Studdert-Kennedy, M., editors, *The Evolutionary Emergence of Language: Social function and the origins of linguistic form*. Cambridge University Press.

Kirby, S. (1999b). Learning bottlenecks and the evolution of recursive syntax. In Briscoe, T., editor, *Linguistic Evolution through Language Acquisition: Formal and Computational Models*. Cambridge University Press.

Tonkes, B., Blair, A., and Wiles, J. (2000). Evolving learnable languages. In Solla, S. A., Leen, T. K., and Muller, K. -R., editors, *Advances in Neural Information Processing Systems 12*. MIT Press.

Evolving language, I-consciousness and free will

Rüdiger VAAS

Institute of Philosophy, University of Stuttgart & Institute of Zoology, University of Hohenheim, Posener Str. 85, D – 74321 Bietigheim-Bissingen, Germany
Ruediger.Vaas@t-online de

This paper has two goals: (1) to extend explanations of the evolution of language, I-consciousness and our impression of having free will in the light of what is now called the “social intelligence hypothesis”: the evolution of language is forced by natural selection mainly because of its advantage as a tool and weapon for and within the social struggle of our ancestors; (2) to show how biological and linguistic insights may contribute to the understanding of one of the most puzzling philosophical issues – and indeed of our conception as human beings – i.e. (the possibility of) our experience of ourselves and as autonomous agents. The philosophical problems of I-consciousness and free will cannot be *solved* as it would require the reconcilation of apparently inconsistent premises; but it may be *dissolved* by eliminating one of the premises, namely the claim that there are irreducible entities like free-floating selves or Cartesian egos with the ability to act due to their own non-physical power. Nevertheless our misleading *conception* of being such selves with free will has to be explained. And evolutionary biology and linguistics seem to be able to do this: The ego-illusion of systems which permanently confuse themselves with their own self-model, and the (in some sense inadequate) belief of having free will are sophisticated tools with great evolutionary advantages – they are the most subtle form of deception that was rewarded by natural selection, namely, a systematic and stable deception of our own.

Obviously, organisms need not be very mindful to live and reproduce. But some are. Why? Considering social factors are the most promising approach for an answer (Byrne & Whiten 1988, Whiten & Byrne 1997). A main starting point was the observation that primates appear to have more intelligence than is required for their everyday wants of feeding and ranging. Since evolution is unlikely to select for superfluous capacities, Nicholas Humphrey (1976) conjectured that something had been fergotten, namely the social complexity inherent in many primate groups, and suggested that *the social environment might have been a significant selective pressure for primate intelligence*. Since better access to food or a safer place to sleep or a higher rank in the complex hierarchies of primate societies normally increase the probability of producing more offspring than other group members, social intelligence pays off pretty well. Natural selection therefore favours it (or its inherited requirements). And since this selective pressure applies to all group members, an *evolutionary arms race* is set up, leading to a further increase of intelligence. This development probably corresponds to the *rapid expansion of our ancestors' neocortex* – especially the frontal parts, which are most important for working memory and planning (Goldman-Rakic 1992) and probably consciousness (LeDoux 1996). This cortical enlargement – about a factor of three to four during the last five million years – is otherwise hard to explain. And it is biologically expensive, because the brain consumes about 20 percent of the energy when the body is idle but accounts for only two percent of its mass. Furthermore, there is evidence for a correlation between neocortical size and group-size or social complexity (Barton & Dunbar 1997).

Thus, social interactions might have been the most important driving force for the evolution of primate intelligence. The elaborated mental abilities of higher primates are conceived as the product of a cognitive arms race leading to more and more sophisticated representational capabilities (representation of complex social relationships, higher-order intentional stance, theory of mind, mindreading). This climate of competition and conflict favours the use of social manipulation to achieve individual benefits at the expense of other group members. Observing social relationships carefully, struggling for influences, making alliances, or deceiving more powerful leaders got more and more important. Particularly useful for this are manipulations in which the losers are unaware of their loss (as in some kinds of deception), or in which there are compensatory gains (as in some kinds of co-operation). Therefore, egoistic intentions remain hidden. A lot of zoo and field experiments as well as behavioral studies in the wild have already confirmed (and reinforced) these hypotheses. It was shown, for example, that apes – and to a lesser degree perhaps also monkeys – may be able to respond differently, according to the beliefs and desires of other individuals (rather than according only to the other's overt behavior). Hence, they possess a theory of mind (Premack & Woodruff 1978) and can assume what Daniel Dennett (1988) has called the *intentional stance*: They ascribe intentions to others and take them into consideration for their own actions.

Language is, among other things, a very useful tool and medium for explicit representations and metarepresentations including an intentional stance, self-attributions, I-consciousness, higher-order volitions, autonomous agency etc. These are not an epistemic luxury but have a function, i.e. a causal role. They allow a more precise representation of the external and internal states and their rational and emotional evaluation. They allow a broader range of reactions in complex situations, especially in social contexts. The concept of self reifies the organizing activity of an organism that incorporates its experience into its future actions. These capabilities are – at least at the higher-order level of human beings – based on and boosted by language, and this is probably the main reason for the development of larger brains and linguistic capabilities (cf. Goody 1997). Thus, it is reasonable to assume that these cognitive capabilities are an important factor for the origin and evolution of language and cannot be excluded by any elaborated theory trying to explain this still rather mysterious issue (cf. e.g. Aitchison 1996, Jablonski & Aiello 1998, Noble & Davidson 1996): Language was incorporated in cognitive representations of own's and others' intentions and offered more abstract and efficient ways to use these representations; language permits more effective classification, storage and distribution of information, and thus more efficient use of memory and communication; language is an important means to envisage the future; and language-in-use is a new and very effective sort of tool for co-operation between individuals, because it makes information explicit and easily communicable even in the absence of visual contact. Language also paved the way for even more sophisticated deceptions (i.e. lies) and influencing others to act in accordance with one's own goals. Language is based on symbolic and abstract thought, but conversely it also enhanced their further development. Finally, language lead to more and more sophisticated models of the world and of ourselves.

Self-consciousness is a rather shaky term with many different meanings which often depend on each other, e.g. notions like self-awareness, self-knowledge, self-recognition, sense of ownership etc. (cf. Frank 1994, Bermúdez, Marcel & Eilan 1995). Self-consciousness is not a single ability or property but a complex entanglement of different features creating a special kind of knowledge. As a premise, it is assumed here that self-consciousness does not come ready-made into existence, but bootstraps itself with the help of other minds in a complex interplay of the infant with the social and physical environment starting from inborn dispositions. It depends on perspectivity due to centered information acquisition, bodily awareness due to proprioception and feedback from results of one's own actions (including the experience of resistance). These are crucial ingredients for a higher-order form of self-consciousness, i.e. *I-consciousness*. It is conceptualizable and verbalizable. It is based on a feature which is called a *self-model*. This is an episodically active representational entity (e.g. a complex activation pattern in a human brain), the contents of which are properties of the system itself. It is embedded and constantly updated in a global model of the world created also by the brain based on perceptions, memories, innate informations etc. (Metzinger 1993). Self-models are limited in a crucial way. They *cannot* represent their own representations as their own representations as their own representations and so on *ad infinitum*. But there is (or at least was) also no need for that. From an evolutionary perspective, it would have been quite disadvantageous for our ancestors to forget their physical and social environments and plunge into a self-amplifying spiral of self-reflection. Hence, there is a – probably hard-wired – *self-referential opacity*: The phenomenal mental models employed by our brains are semantically transparent, i.e. they do not contain the information *that* they are models on the level of their content (Van Gulick 1988). Possibly these phenomenal mental models are activated in such a fast and reliable way that the brain itself is not able to recognize them *as* such anymore because of a lower temporal resolution of metarepresentational processes due to limited temporal and physical ressources. If so, the system "looks through" its own representational structures as if it was in direct and immediate contact with their contents, creating a special sort of self-intimacy. This leads us to a rather dramatic – and possibly offending – hypothesis: We are systems which are not able to recognize their self-model as a self-model. For this reason we are permanently operating under the conditions of a "naive-realistic self-misunderstanding". We experience ourselves as being in direct and immediate epistemic contact with ourselves. Hence, we are systems which permanently confuse themselves with their own self-model (Metzinger 1996). In doing this, we generate an *ego-illusion*, which is stable, coherent, and cannot be transcended on the level of conscious experience itself.

Another controversial issue is the problem of free will (see e.g. Honderich 1988, O'Connor 1995, Walter 1998 for an introduction). To define free will in the strongest sense, Libertarians often presume three necessary conditions which, taken together, are sufficient: intelligibility, freedom, and origination. *Intelligibility* means that a person's free choices are based on intelligible reasons. *Freedom* means that this person can make different choices under completely identical conditions, i.e. that this person could act otherwise even if all natural laws and boundary conditions (including his or her own physical states) are the same. *Origination* means that the person is able to create his or her choices and acts according to these choices in a nonphysical way. But this presupposes an ontology (e.g. a kind of dualism or idealism) which goes beyond and is at least partly independent of the physical world. However, even such an ontology won't offer what Libertarians want, for it cannot avoid the dilemma of plunging into an *infinite regress* or abruptly step on the brake at a *mysterious causa sui*. This is because in order for me to be truly or ultimately responsible for how I am, so that I am truly responsible for what I want and do (at least in certain respects), something impossible has to be true: There has to be a starting point in the series of acts that made me have a certain nature – a beginning that constitutes an act of ultimate self-origination. But there is no such starting point. Therefore, even if I can act as I please, I can't please as I please. That is not to say that there are no higher-order volitions, for instance wanting to want not to stay that lazy anymore. But ultimately my reasons, beliefs and volitions are non- (our sub-)consciously determined – by earlier experiences, heredity, physiology or external influences – and therefore not *ultimately* up to me. Thus, in order to be ultimately autonomous and responsible, one would have to be the ultimate cause of onself, or at least of some crucial part of oneself (Strawson 1986). But this would strangely promote man to something like an Aristotelian God, a prime mover. (This is no polemic exaggeration but what Libertarians have actually conceded, see e.g. Chisholm 1964, p. 32, Kane 1989, p. 121.)

However, there is no hint for the existence of humans as prime movers and nonphysical forces interacting with our physical world through causal loopholes. Nevertheless we do conceive ourselves, at least sometimes, as being free. We have the *feeling* that it is up to us to decide between alternatives. This feeling depends on second-order emotions (without which we cannot act and choose in complex situations despite of rationality), an intentional stance, a "healthy" (non-deprived) development, non-predictability or epistemic indeterminism (that is to say we cannot know the future for certain, and especially not our own future), rationality (the ability to reflect and reason), planning (and hence higher-order thoughts, a concept of the future et cetera), higher-order volitions, and sanity. These features are compatible with a naturalistic world view (Vaas 1996 & 1999) and even with determinism. Therefore it is not to deny a weaker form of free will. But this does not imply the existence of the kind of freedom and origination for which Libertarianism is argueing. The Libertarian will still insist that our subjective impression of freedom is a powerful argument for free will. Thus, a sceptic should be able to explain such an impression within a naturalistic framework. And this is what an evolutionary perspective might achieve: *Ascribing intentional states to others necessarily includes ascribing volitions to them and assuming that they have the power to transfer their volitions into actions somehow*, because this is the only way to get advantages from the intentional stance at all. For, if other beings are thought to have intentions but they would be causally inert, that is to say their behavior has nothing to do with their volitions, this ascription of intentions and hence volitions simply wouldn't matter. However the intentional stance is not an irrelevant luxury. It is a powerful tool to get along with the complexity of the social world and even an anthropomorphically-conceived nonsocial world (up to highly restricted activities – e.g. in playing computer chess nowadays it is common and helpful to think and act as if the computer "wants" and "plans" something). Individuals endowed with this tool are better prepared for the struggle of social life. And it is advantageous to assume the volitions of others as somehow being independent of the environment or the past. Not absolutely independent of course, but in an approximate sense – because this makes it a lot easier to deal with them due to the fact that complex organisms can act (or react) quite differently in similar circumstances and quite similar in very different circumstances. There is another reason to take a concept of volition as evolutionarily advantageous, and this is just the other side of the coin: To deal with other individuals in a complex way means also *to plan one's own actions carefully and evaluate their effects*. This presupposes some kind of awareness of one's own volition, hence a concept of will and self. Higher-order representations also take one's own mental states into account – not only for decisions and follow-up analyses but also as a parameter in the plans of others regarding oneself. Thus, it is reasonable or even necessary to ascribe volitions to oneself, too – because otherwise one cannot reason about the mental states of others who are presumably dealing with oneself. This makes one's own volitions explicit – and much more flexible. For instance, an individual may think: "She believes that I want to do this, and she will react to this in a certain way to get an advantage over me – and therefore I will act otherwise and not do this but that." At least since the point from which there has been *language* with an inbuilt grammatical structure distinguishing between subjects and objects, active and passive, present and future – but probably much earlier –, such concepts of volition, actions and self-notions have been flourishing. This was not only the case in contexts of cheating, however! In the course of time *co-operation* became more and more important among our early ancestors. And the existence of some form of language already implies a high degree of co-operation (Calvin & Bickerton 2000) – spoken language would never have emerged unless most people, most of the time, followed conventional usage. But co-operation in complex, not inherited forms also presupposes an intentional stance and the capacity to ascribe volitions to others.

In conclusion, evolution shaped our minds respectively our brains to cope with our complex social lives. We are forced by our very nature to interact with other people in a fundamentally different way than to interact with, say, stones and sticks (Strawson 1962). From this it is no longer a big step to a notion of free will which is a powerful tool to act in consonance with or opposition to others and to establish some kind of moral responsibility – a very effective way to influence the behavior of others and justify punishments. Thus, free will even succeeded to become an entitiy of religious, philosophical or political theories and a postulate for jurisdiction. Of course we need not dismiss an intentional and personal stance. It is, obviously, crucial for our survival. We cannot leave our subjective standpoints, turning exclusively to an objective, perspectiveless view. We may accept that we have, ultimately, no free choice. Nevertheless, in our everyday life we think and act as if we did. Even sceptical philosophers do – or they might find themselves out of the race quickly. Nature is stronger than insight and “the human brain is, in large part, a machine for winning arguments, a machine for convincing others that its owner is in the right – and thus a machine for convincing its owner of the same thing. The brain is like a good lawyer: given any set of interests to defend, it sets about convincing the world of their moral and logical worth, regardless of whether they in fact have any of either. Like a lawyer, the human brain wants victory, not truth; and, like a lawyer, it is sometimes more admirable for skill than for virtue” (Wright 1994, p. 280).

For some helpful comments or discussions I am grateful to André Spiegel, Henrik Walter and Thomas Zoglauer.

References

Aitchison, J 1996: The Seeds of Speech. Cambridge: Cambridge University Press.

Barton, RA & Dunbar, RIM 1997: Evolution and the social brain. In: Whiten & Byrne 1997.

Bermúdez, JL, Marcel, A & Eilan, N (eds) 1995: The Body and The Self. Cambridge: MIT Press.

Byrne, R & Whiten, A 1988: Machiavellian Intelligence. Oxford: Clarendon Press.

Calvin, W & Bickerton, D 2000: Lingua ex Machina. Cambridge: MIT Press.

Chisholm, R 1964: Human Freedom and the Self. In: Watson, G (ed): Free Will. New York: Oxford Univ. Press 1982.

Dennett, D 1988: The Intentional Stance in Theory and Practice. In: Byrne & Whiten 1988.

Frank, M (ed) 1994: Analytische Theorien des Selbstbewußtseins. Frankfurt am Main: Suhrkamp.

Goldman-Rakic, P 1992: Working Memory and the Mind. Scientific American 267, 110-117.

Goody, EN 1997: Social Intelligence and Language: Another Rubicon? In: Whiten & Byrne 1997.

Honderich, T 1988: A Theory of Determinism. Oxford: Clarendon Press.

Humphrey, N 1976: The Social Function of Intellect. In: Consciousness Regained. Oxford: Oxford University Press 1998.

Jablonski, NG & Aiello, LC (eds) 1998: The Origin and Diversification of Language. San Francisco: Univ. of Calif. Pr.

Kane, R 1989: Two Kinds of Incompatibilism. In: O’Connor 1995.

LeDoux, J 1996: The Emotional Brain. New York: Simon & Schuster.

Metzinger, T 1993: Subjekt und Selbstmodell. Paderborn: Schöningh.

Metzinger, T 1996: Niemand sein. In: Krämer, S (ed): Bewußtsein. Frankfurt am Main: Suhrkamp.

Noble, W & Davidson, I 1996: Human Evolution, Language, and Mind. Cambridge: Cambridge University Press.

O'Connor, T (ed) 1995: Agents, Causes, Events. Oxford: Oxford University Press.

Premack, D & Woodruff, G 1978: Does the Chimpanzee have a Theory of Mind? Behav. and Brain Sciences 4, 515-526.

Strawson, G 1986: Freedom and Belief. Oxford: Clarendon Press.

Strawson, PF 1962: Freedom and Resentment. Proceedings of the British Academy 48, 187-211.

Vaas, R 1996: Mein Gehirn ist, also denke ich. In: Hubig C & Poser H (eds): Cognitio humana. Leipzig, Vol. 2.

Vaas, R 1999: Why Neural Correlates of Consciousness are fine, but not enough. Anthropology & Philosophy III, No. 2.

Van Gulick 1988: A functionalist Plea for Self-consciousness. The Philosophical Review XCVII, 149-181.

Walter, H 1998: Neurophilosophie der Willensfreiheit. Paderborn: Schöningh.

Whiten, A & Byrne, R 1997: Machiavellian Intelligence II. Cambridge: Cambridge University Press.

Wright, R 1994: The Moral Animal. New York: Pantheon.

The origin of phonetic abilities: A study of the comparative data with reference to the aquatic theory

Marc VERHAEGEN & Stephen MUNRO

Studiecentrum Antropologie
Mechelbaan 338, 2580 Putte, Belgium
marc.verhaegen@village.uunet.be
http://allserv.rug.ac.be/~mvaneech/Verhaegen.html

Human language has three major components: phonology, semantics, and grammar. These components are acquired successively from about the first, the second and the third year of life (Hirsch-Pasek & Golinkoff 1996). This succession may reflect the human linguistic evolutionary stages: pre-language, one-word sentences, and grammatical or "true" language. This paper discusses the first stage, the phonetic pre-adaptations for language, and is based mostly upon comparative data with other mammals.

An amphibious scenario

Most researchers agree that our remote primate ancestors lived in trees. A number of different hypotheses, however, have been put forward to explain the latter part of our evolutionary history. Recently, there has been a steady accumulation of evidence suggesting that human evolution may not have started in a warm and dry milieu, as the savanna theory used to assert. Instead, the anatomical, physiological, biochemical and palaeo-environmental data suggest that humans may have evolved in warm and wet conditions, perhaps at the edge between land and water (Hardy 1960, Morgan 1997, Bender *et al.* 1997, Verhaegen 1997, Tobias 1998). More specifically, at the recent symposium on *Water and Human Evolution* (Ghent, Belgium, 1999), it was proposed that our ancestors may have been coastal or riverside omnivores, which not only consumed terrestrial plants and animals but also collected part of their food in the water. That means their lifestyle may have included wading and even diving for shellfish, seaweeds, crustaceans and fish.

Typically human features such as bipedal locomotion and frequent tool use are occasionally seen in some nonhuman primates. Proboscis monkeys, for example, cross shallow stretches of water on two legs when moving from one mangrove tree to another, and lowland gorillas go wading on their hindlimbs through forest swamps in search of what researchers call *aquatic herbaceous vegetation* or *AHV* (Chadwik 1995). In the same way, Pliocene hominids might have waded in shallow waters in forest clearings, gallery forests or mangrove areas, in search of floating fruit, sedges, reeds, AHV, fish and/or shellfish, all of which were probably available, and all of which hominids were probably capable of eating (Sponheimer & Lee-Thorp 1999, DuBrul 1977, Puech *et al.* 1986, Puech 1992, Broadhurst *et al.* 1998).

Tool use is seen in sea otters, capuchin monkeys and chimpanzees, which all use stones to crack open hard-shelled foods such as shellfish and nuts (Fernandes 1991). It is possible that human Pliocene ancestors, like mangrove capuchins, manipulated hard objects to remove and open the oysters which grew on the mangrove trees in the areas in which they lived.

Later, during the late Pliocene or early Pleistocene, members of the genus *Homo* – as opposed to our more distant relatives the australopithecines – might have also learned how to dive and collect deeper shellfish and other aquatic resources. This could explain why humans have such remarkable diving capabilities, particularly when compared to all nonhuman primates (Schagatay 1996). Indeed, *Homo* fossils – as opposed to *Australopithecus* – are typically found near shellfish (Chiwondo, Chemeron, Nariokotome, Zhoukoudian, Boxgrove, Terra Amata, Rabat, Hopefield, Gibraltar and others). *Homo erectus* remains especially have been discovered in marine sediments amid shellfish, barnacles and corals, from the early Pleistocene skull of Mojokerto at Java (Ninkovich & Burckle 1978), to the late Pleistocene Acheulean tools in Eritrea (recent discoveries of Robert Walter and co-workers). Possibly, *Homo erectus* spread along the Indian Ocean coasts, where they gathered foods from the land and the sea, and where they later followed rivers inland to colonise the interiors of Africa and Eurasia.

Whereas stone tool use for cracking hard-shelled foods may have been a preadaptation for the development of lithic technologies, the diving abilities of our ancestors might have been a preadaptation for the development of voluntary speech. However, like Darwin (1871), we believe human sound production probably has deeper roots, beginning at a time when our ancestors were still arboreal. Comparative studies suggest there may be several, perhaps overlapping, preadaptations for speech, including musical abilities, swallowing musculature, the ability to close the airways, and voluntary control of breathing.

Singing – Vocal Chords

Babies of two or three months produce cooing sounds. This is called *vocalising* and is performed with the vocal chords in the larynx, without much oral movement. Soon thereafter, even in deaf children, the *babbling* starts to include labial consonants, and syllables are produced (consonant plus vowel). In babies older than six months, the sound pattern already resembles the native language, and “dialogues” with the mother stimulate the utterances.

The early prelingual sounds, without symbolic meaning, may correspond with the elaborate songs of nonhuman primates like gibbons. Music powerfully affects our emotions (anthems, hymns, marches, love songs), and has always been a territorial and pair- or group-binding behaviour, as it is in other animals. Well-developed musical abilities and duet singing are seen in several monogamous primates like indris, tarsiers, titi monkeys, and gibbons (Darwin 1871). Bonobos engage in group chorusing, and rival males of this species have been observed engaging in vocal duels (De Waal 1997). Some aquatic species like humpback whales (polygynous) also use complex melodic utterances for territorial behaviour. It is known that musical training in young children induces an enlargement of the *planum temporale* and the auditory cortex in the left brain hemisphere, and can also lead to an improvement in a child’s ability to hear absolute tones (Schlang *et al.* 1995). Intonation is an indispensable element of all spoken languages, and almost half of the world’s langauges are tonal.

Diving – Airways

An important development in human infants between four and six months is the descent of the larynx. One possible explanation for laryngeal descent in humans could be the need to breathe a large amount of air in a short period of time to facilitate diving (Morgan 1997).

All humans, unlike nonhuman primates, can learn to dive. Several human populations, such as some Indonesian and Oceanic populations, as well as the Ama of Korea and Japan, still collect shellfish through breath-hold diving, and there is reason to believe that the ancestors of humans, at least since the time of *Homo erectus*, also had this ability (Schagatay 1996, Verhaegen 1997).

Diving, as seen in aquatic or semi-aquatic mammals, requires a voluntary control of breathing. In contrast with land mammals, divers must be able to take a deep breath just before they intend to dive, and breathe deeper and faster between dives. They must also be able to hold their breath underwater, exactly when their oxygen needs are highest. In contrast, terrestrial mammals intensify their breathing at the moment they need more oxygen; while running, for example. Diving also requires the complete closure of the airways underwater, so that water can be kept out of the lungs.

Feeding – Mouth and tongue

A descended larynx may also have been a useful adaptation for swallowing certain aquatic foods, such as oysters, and perhaps also for feeding underwater. Other animals featuring a descended larynx include so-called "suction feeders" of seafood or fruits, such as some pinnipeds and some bats (Fay 1982, Negus 1962, Rosewear 1965). Possibly, the laryngeal descent allows considerable retraction of the tongue so that the pressure in the oral cavity can be lowered, which is one possible way to accommodate underwater suction, as in walruses, or for sucking juicy fruits, as in some bats (Fay 1982, Hildebrand 1974, Rosewear 1965, Sprague 1943).

Other adaptations seen in mammals that regularly suction-feed are a small mouth, a smooth and vaulted palate, and a smooth and round tongue that fits nicely in the palate. These features, in different combinations, are seen in sloth bears, some bats and primates that suck insects or fruit pulp, and in particular in walruses and some other pinnipeds that suck shellfish, squid or fish. These features also typically distinguish humans from apes. Chimpanzees have a larger mouth, a flat and rough palate with more transversal ridges, a flat tongue, and no descended larynx (Hocket 1967, Laitman 1985).

The human sucking adaptations could have been used for fruits and/or for smooth aquatic foods. Humans are able to swallow food underwater, and can also keep their mouths open underwater without swallowing or inhaling water. Feeding underwater requires a fine co-ordination of the lips, mouth, tongue and throat in order to keep the water out of the airways. The human tongue is extremely flexible and is well adapted for manipulating objects within the mouth.

Some of these feeding adaptations might explain why our tongue is able to close the oral cavity at different places, allowing a diverse number of consonants to be produced, for example, at the alveolar, palatal, velar and uvular articulation places.

Concluding remarks

We believe that at least some of these adaptations of the oral cavity and airways were important for the evolution of human speech, and probably came about as a result of our ancestors' increased diving abilities (Diller 2000). The anatomical mechanisms required for human speech, including voluntary breath control and a well-developed ability to control the lips, tongue and mouth, are, in our opinion, the types of adaptations we could expect to see in a creature shifting from a predominantly climbing/wading existence to an existence based more on wading and diving.

This phase of our evolutionary past was probably also when a major period of encephalization occurred, and may have corresponded with a time when our ancestors were in need of a new, or at least modified, form of communication system. After all, traditional primate communicative devices such as smell and body language – with the exception of facial expression – were probably less effective in a semi-aquatic milieu when compared to a non-aquatic one (Morgan 1997).

In conclusion, we believe that the modifications to our ancestors' mouth and airways to allow greater breath control for diving, as well as the restructuring of the tongue and mouth to accommodate a more varied diet, probably coincided with an early stage in the expansion of the human neocortex, in particular the areas that controlled the fine movements of the mouth and throat muscles (Brodmann areas 4 and 44). These adaptations, in our opinion, may have been important precursers to the evolution of human speech, particularly when combined with the already well-developed rhythmical, melodic and duetting abilities of our ancient primate ancestors.

References

Bender R, Verhaegen M & Oser N 1997. Der Erwerb menschlicher Bipedie aus der Sicht der aquatic ape theory. *Anthrop Anz* 55: 1-14.

Broadhurst CL, Cunnane SC & Crawford MA 1998. Rift Valley fish provided brain-specific nutrition for early *Homo*. *Brit J Nutr* 79: 3-21.

Chadwik D 1995. Ndoki, last place on earth. *Natl Geogr* 188: 2-43.

Darwin C 1871. *The descent of man.* Appleton NY.

De Waal F 1997. *Bonobo: the Forgotten Ape.* Univ Calif Press.

Diller K 2000. Why linguists and anthropologists should be interested in the aquatic ape theory for the origin of speech. *Evolution of Language Conference*. April 4-6 Paris.

DuBrul EL 1977. Early hominid feeding mechanisms. *Am J phys Anthrop* 47: 305-320

Fay FH 1982. Ecology and biology of the Pacific walrus *Odobenus rosmarus divergens*. *N Am Fauna* 74: 1-279.

Fernandes MEB 1991. Tool use and predation of oysters by the tufted capuchin in brackish water mangrove swamp. *Primates* 32: 529-531.

Hardy AC 1960. Was man more aquatic in the past? *New Sci* 7: 642-645.

Hildebrand M 1974. *Analysis of vertebrate structure.* Wiley NY.

Hirsch-Pasek K & Golinkoff RM 1996. *The origins of grammar.* MIT CA.

Hocket CF 1967. The foundations of language in man, the small-mouthed animal. *Scient Am* 217: 141-144.

Laitman JT 1985. Evolution of the hominid upper respiratory tradt: the fossil evidence, pp. 281-286 in Tobias ed. *Hominid evolution: past, present and future.* Liss NY.

Morgan E 1997. *The aquatic hypothesis.* Souvenir London.

Negus VE 1962. *The comparative anatomy and physiology of the larynx.* Hafner NY.

Ninkovich D & Burckle LH 1978. Absolute age of the base of the hominid-bearing bed in Eastern Java. *Nature* 275: 306-308.

Puech P-F 1992. Microwear studies of early African hominid teeth. *Scann Microsc* 6: 1083-1088.

Puech P-F, Cianfarani F & Albertini H 1986. Dental microwear features as an indicator for plant food in early hominids: a preliminary study of enamel. *Hum Evol* 1: 507-515.

Rosewear DR 1965. *The bats of West Africa.* Brit Mus London.

Schagatay E 1996. *The human diving response: effects of temperature and training.* Univ Lund.

Schlang G, Jäncke L, Huang Y & Steinmetz J 1995. In vivo evidence of structural brain asymmetry in musicians. *Science* 267: 699-701.

Sponheimer M & Lee-Thorp JA 1999. Isotopic evidence for the diet of an early hominid, *Australopithecus africanus*. *Science* 283: 368-370.

Sprague JM 1943. The hyoid bone region in placental mammals. *Am J Anat* 72: 385-472.

Tobias PV 1998. Water and Human Evolution. *Out There* 35: 38-44.

Verhaegen M 1997. *In den beginne was het water.* Hadewijch Antwerp.

The role of narration in the emergence of human language

Bernard VICTORRI

LTM-CNRS
ENS, 1 rue Maurice Arnoux, F-92120 Montrouge, France
Bernard.Victorri@ens.fr

Introduction

What is at stake in this paper is the emergence of the human species-specific language from an evolutionary point of view. Following Bickerton (1990), I shall assume that pre-human hominids were endowed with a protolanguage without syntax and some other important properties of human language but which was fitted for simple acts of communication about 'here-and-now' perceptible reality, in which pragmatic considerations can compensate for the lack of these properties. I will leave aside the problem of the emergence of protolanguage as such: several authors (Dunbar 1996, Dessalles 1998, among others) have shown that this problem (bringing up the 'Machiavellian Intelligence' paradox - see Byrne & Whiten 1988) can be dealt with from a Darwinian perspective, even if there is no easy solution readily available. The point I want to make in the article deals with the transition from protolanguage to language. I will argue in favour of the emergence of a novel function of communication which could have started the process that led to the acquisition of the very specific properties the human language is endowed with. In this respect, what I suggest belongs to the category of the 'function-centred' scenarios, in which approaches such as Donald (1991) and Knight (1998) can be placed, contrasting with the 'mechanism-centred' scenarios which can be found in Pinker (1994) or in Bickerton (1998).

Language properties and narrative structure

Human language differs from other communication systems by many specific syntactic and semantic features. Many authors focus exclusively on syntax, but semantics deserves as much consideration.

To give only one example, all human languages possess complex systems of markers for expressing aspect (durative, perfective, etc.). The category of aspect is very different from that of tense: aspect refers, not to the time relation between a situation and the moment of its being mentioned in speech, but rather, to how the situation itself is being viewed with respect to its own internal makeup (cf. Comrie 1976). The importance of aspect in human languages (Chinese, for instance, has no grammatical markers for tense, though it does have aspect morphemes) points to the central place of the narrative function of language.

Actually, aspect is not really needed when the goal of the communication is reduced simply to transmission of factual information. It is also worth noticing that there is no trace of aspectual notions in logical languages, even in those dealing with temporal logic. But aspect is absolutely necessary for the purposes of narration: it allows the narrator to present an event from various points of view, depending on the relation between the event and the characters on which the narrator focuses.

As a matter of fact, semantics of tense itself is best explained in terms of narrative mechanisms. As pointed out by Turner (1996:149-154), tense expresses the relation between a temporal focus and a temporal viewpoint, and the temporal viewpoint depends on the choice of the narrator: nothing in language forces the narrator to choose the moment of speech as the temporal viewpoint. This is why the present tense can so easily be used to evoke past or forthcoming events.

Observations of the same kind can be made about other semantics domains, including expression of spatial relations, modalities, animacy and agency, etc. Semantic systems may seem imperfect to a logician or an information theorist, but they are perfect tools for a narrator who wants to present in his own way past or imaginary scenes and events, with no direct link to the 'here-and-now' situation.

At last, syntax as such can be explained in the same framework. Turner (1996:143-148) shows that grammatical constructions, with thematic roles, can be viewed as projection of "basic abstract stories". Embedding and recursion can also be viewed as projections of story nesting, i.e. combining one basic abstract story inside another.

On the basis of such considerations, which fit into the general framework of cognitive linguistics (Lakoff 1987, Langacker 1986, Talmy 1988, Fauconnier 1997, Goldberg 1995, Sweetser 1990, Fauconnier & Sweetser 1996), Turner (1996:160-168) proposes a novel scenario of the emergence of language. He assumes that narrative imagining first arose as an individual cognitive ability, resulting from the increase of hominid general-purpose intelligence. Then, the narrative structure, already present in individual minds, has been projected into the communication system, transforming it into human language.

The problem of the nearly total extinction of archaic Homo Sapiens

Turner's scenario is appealing because it can be put in relation with the last steps of hominid evolution. One of the most puzzling facts about hominisation is the extinction of almost all the archaic Homo Sapiens, such as Neanderthals. According to contemporary theories supported by both paleoanthropology and population genetics (see for instance Lewin 1989, Cavalli-Sforza 1996), all the descendants of Homo Erectus, who dispersed all over the Ancient World one million years ago, and who evolved locally into different groups of archaic Homo Sapiens, disappeared subsequently (some thirty thousand years ago for Neanderthals), except, of course, our species, coming from a small group living in East Africa one hundred thousand years ago. It is hard to understand why such intelligent beings (the brain size of Neanderthals was even slightly larger than ours), well adapted to various environmental conditions, were erased from the earth without any descent. External reasons, such as climatic changes, epidemics, or competition with our own species, are not really convincing: in particular, they cannot explain why this extinction took place all over the Ancient World.

So we must look at endogenous reasons, such as social disturbances (Victorri 1997). At one stage, increase in individual intelligence could have become contradictory with social organisation. It is well known that intelligence is not at all a defence against antisocial behaviour. In animals behaviours that are dangerous for the survival of the species, such as killing siblings or eating offspring, are prevented by instinctive reactions. But one consequence of the great development of the neocortex that characterises hominisation is precisely the control of instinctive behaviour: even if we feel that an action is repulsive, we can carry it out anyway. Killing his brother or his father to replace him as the chief of the tribe, killing and eating children during a lengthy period of food shortage, are examples of intelligent behaviour that ensures a short-term individual well-being, which prove fatal to the group prosperity in the long range. Of course it does not go without conflicts, internal conflict for the one who is ready to act, and social conflict within the group as each member senses and reacts to the forthcoming drama. So such acts could provoke real crises, triggering off uncontrollable violence with disastrous consequences to the life of the group.

Emergence of the narrative function

Then, the hypothesis that can be put forward is that language emerged in such situations of crisis, allowing to avoid them. We can assume that when a crisis was about to break out, most members of the group remembered the preceding crises, adding the recollection of past disasters to the instinctive repulsion. If an individual was able to evoke what happened in the past by his voice and his gestures, he had a chance of winning the support of the whole group and stopping the imminent catastrophe. Expressing what happened in the past was also expressing what could happen again and what should never happen any more. This behaviour would bring a new cohesion to the group, constructing a new collective conscience able to offset individual desires. It would open the path to a new social order, with laws imposed from above by the awareness of belonging to a same group endowed with a collective history. All the myths and religions of humanity rest the fundamental prohibitions upon narrative accounts of origins that put precisely on stage the prohibited behaviours.

Thus, the first use of the narrative function could have consisted in expressing specific social events, creating a totally novel social organisation in the animal kingdom, which permitted our species to control the social disturbances that could explain the extinction of the other archaic Homo Sapiens. Human language, with all its syntactic and semantic properties, would then have stemmed from the needs of narration, leading to what Donald (1991) calls 'the mythic culture'.

We can even assume that language proper was first limited to this particular function, and that during a long period its development was independent (to a certain extent) from the utilitarian protolanguage that our species must have possessed, like the other archaic Homo Sapiens. Afterwards, the extension of the use of language to all aspects of the life could have triggered the “symbolic revolution” that took place some forty thousand years ago (cf. Tattersall 1998).

Conclusion

The hypothesis of the emergence of the narrative function as the starting point of the emergence of language takes three main points into account:

- first, it can explain the unique syntactic and semantic properties of language, specially those semantic properties that make human language different not only from animal communication systems but also from logical and other formal languages.

- second, it is compatible with our current knowledge of the last steps of hominid evolution, including the puzzling extinction of almost all the archaic Homo Sapiens, which contrasts with the great evolutionary success of our species.

- third, it permits to understand the development of a novel level of organisation, specific to our species, in which socio-cultural laws replace, to a large extent, the socio-biological constraints governing all the rest of the animal kingdom.

References

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

Bickerton D., Catastrophic evolution: the case for a single step from protolanguage to full human language, in Hurford et al., 341-358, 1998.

Byrne R. & Whiten A. (eds.), *Machiavellian Intelligence. Social Expertise and the Evolution of Intellect in Monkeys, Apes and Humans*, Oxford, Clarendon Press, 1988.

Cavalli-Sforza L., *Gènes, peuples & langues*, Paris, Odile Jacob, 1996.

Comrie B., *Aspect*, Cambridge University Press, 1976.

Dessalles J.L., Altruism, status and the origin of relevance, in Hurford et al., 130-147, 1998.

Donald M., *Origins of Modern Mind: Three Stages in the Evolution of culture and Cognition*, Harvard University Press, 1991.

Dunbar R., *Grooming, Gossip and the Evolution of Language*, London, Faber & Faber, 1996.

Fauconnier G., *Mappings in Thought and Language*, Cambridge University Press, 1997.

Fauconnier G. & Sweetser E. (eds.), *Spaces, Worlds, and Grammar*, University of Chicago Press, 1996.

Goldberg A., *Constructions - a Construction Grammar Approach to Argument Structure*, University of Chicago Press, 1995.

Hurford J.R., Studdert-Kennedy M., Knight C. (eds.), *Approaches to the Evolution of Language*, Cambridge University Press, 1998.

Knight C., Ritual/speech coevolution: a solution to the problem of deception, in Hurford et al., 68-91, 1998.

Lakoff G., *Women, Fire and Dangerous Things*, University of Chicago Press, 1987.

Langacker R., *Foundations of Cognitive Grammar*, Stanford University Press, 1986.

Lewin R., *Human Evolution*, Blackwell, 2° ed., 1989.

Pinker S., *The Language Instinct*, William Morrow and Co, 1994.

Sweetser E., *From Etymology to Pragmatics*, Cambridge University Press, 1990.

Talmy L., Force Dynamics in Language and Thought, *Cognitive Science*, 12:49-100, 1988.

Tattersall I., *Becoming Human, Evolution and Human Uniqueness*, New York, Harcourt Brace and Co, 1998.

Turner M., *The Literary Mind*, Oxford University Press, 1996.

Victorri B., Débat sur la langue mère, in *Les Langues du monde*, dossier hors série, Paris, *Pour la Science*, oct. 1997.

Just say no: The evolutionary and developmental significance of negation in behavior and natural language.

Chris WESTBURY

Department Of Psychology, P220 Biological Sciences Bldg.,
University Of Alberta T6G 2E1 Canada
E-mail: chrisw@ualberta.ca

> “If we are by definition the animals that talk, we are ipso facto the animals that deny, for as Spinoza and Hegel argue, any linguistic determination directly or indirectly involves a negation.”
>
> L. R. Horn, *A Natural History of Negation*, p. xiii

The concept of negation has been described by those who have grappled with it as “curiously difficult” (Wilden, 1980) and “far from simple and transparent” (Horn, 1989). Negation has been alleged by some to play a fundamental role in the human ability to symbolize. In this paper the role of negation in the evolution and development of symbolization and language is considered in a neurobiological context. The argument has two sections.

In the first section, the problem of negation is defined and described, with emphasis on the well-defined stage at which the human ability to use negation diverges from the analogous abilities of our closest living primate relations.

In the second section, this functional divergence is attributed mainly to differences in cognitive functions whose initial adaptive utility was related to behavioral control. Much evidence suggests that those functions are dependent on the development of a specific region in the human prefrontal neocortex. This is of particular interest in light of the relation that has been proposed between negation and symbolization.

In this abstract, the contents of the two sections are briefly summarized.

What is negation?

It has been independently proposed many times (from Plato to Spencer Brown) that negation is best defined not as *enantion* (contrary) but as *heteron* (other): i.e. that negation is a positive assertion of the existence of a relevant difference. If we define a 'relevant difference' as one which requires an active choice, a framework for ranking the increasing complexity of six forms of natural language negation is apparent. The highest and most complex form of negation- propositional negation- is the only negational form which necessarily pre-assumes fully-developed language production and comprehension, and is therefore of the least interest for those interested in the evolution and development of those abilities. The prior five forms appear in human development in a sequence of increasing complexity.

The first three forms of negation are:

i.) Negation as rejection/ emphasis of rejection of external entities

ii.) Negation as a statement of refusal to stop or start action

iii.) Negation as an imperative

All these forms may be dismissed as variations of a simple 'forced choice'- i.e as requiring the negator to choose a desirable alternative over an alternative that is clearly less desirable. These forms of negation all have analogues in many non-human animals.

The fourth form of negation- failure negation, or negation of a self-generated or planned action- is only a small variation on the third form of negation (imperative negation), in which the negating agent rejected the actions of another, and is also closely related to the the first two forms of negation. However, the differences between the third and fourth form necessitate a number of extremely important changes in the way concept of negation is used. This fourth form of negation is the first form of negation that is elicited without external coercion. More importantly, it is the first form that allows its user to reject an entity that he has himself generated. It is therefore the first form of rejection that can (and, in the case of past or future action, must) operate on a non-existent entity. The use of negation as a comment on ones own unsuccessful or prohibited action is therefore an intermediary case between negation as rejection of a salient external object, message, or action, and negation as rejection of an wholly internal proposition.

Although failure negation marks the beginning of the gulf separating humans from other animals, it is not totally absent in non-human animals. For example, De Waal (1982) describes a chimpanzee who tried to 'erase' a fear grin while it was happening. Note that two reasons the chimp was able to negate this particular behavior is because it was long-lasting enough to be negated while it was happening, and because it was a facial behavior, which made it possible to use the hands to negate it. The negation by the hands of an on-going behavior may be contrasted to the failure negation of a human infant, who uses this form of negation after the negated behavior has occurred. The anecdotal nature of the evidence and the slightly different quality of the human infant’s and the chimpanzee’s negation shows the gulf that is appearing between man and his most intelligent primate relative. There is no consistent situation in which self-generated negation of their own actions is seen in non-human primates.

Scalar negation and the prefrontal cortex

The fifth form of negation – 'scalar negation' – is the first form of negation, which must necessarily be detached: i.e. used against an absent entity. For this reason, it is also the first form of negation that requires a stated predicate. Scalar negation appears very early in normal human children’s language development, and also appears in human children raised without exposure to a formal language model. However, it has no indisputable analogue among other primates.

Since a recognized absence only exists in contrast to a recalled presence, scalar negation is necessarily dependent upon memory. However, an infant does not comment on *every* absence- only absences which make a difference. Therefore, this form of negation is more specifically dependent on memory that enables what is present to shift attention to a stimulus that is (relevantly) not present. When a child reacts to an absence, he is in effect reversing the significance of the stimuli which are present, because it is only in the context of those present stimuli that any absence can be noticed- what is absent doesn’t look like anything. In reacting to an absence, what was previously ground (irrelevant context, in virtue of not being the relevant absent entity) must now be interpreted as figure (relevant context, in virtue of pointing to the relevant absent entity).

The ability to make such internally-controlled shifts of significance has been well studied. There is strong consensus about the primate neurological substrate that underlies it. It is primarily reliant on two closely interconnected regions, the dorsolateral and orbital prefrontal cortex. The evidence implicating these areas comes from localized electrical stimulation, localized dopamine depletion studies, single cell recording studies, and brain imaging studies (see Diamond, 1988, 1991 & Deacon, 1997 for reviews). All these techniques suggest that the prefrontal cortex plays a specific role in choosing the appropriate behavior from a set of possibilities under conditions which require the subject to modulate his current behavior based on changing contexts, especially when those changes require the overcoming of an innate response tendency. Inasmuch as “Attentional development can be broadly characterized as a systematic increase in the child’s ability to override innate tendencies” (Kinsbourne, 1992, p. 261), these cortical regions may be characterized in general terms as controlling attentional shifting.

The appearance of displaced reference in scalar negation is a symptom rather than a cause of the ‘language insight’. The 'language insight' is the result of an incremental increase in the human infant's ability to make actions contingent upon internal markers. Most animals are limited to differentiating only those dimensions in the world that they are born ‘carrying’ or learned dimensions that have direct biological significance and are marked in the environment. Human beings are unique in our ability to construct and project onto the world an enormous number of abstract (displaced) dimensions of relevance. The 'cognitive clay' that we use to construct those dimensions is negation as *heteron:* our ability to formulate and follow rules about how to make the either/or distinctions that underlie the use of what is not present to shift attention.

Clearly many of the distinctions we make are made possible *by* language. The opposite relation holds true for some early forms of negation. Rather than being made possible by language, those forms of negation make language possible, in virtue of their role as necessary precursors to displaced symbolic reference. The fact that these abstract forms of negation have a specific behavioral utility suggests a plausible incremental adaptive pathway that could have spanned the 'adaptive gulf' separating the symbolic power of human language from the poor imitation of those powers that we find in other primates.

Acknowledgments: This paper was written while I was a research associate at the Center for Cognitive Studies at Tufts University. Thanks to Richard Griffin and Elena Nicoladis for valuable comments on earlier drafts.

References

De Waal, F. (1982). Chimpanzee Politics, Power and Sex among Apes. New York, N.Y.: Harper.

Deacon, T. (1997). The Symbolic Species: The Co-evolution of Language and the Brain. New York, N.Y.: W.W.Norton and Co.

Diamond, A. (1988). Differences between adult and infant cognition: Is the crucial variable the presence or absence of language? In: L. Weiskrantz, Ed. (1988). Thought Without Language. Oxford, England: Clarendon Press.

Diamond, A. (1991). Neuropsychological insights into the meaning of object concept development. In:The Epigenesis Of Mind: Essays On Biology and Cognition. S. Carey & R. Gelman, Eds. New York, N.Y.: Lawrence Erlbaum.

Horn, L.R. (1989). A Natural History Of Negation. Chicago, Ill.: The University Of Chicago Press.

Kinsbourne, M. (1992). Development of attention and metacognition. In: S. J. Segalowitz, & I. Rapin, I., Eds. Handbook of neuropsychology, Vol. 7. Amsterdam, Netherlands: Elsevier Science Publishing Co, Inc.

Wilden, A. (1980). Analog and digital communication: On negation, signification, and meaning. In: Wilden, A. System And Structure: Essays In Communication And Exchange. 2nd Edition. London, England: Tavistock Publications.

Interjections, language and the showing/telling continuum

Tim WHARTON

Dept. of Phonetics and Linguistics
University College London
tim@tcody.freeserve.co.uk

Interjections (*ouch*, *oops*, *wow*, *yuk*, *eh*, *oh* etc.) are a fairly broad class of items which have been variously defined. The questions of what and how they communicate, and whether or not they are part of language, have been addressed from two largely dichotomous viewpoints. Ameka 1992, Wierzbicka 1992, Wilkins 1992 treat interjections as encoding rich conceptual structures which are part of the semantics of natural language. Goffman 1981 regards them as falling outside the linguist's domain, analysing them in terms of the socio-communicative roles they play. My aim in this paper is to assess the relative strengths and weaknesses of these two contrasting approaches, and to suggest a new analysis of interjections which preserves the insights of both. This analysis has implications for accounts of the evolution of language.

While Ameka *et al*. are agreed that since they have semantic structure, interjections are part of language, they do not agree on what exactly constitutes an interjection; Goffman too has his own criteria. There is, however, broad agreement from both camps that interjections satisfy the following two criteria: firstly, they are capable of constituting an utterance by themselves in a unique non-elliptical manner; secondly, they express a speaker's mental or emotional state or attitude. There is also agreement that one can distinguish two types of interjections: *primary* interjections (*ouch*, *oops* etc.) – words that cannot be used in any other sense than as an interjection, and secondary interjections (*shit*, *damn* etc.) – words which have an independent semantic value but are often used as interjections. Primary interjections are the main focus of this paper.

For the conceptualists, interjections have " 'real' semantic (i.e. propositional / conceptual) content" (Wilkins 1992: 119). The kind of analysis they propose is exemplified below – (Wilkins' proposed conceptual structure for *wow* (151) (1):

(1) "*wow!*"
I have just now become aware of this something,
that I wouldn't have expected
[or 'that I wouldn't have thought I would become aware of']
This something is much more X than I would have expected,
and this causes me to feel surprised,
and to feel that I could not imagine this something being more X
than it already is now.
I say '/wau/!' because I want to show how surprised (and impressed)
I am feeling right now.

There are four main problems with this approach. Firstly, the communicative content of interjections is highly context-dependent, and an utterance of *wow* seems to communicate something altogether more vague than the elaborate conceptual structure in (1). Secondly, humans use a wide range of para- and non-linguistic behaviours to communicate: facial expressions, gestures, some intonation, for example. The conceptualist approach overlooks the fact that interjections share with these the property of being partly *natural* as well as partly coded. Thirdly, interjections do not contribute to the truth-conditions of sentences that contain them, and this is surprising if they encode fully-conceptual structures. Conceptual representations have logical properties (they can contradict or imply other conceptual representations); one would therefore expect the constituents of these representations to be truth-conditional. Fourthly, intuitions do not support the claim that interjections encode conceptual structure. Whilst one is happy to concede that the italicised expressions in (2) encode the same (or similar) concepts, one is less happy to say the same about those in (3ab), which do not appear to be synonymous in the same way:

(2) Be careful with that *needle* / *hypodermic* / *syringe*.
(3) a. *I feel pain*, what did you do that for?
b. *Ouch*, what did you do that for?

Furthermore, (4a) below intuitively involves a conceptual repetition not evident in (4b):

(4) a. I feel pain, I feel pain.
b. *Ouch*, I feel pain.

Goffman defines interjections, or *response cries*, in functional terms. Some are entirely instinctive, natural reactions – *brrrr*, for example – the purpose of which is to restore some kind of physical equilibrium. Others are less instinctive; a person uttering *oops* on dropping something might do so because it has the effect of “downplaying import and hence implication…of…incompetence” (1981: 102). Goffman also regards swear words (imprecations) as response cries and introduces the notion of a continuum: “response cries such as *eek!* might be seen as peripheral to the linguist’s domain…but imprecations …are more germane, passing beyond semi-word segregates to the traditional material of linguistic analysis” (p.121). However, Goffman’s account fails to address the question of how interjections communicate what they do.

Over the last 30 years philosophers of language and linguists have broadly agreed that there are two distinct kinds of linguistically encoded information. The distinction has appeared in various forms – *truth-conditional* versus *non-truth-conditional* meaning, *describing* versus *indicating* – and is one we might exploit in developing a more adequate analysis of interjections. One version of this distinction is the one made in *Relevance Theory* (Sperber & Wilson 1986, 1995) between *conceptual* and *procedural* encoding. Relevance Theory is an *inferential* account of communication within a cognitive framework. The linguistically encoded content of an utterance (characterised as the output of a modularised language faculty (Chomsky 1995)) is seen as providing no more than a starting point for rich inferential comprehension processes guided by expectations of relevance, cognitively defined: linguistic communication is massively inferential. This model can be contrasted with a *code* model of communication, where communication is achieved simply by coding and decoding. The conceptualist approach is reminiscent of such a code model, in which inference plays a rather minimal role.

The idea behind the conceptual-procedural distinction is that while most words encode concepts, constituents of conceptual representations, the function of certain others is to guide the hearer’s search for the speaker’s intended meaning by constraining the inferential processes that construct or manipulate those conceptual representations. As a first step toward a new account I propose that interjections encode procedural information. Rather than encoding the specific conceptual structure in (1), *wow* might encode a procedure which points the hearer in the direction of attributing to the speaker an attitude of delight or astonishment, without this being encoded. This approach solves three of the four above problems with conceptualist approach. We are left with the problem of how best to characterise the fact that interjections share certain properties of para- or even non-linguistic behaviours: even if we can analyse interjections as encoding procedural information, they are surely not deserving of the same kind of linguistic status as other (non-interjectional) procedural expressions.

We may be able to capture the uncertain linguistic status of interjections (and also the heterogeneous nature of the class) by drawing on Goffman's notion of a continuum. Thus we might propose that interjections are located at various points along a continuum between *showing* and *telling*, where showing is relatively natural behaviour and telling is properly linguistic. Consider an example: to show Jack that she is delighted with a gift, Lily allows him to see her natural reaction, a smile; to tell him, she might utter something like '*it's wonderful*'. To utter an interjection such as *wow* is to communicate delight with a degree of encoding which takes it beyond mere showing, but falls short of telling. I further suggest that the procedural approach might be extended to some of the para- and non-linguistic behaviours discussed above: it follows, therefore, that interjections may encode procedural meaning whether or not they are properly linguistic.

Fundamental to the relevance theory model of communication is the ability of humans to attribute thoughts, intentions and beliefs to one another: the ability to entertain *metarepresentations*, that is, representations of other representations. An act of ostensive communication is achieved by a speaker providing evidence of her intention to inform the hearer of something. In any ostensive act there are two layers of information to be retrieved. The first layer is the information being 'pointed out'; the second is the information that the first layer has been pointed out intentionally.

When Lily smiles, she provides direct evidence of the first layer of information she wishes to convey: she shows Jack that she is delighted with the gift. He can see that she is delighted without necessarily paying attention to her intention to inform him.Lily's linguistic utterance constitutes a very different kind of meaning, however. Here, Jack's recognition of Lily's intention to inform him is of crucial importance. Since the relationship between her utterance and the fact that she is delighted is arbitrary – it is a coded signal – her utterance provides only indirect evidence of the first layer of information. Instead, all the evidence provided by her utterance bears directly on her intention to inform Jack. Seen in this way, the distinction between showing and telling reflects Grice's (1982) distinction between what he termed 'natural' and 'non-natural' meaning, one feature of which was that increasingly complex metarepresentations might be entertained in a communicative act.

Metarepresentation has a psychological function independent of communication, that of predicting the behaviour of con-specifics (Byrne and Whiten 1988), perhaps in such a way as to better integrate social groups (Dunbar 1998). Since the metarepresentational ability could have evolved independently of communication, one might argue that it preceded, rather than followed, the evolution of a language faculty (Sperber (forthcoming)); having some metarepresentational ability, after all, makes inferential communication possible in the absence of language. To individuals already deeply involved in inferential communication a language faculty may have proved adaptive by making the input to evolving inferential abilities more explicit. In the absence of language, however, communication would have involved the use of other behaviours such as gesture, facial-signalling or mime (Donald 1998). On this approach, the evolutionary significance of the showing/telling continuum is clear: showing came *before* telling or saying that. The showing/telling continuum may represent an evolutionary progression from natural communicative behaviour to language proper.

There is evidence to support this claim. Neurological evidence suggests that interjections are associated with phylogenetically ancient sub-cortical circuitry, as opposed to more recent cortical structures implicated in the production of language. The relative phylogenetic histories of the two systems may differ in a manner analogous to that proposed for human vision

(Ramachandran and Blakeslee 1998), where as well as the evolutionarily more recent 'visual pathway' to the *visual cortex*, responsible for the lion's share of visual processes, there exists a phylogenetically primitive pathway to the *superior colliculus*, which mediates unconscious orienting behaviour. This, together with the fact that interjections such as *ouch* remain within the repertoire of certain serious aphasics (Goodglass 1993), demonstrates a clear dissociation between at least some interjections and language proper. The degree to which the showing/telling continuum might be reflected in the neural structures that mediate communicative events needs further investigation, and may be illuminated by increasingly revealing techniques of neural-imaging.

Adopting a broader perspective, this paper supports the view of Jackendoff (1999) who suggests that some sort of "middle ground" needs to be defined between Chomskyans, who, despite advocating an innate language faculty, have in the past tended to "devalue evolutionary argumentation" (p.272), and evolutionists, who have been inclined to deny the possibility of an innate language capacity in the absence of any evolutionary justification. Jackendoff also suggests that interjections might represent "*fossils* of the one word stage of language evolution" (p. 273). The proposals here develop this suggestion.

References

Ameka, F. (1992) Interjections: The universal yet neglected part of speech. In *Journal of Pragmatics* 18, 101-118.

Byrne, R. and A. Whiten (eds.) (1988) *Machiavellian Intelligence: Social Expertise and The Evolution of Intellect in Monkeys, Apes and Humans*. Oxford: Clarendon Press.

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

Donald, M. (1998) Mimesis and the executive suite. In Hurford, J., Studdert-Kennedy, M and C. Knight (eds.) *Approaches to the evolution of language*. Cambridge: Cambridge University Press.

Dunbar, R. (1998) Theory of mind and the evolution of language. In Hurford, J., Studdert-Kennedy, M and C. Knight (eds.) *Approaches to the evolution of language*.Cambridge: Cambridge University Press.

Goffman, E. (1981) *Forms of Talk*.Oxford: Blackwell.

Grice, H. P. (1982) Meaning revisited.In *Mutual Knowledge*.London: Academic Press.

Jackendoff, R. (1999) Possible stages in the evolution of the language capacity. In *Trends in Cognitive Science*, Vol. 3, No. 7.

Goodglass, H. (1993) *Understanding Aphasia*.New York: Academic Press.

Ramachandran, V. and S. Blakeslee (1998) *Phantoms in the brain*.London: Fourth Estate.

Sperber, D. (forthcoming) Metarepresentations in an evolutionary perspective.To appear in S. Davis & D. Sperber (eds.) *Proceedings of the Tenth Vancouver Cognitive Science Conference*.

Sperber, D. and D. Wilson. (1986/1995) *Relevance: Communication and Cognition.*Oxford: Blackwell.

Wierzbicka, A. (1992) The semantics of interjection. In *Journal of Pragmatics* 18, 159-192.

Wilkins, D. (1992) Interjections as deictics. In *Journal of Pragmatics* 18, 119-158.

The contribution of speech perception for the evolution of the capacity for phonological systems in *homo sapiens.*

Henning WODE

Englisches Seminar der Universität Kiel
Leibnizstr. 10
24098 Kiel, Germany
OfficeLing@anglistik.uni-kiel.de

ABSTRACT

Attempts to reconstruct the evolution of the capacity for phonological systems in *homo sapiens* require consideration of the anatomy of the vocal tract (e.g. Liebermann 1975, 1984, 1990, Kent 1992), perception as well as the interaction between the two. In this paper I will review the impact of the auditory system on these issues along lines originally suggested by Kuhl (e.g. Kuhl/Miller 1978, Kuhl/Padden 1982, 1983, Kuhl 1988, 1993). I will argue that the structure and the functioning of the auditory system as required for natural human languages to function the way they do, antidates the evolution of the vocal tract as required for speech by quite a margin. The data to be considered include speech perception in human neonates, pre-speaking infants, children and adults, as well as non-human animals. The argument is focussed primarily on distinctive features, their typology, and the ability of human beings to categorize the sound wave into segments. It is argued that the distinctive features are based on the heightened sensitivities of the auditory system. They restrict the universal set of sound contrasts that are used for phonemic purposes, i.e. these auditory sensitivities constitute the biological basis for the distinctive features; and categorical perception constitutes the biological basis for the phoneme.

The first step in the argument is to point out that the distinctive features and phonemes are based on the functional potential of the auditory system and not on the anatomy of the vocal tract. That is, perception is primary and articulation secondary. The crucial evidence for this point is the typology of distinctive features and the fact that there is nothing in the anatomy of the vocal tract to explain why, for example, only certain places of articulation are used for phonemic purposes and not others.

The second step is to show that human neonates and infants can do two things: (a) They have categorical perception; and (b) during the first half of their first year of life they can detect those sound contrasts that are used for distinctive purposes in the languages of the world, including those that do not occur in their native language (e.g. Streeter 1976, Trehub 1976, Lasky et al. 1975, Werker/Tees 1984, Best et al. 1988, Best 1994). Apparently, these ablities have an innate basis, and can, therefore, serve as the biological basis for the functional potential of phonological systems.

Two objections are possible here. One is that categorical perception is a phenomenon derived from the research on synthetic speech and it needs to be shown to also apply to non-synthetic, i.e. natural stimuli. The second objection is that so far we are far from having checked all or most of the sound contrasts that are utilized phonemically in the languages of the world. The first objection is valid. However, there are some experiments that have relied on natural stimuli,

such as the dental vs. retroflex stop contrast of Hindi (Werker/Tees 1984). Moreover these findings have been replicated for synthetic stimuli (Werker/Lalonde 1988). Although these studies were not on categorical perception, they do show that the perception does not neccessarily differ for natural and synthetic stimuli.

The third step is concerned with the animal evidence. The earlier research on chinchillas, monkeys, Japanese quail, and European starlings (e.g. Kuhl/Miller 1978, Kuhl/Padden 1982, 1983, Kluender et al. 1987, Kluender et al. 1997 etc.) had shown that these animal species had categorical perception and that their points of heightened auditory sensitivity along a given acoustic dimension parallelled those of human beings. This kind of evidence allowed for several conclusions. First, categorical perception and the points of heightened auditory sensitivity are not unique to *homo sapiens*, hence they are not species-specific. It followed, second, that they are not language-specific either. Third, since the same properties occur in such diverse species including *homo sapiens*, it is reasonable to assume that these species share a common ancestor. Fourth, given the differences in the anatomy among the present-day species, the development of their anatomical differentiation needs to be taken as postdating the shared auditory system by a considerable margin. Fifth, it is not quite clear at the present time whether it can also be claimed from the above evidence that perception has had an impact on shaping the vocal tract in such a way that the latter was brought into line to closely match perception in the sense of Stevens' notion of the quantal nature of speech (Stevens 1972, 1989).

Whereas the earlier animal evidence referred to above suggested that the points of heightened auditory sensitivity were the same for humans and non-humans, there is a recent study by Sinnott/Brown 1997 that indicates that that may not neccessarily hold for all sound contrasts used for phonemic purposes in the languages of the world. Sinnott/Brown studied the perception of the /r/-/l/ contrast in American English by adult native and non-native speakers and Japanese macaques using a synthectic continuum ranging from /ra/ to /la/. The contrast was discriminated by both groups of subjects. But the category boundary was located in different areas of the cotinuum. With the monkeys it was shifted more towards the /r/ end of the continuum, with the humans it was located more towards the /l/ end.

Obviously, it is no longer correct to assume that animals and humans share all points of heightened sensitivity. Apparently, this is correct only for some contrasts, notably, VOT in stops, place contrasts like labial, dental, velar, the (some?) vowels, and probably others. Nonetheless, it still seems justified to maintain the view that natural human languages and the capacity for them have evolved in a natural way from common, i.e. non-language-specific properties of the organisms according to Neo-Darwinian principles.

The co-evolution of language and emotion

Robert WORDEN

Charteris Ltd.
rpw@charteris.com

Summary

Emotions and language have important features in common:

- They both depend on complex assessments of the current social situation
- They both depend on a theory of mind, to assess how others perceive that situation
- They both serve a purpose of social communication
- They both depend on rapid unconscious calculations
- They are both learned and culturally transmitted

I propose that language and emotion are facets of the same symbolic social intelligence faculty, which we share with other primates. At some time within the last four million years, there has been strong selection pressure to increase the power of this faculty to support rapid fluent language. This increase in the horsepower of social intelligence, and the growth of the theory of mind enabled by language, have had major impacts on our emotional reactions, not always for the best.

Origins of language and emotion

In a previous paper at this conference (Worden 1998) I proposed that the language faculty has its origins in primate social intelligence. Evidence in support of this hypothesis is:

- Social intelligence requires us to form structured, composite representations of social situations. These representations are well suited to encode language meanings.
- Social intelligence takes inputs of all sensory modalities, and language meanings can refer to all sensory modalities.
- The primary usage of language is for social communication, which requires social intelligence
- Language requires a theory of mind, which is a facet of social intelligence.

It can be similarly argued that human emotion depends primarily on social intelligence: Apart from a few ‘primitive’ emotions like fear, many emotions serve a purpose of preparing us physiologically for social interactions with our peers. Emotions are driven by a cognitive assessment of our current social situation, depending on sense data of all modalities (Frijda 1986; Oatley 1992). Many emotional reactions produce effects (e.g. in facial expression or posture) which communicate social information to our peers. Finally, complex emotional reactions (e.g pride, shame, embarrassment) depend not just on our assessment of an actual social situation, but also on our assessment of how others see that situation (e.g. Harre 1986, Lutz 1988); they depend on our theory of mind (Carruthers and Smith, 1996).

The common antecedent of human language and human emotion can be seen in primate social behaviour. This largely consists of emotional reactions, driven by social assessments of interactions with their peers.

We can build a computational model of primate social intelligence which then extends to be a computational model of language. In this model, primates represent social situations internally by tree-like feature structures. A representation as complex as a tree structure – rather than, for instance, a simple list of attributes – is needed to capture the composite and structured nature of primate social situations, which primates are sensitive to.

Similarly, the meanings of sentences can all be represented by tree-like feature structures of nodes, links and slots. This is commonly done in computational models of language (Shieber 1986). Doing so enables the feature structures to represent the complex structured meanings of language, and to capture empirical findings about language meanings.

We assume, then, that both language meanings and social situations are represented by the same tree-like feature structures, or 'scripts', in the brain. These feature structures are the core of primate social intelligence, and are therefore the drivers of many primate emotional reactions. We can describe a typical primate social/emotional reaction in terms of this feature structure model:

- A primate observes the current social situation and encodes this as a feature structure (e.g. he observes an ally being threatened by another of superior rank).
- By unification with other learned feature structures, which describe the regularities of social life, he or she works out the implications of the current situation for himself (e.g. it may threaten physical danger, or a decrease in rank)
- These implications trigger emotional reactions such as hormone changes, and changes in posture and expression, which both communicate to others and prepare the individual for any required actions (e.g. appeasement).
- This assessment lasts for a long enough period (several seconds or minutes) to affect the animal's physiology appropriately and to communicate effectively with others by posture and expression

A key concern in primate social life is the achievement of higher rank in the group, as this greatly enhances reproductive chances. Many primate social/emotional reactions are triggered by situations with potential impact on the primate's rank, and emotional reactions have evolved to improve the chances of achieving high rank (de Waal, 1982; Byrne & Whiten 1988).

Just as many non-human primate social/emotional reactions are triggered by situations with potential impact on rank in the group, so are human emotions. For our assessment of our own rank, we may substitute the phrase 'self-esteem'. Situations with a bearing on our self-esteem are powerful releasers of emotion. Our self-esteem depends largely on what we think others think about us; so our theory of mind, which enables us to partially understand what others think about us, is closely linked to our emotional reactions. Theory-of-mind knowledge, like other social knowledge, can be represented in tree-like feature structures.

Human emotional reactions have evolved from this primate model and are based on it – but the impact of language and a theory of mind has radically changed them, as discussed in the next section.

The evolutionary impact of language on emotion

The emergence of language has coincided with – and perhaps caused - a great increase in the size and power of the human brain, relative to other primates (Deacon 1997). In order to produce and understand language, the human brain must perform several complex calculations per second. This is probably much faster than the rates of computation required for most primate social intelligence. Also the feature structures required for modern language are probably more complex than those used by most primates.

As we know from introspection, language is capable of serving as a kind of scaffolding for extended trains of conscious thought – which may reach conclusions very different from the simple automatic social inferences made by other primates. Language enables us to be consciously aware of a line of reasoning, to recall and modify it.

Whatever selection pressures drove the emergence and refinement of language, therefore, they have had important side-effects on human social intelligence:

1. They have increased the complexity of the feature structures we can use to represent social situations
2. They have increased the speed with which social inferences can be made
3. They have led to a powerful theory of mind
4. They have enabled conscious, language-based trains of thought

What has been the impact of these major changes on the human emotional system? As for other primates, our emotional reactions are driven by assessments of our social situation - using the same set of internal representations (feature structures) and computations. Increasing the horsepower of this social computation engine must have a powerful impact on those emotional reactions.

I suggest that the first impact of these changes may have been to take a pre-existing system of emotional reactions which had evolved over many millions of years, and which was well-matched to the task it performed, and to de-stabilise it – in that the new speed and power of social cognition was no longer well-matched to the task of driving emotional reactions.

A typical primate brain can form a simple feature structure representation of the current social situation, and can (probably slowly) perform various simple inferences on it which lead to simple emotional reactions. However, since the emergence of language the human brain has been capable of doing more, and doing it faster. This increase in speed and complexity threatens to undermine the previous primate emotional mechanisms, of a relatively slow and simple social computation leading reliably to an appropriate emotional reaction.

If the selection pressure leading to language (and to increased social intelligence) was a strong one, then such a de-stabilisation of emotions will have at least started to occur as the changes required for language began to take place. This de-stabilisation of emotions would lead to decreased fitness in the individuals concerned. In the first instance, there would then be an evolutionary ‘tug of war’ between the two selection pressures, for language proficiency and for stable appropriate emotions.

I suggest that a straight evolutionary tug-of-war may not have taken place, or in any case did not go on for long, because our minds evolved so as somehow to re-stabilise (if only incompletely) our emotional reactions. We have evolved mechanisms which somehow ‘fixate’ our emotional reactions into a comparatively small number of more stable states, leading to sustained emotional reactions, in spite of the rapid possible oscillations of different social inferences and conscious thoughts (see e.g. Stewart and Joines 1987). The cognitive model of emotion is not yet well defined enough to be able to know much about these possible re-stabilisation mechanisms.

The cost of this de-stabilisation and re-stabilisation has perhaps been rather high – resulting in the apparently irrational nature of many of our emotional reactions, with their ability to override conscious thought. Possibly the best way to understand our emotions will be to understand how they share common cognitive mechanisms with the language faculty (which is better understood and easier to study) – and then to understand how the evolution of language has forced the evolution of our emotions. From where we are now, this may be a long and difficult research programme – but given the great potential value of understanding the emotions, it is worth the attempt.

References

Byrne, R. and Whiten, A., (eds) (1988) Machiavellian Intelligence: social expertise and the evolution of intellect in monkeys, apes and humans, Clarendon Press, Oxford

Carruthers, P. and Smith, P.K. (eds) (1996) Theories of theories of mind, Cambridge University Press, Cambridge UK

Deacon, T. (1997) The Symbolic Species: the co-evolution of language and the human brain, Penguin, London

Frijda, N. H. (1986) The Emotions, Cambridge University Press, Cambridge, UK

Harre, R. (1986) (ed) The Social Construction of Emotions, Blackwell, Oxford

Lutz, C. A. (1988) Unnatural Emotions: everyday sentiments on a Micronesian atoll and their challenge to western theory, University of Chicago Press

Oatley, K. (1992) Best Laid Schemes: the psychology of emotions, Cambridge University Press, Cambridge, UK

Shieber, S. (1986) An Introduction to Unification-Based Approaches to Grammar, CSLI, Stanford CA

Stewart, I. And Joines, V. (1987) TA today: a new introduction to transactional analysis, Lifespace, Nottingham

De Waal, F. (1982) Chimpanzee Politics: power and sex among apes, Johns Hopkins, Baltimore

Worden, R.P. (1998) The evolution of language from social intelligence, in Approaches to the Evolution of Language, J. R. Hurford, M. Studdert-Kennedy and C. Knight (eds) Cambridge University Press, Cambridge.

A protolanguage with no declaratives and no names

Alison WRAY

Centre for Language and Communication Research
Cardiff University
wraya@cf.ac.uk

In previous work (Wray 1998, 2000) I have explored the idea that socially sophisticated hominid populations could have communicated in complex ways without grammar, by using a holistic protolanguage: a phonetically-articulated system of discrete, agrammatical messages. This paper tests that scenario, by considering the ways in which such a protolanguage would restrict the expressive scope of its speakers, and the effects that this could have on the nature of hominid life and on the timing of further evolution. I shall demonstrate that, specifically, a holistic system is incompatible with the coining of names for things and people, and with the extensive use of declarative statements. The absence of these two features of referential expression would inhibit information exchange and could explain the long period of technological and cultural stagnation between 1.4 million and 100,000 years BP (Mithen 1996:116).

The primacy of holistic processing

It seems reasonable to look for a role for holistic processing in protolanguage, since, not only did protolanguage presumably develop *out* of a holistic communication system of some sort, but it also developed *into* a communication conglomerate which, besides words and grammar, continues to support holistic processing (Wray 1999, Wray & Perkins 2000). Recent research on the nature and extent of formulaicity in human language (for reviews see Wray 1999, forthcoming) has revealed it to be a common and essential feature of our every day communication. Formulaic language has a number of communicative functions, of which almost all directly correspond with the manipulative socio-interactional functions observed in ape communication (Reiss 1989). Our socio-interactional formulaic expressions are directed towards the physical, emotional and perceptual manipulation of the hearer: commands, requests and warnings can all be expressed using fixed phrases; politeness is associated with fixed forms; we use in-phrases, chants and songs, etc. to indicate our group identity and so on (see Wray 1998, 2000, forthcoming, Wray & Perkins 2000 for a full exploration of these and the other functions of formulaic language). A significant advantage to us of using such fixed forms for common interactional functions, rather than constructing a novel sequence to express the same idea, appears to be the ease with which they can be recognised and decoded by hearers who share our language variety. The easier the decoding, the greater the likelihood of the hearer reacting in the desired way.

What a holistic protolanguage might look like

In this model, protolanguage messages are semantically complex and agrammatical. They are holistic, which means that a complete message is uniquely associated with an arbitrary form, not made out of smaller recombinable units of meaning. For instance, we might say that *tebima* means *give that to her* and *mutapi* means *give that to me*. Note that “there is no phonological similarity between sequences with similar meanings, because they are holistic. There is no part of *tebima* that means *give* or *her*. Simply, the whole thing means the whole thing” (Wray 2000).

No names

Much has been written over the years about the power of naming, and it has seemed logical to assume that the forerunners of language included amongst their attributes, from the earliest stages, names for things and people. Hypothetical protolanguages are characteristically depicted as heavy on referential items when still lacking any means for expressing the complex relationships between them. But in a holistic protolanguage there is no place for naming, and the use of pronouns in the glosses of *tebima* and *mutapi* above is significant. If a holistic system like this tried to be too precise, such as referring to each member of the group and to each common object individually, it would soon run into difficulties. Without recombinable constituents, you would need an entirely new sound-string for each possible utterance about each person and thing – different ones for *give the stick to Mary*, *give the stick to Joyce, give the stone to Mary* and so on. Although highly precise in meaning, each individual string would have such a low functional load as to be rarely heard and used. This would make it difficult to remember, and once forgotten it could not be reconstructed by rule since there are no rules or individual words. Furthermore, this level of precision would generate so many strings that differentiating them would require an excessively large pool of phonetic variants in lengthy combinations.

These difficulties are resolved by using the same string for many speech events. By grouping together, say, all the female persons into one set, a single string *give that to her* can stand for *give that to Mary, ... to Joyce, ... to Brenda* and all the rest. Even now, the inventory of such strings will still be large, and the need to differentiate them could explain why we possess today a capacity for making articulatory distinctions far in excess of what we need for our modern analytic language of small, recombinable units (Wray 1998, 2000). There is a price to pay, however, for not being able to express the specific identity of individuals or objects. If you say *give that to her*, how will the hearer know to what, or whom, you are referring? For the purposes of disambiguation, such messages must be supported by indicative gesture, eye-gaze, or whatever. In short, although the speaker and the hearers all perfectly well know the difference between Mary, Joyce, Brenda and the other females, and between stones, sticks, meat, leaves and other portable objects, they are handicapped in their ability to express linguistically these fine distinctions by the limitations in the expressive potential of the language. The shortfall is made up using holistic paralinguistic strategies which we still put to the same purposes today.

No declaratives

Hot on the heels of having no names comes the effective absence of declaratives in all but one limited context. Even though there is no cognitive restriction on their wider use, there is a linguistic one. Because new items in the language can neither be created nor decoded by rule, the cost of introducing them is high, and there must be a good reason for bothering. Manipulative messages such as commands and threats offer a practical pay-off, and also are relatively easy to understand using pragmatic cues or by observing the behaviour of speaker and hearer. In contrast, the function of declaratives is to change others' knowledge and, for them to be useful, the speaker must be imparting some information that the hearer does not possess, and could not have deduced. If you cannot specifically name, the effectiveness of declaratives is drastically compromised. Since an utterance that could refer to any of several referents is incomprehensible unless the speaker can indicate who or what is meant, the referent must be present[1]. Yet if the referent is present, any information that is observable need not be explicitly imparted. A general message such as *look at her* will be more economical and just as effective

---

19 Unless an individual is symbolically represented, in absence, by a present object. This has extensive implications for the assumed state of hominid cognition, culture and, possibly, (religious) belief, and is well beyond the scope of this paper.

as separate messages for *she is tall* and *she is limping*. This leaves one condition, where the referent is present but the new information about him/her/it is hidden (e.g. *she is sick; she is sad*). The imparting of privileged information is a type of *gossip* (Dunbar 1996, 1998), oiling the wheels of social cohesion and adjusting the relative social status of speaker and hearer. It is closely allied with the other manipulative uses of language which are central to the continuity of holistic processing from primate to human. In the case of gossip only, then, just as Dunbar (1998:105) suggests, there is both the interactional motivation and the linguistic means for using declaratives.

This position offers a new perspective on Tomasello & Camaioni's (1997) exploration of declaratives, which they, also, view as a special case in language evolution. Their stance differs from the one in this paper, since they consider that the communicative ability to express declaratives is held in check by an absence of the cognitive means, whereas I am proposing the reverse. A problem for their position is that declarative gestures are virtually absent in primates in the wild, but present in some apes taught a human-like language system. This undermines their claim that declaratives are beyond the cognitive capacity of apes. However, it is consistent with the cognitive and expressional functions that underlie the use of declaratives being dissociable, so that untaught apes, like the hypothesised protolanguage users, are prevented from making declarative gestures/utterances not because they lack an appreciation that others have separate mental representations, but because, lacking naming, they have little if any use for declarative gestures. This is what Gómez et al (1993: 419) suggest, though for different reasons: declaratives are a skill for which apes naturally have the *potential*, but normally no *motivation*. Meanwhile, apes that are taught to name gain both the means and the motivation for using declaratives.

Naming and the appearance of grammar

The holistic protolanguage described above is a self-supporting, stable edifice, whose constraints would stifle its own further evolution, perhaps for thousands of generations: specific naming is unsustainable; without naming, declaratives have almost no purpose; without declaratives, information exchange is largely impeded; this minimises technological and cultural innovation, rendering naming unimportant. It is possible to imagine how the combination of strong social hierarchies, themselves very well served by holistic utterances, and this stranglehold on referential expression, could considerably postpone the breakthrough into full human language, relative to the independent evolution of our modern cognitive and intellectual abilities.

By postponing the linguistic expression of the 'naming insight' until hominids are on the very brink of anatomical and mental modernity, there is no need for us to postulate a gradual development of grammar simply in order to accommodate major neurological changes. Provided it is conceivable that our modern mental capacities arose independently of language, naming is unleashed into a powerful cognitive forum that can immediately exploit referentiality by creating argument structure out of the juxtaposition of a word and a holistic utterance, and by segmenting holistic utterances to 'identify' new words and structures *post hoc* (see Wray 1998, 2000). Importantly, the significance of naming in this context is not the coining of an arbitrary phonetic sequence with a consistent meaning, for that hurdle was overcome long ago, when the holistic messages were first coined. What is different about naming is the size of the meaning unit: it refers to a person or thing (and soon an idea, action, attribute, etc) rather than a proposition.

Taking on board all the provisos concomitant with the view that "ontogeny epitomises (although it does not recapitulate) phylogeny" (Studdert-Kennedy 1998:172), it may be noted that first language acquisition follows pretty well the same path at the one outlined above. First, infants use holistic noise and gesture for manipulative and social purposes. They then pick up contours and articulatory features of complete utterances, which they produce as holistic phonologically-governed message-carriers. When they gain the naming insight they become referential in gesture and word, and proceed to segment the holistic strings into smaller meaning units that can be juxtaposed to create grammatical relationships. However, they continue to employ a holistic processing strategy for various ends, including the gamut of socio-interactional functions associated with the physical, emotional and perceptual manipulation of others (Wray 1999, forthcoming). The significance of this similarity between ontogeny and phylogeny lies primarily in the superimposition of naming onto an existing holistic system, and the way in which referentiality and compositionality team up with formulaicity to maximise the advantages novel and routine expression (see Wray 1998, 2000, forthcoming; Wray & Perkins 2000).

This scenario offers several explanatory advantages. It accommodates a slow development of long-term memory storage and of our excessive articulatory prowess, under pressure from a greedy communicational system that takes no hostages (if you can't distinguish messages, or remember them, there is no grammar to fall back on). It provides a long period of phonetically-articulated complex social interaction without a linguistic facility for information exchange, at a time compatible with the technological and cultural stagnation identified from the palaeontological record. It holds back the onset of grammar to a point where it can develop quickly to its present state, while never entirely usurping the ancient communicative roles of holistic messages, thus avoiding scenarios in which simple grammars are of questionable use to speakers with complex communicational needs. It convincingly mirrors the patterns of child language acquisition, particularly in depicting the holistic system as more primitive. And it provides a link from the holistic approach to communication used by primates, and presumably by our earliest ancestors, through to our use of formulaic language as a significant and necessary supplement to our grammatical system today.

References

Dunbar, R. 1996. Grooming, gossip and the evolution of language. London: Faber.

Dunbar, R. 1998. Theory of mind and the evolution of language. In Hurford et al (eds.), 92-110.

Gómez, J.C., Sarriá, E. & Tamarit, J. 1993. The comparative study of early communication and theories of mind: ontogeny, phylogeny, and pathology. In Baron-Cohen, S., Tager-Flusberg, H. & Cohen , D.J. (eds.) Understanding other minds. Oxford: Oxford University Press, 397-426.

Hurford, J., Studdert-Kennedy, M. & Knight, C. (eds.) Approaches to the evolution of language. Cambridge: Cambridge University Press.

Mithen, S. 1996. The prehistory of the mind. London: Thames & Hudson.

Reiss, N. 1989. Speech act taxonomy, chimpanzee communication, and the evolutionary basis of language. In Wind, J., Pulleybank, E.G., Grolier, E.De & Bichakjian, B.H. (eds.) Studies in language origins, vol 1. Amsterdam: John Benjamins, 283-304.

Sinclair, J.M. 1991. Corpus, concordance, collocation. Oxford: Oxford University Press.

Studdert-Kennedy, M. 1998. The particulate origins of language generativity: from syllable to gesture. In Hurford et al (eds.), 202-221.

Tomasello, M. & Camaioni, L. 1997. A comparison of the gestural communication of apes and human infants. Human Development 40:7-24.

Wray, A. 1998. Protolanguage as a holistic system for social interaction. Language & Communication 18:47-67.

Wray, A. 1999. Formulaic language in learners & native speakers Language Teaching 32:213-31.

Wray, A. 2000. Holistic utterances in protolanguage: the link from primates to humans. In Knight, C., Studdert-Kennedy, M. & Hurford, J. (eds.). The evolutionary emergence of language. Stanford, CA: Cambridge University Press.

Wray, A. forthcoming. Formulaic language and the lexicon (working title). Stanford,CA: Cambridge University Press.

Wray, A. & Perkins, M.R. 2000. The functions of formulaic language: an integrated model. Language & Communication 20:1-28.

Can the Baldwin effect really explain the evolution of the LAD?

Hajime YAMAUCHI

Department of Linguistics
University of Edinburgh
hoplite@usa.net

Despite the considerable differences of surface structures of various languages, it can be agreed that all natural languages are equally complex; perhaps the most complex system in any cognitive faculties. The question arises, then, how any child, wherever they may be in the world, can begin to acquire such a complicated system. Everyone intuitively knows linguistic input plays a crucial role in language acquisition. Linguistic input, however, employed to construct knowledge of a language is often ill-formed, incoherent, and most importantly, insufficient (Chomsky 1965). Thus claims have been made that the process of language acquisition is neither completely a process of postnatal learning nor a product of innately fully specified linguistic knowledge. Chomsky's original formulation of the nature of the language acquisition device (LAD) and its core theory –Principles & Parameters theory- (Chomsky 1981) were introduced to solve this embarrassing complication of language acquisition. This strategy tries to come to terms with this complication by putting forward both aspects of language acquisition as being equally significant.

Recent surveys in the field of computational simulations reincarnate a more-than-100-years-old argument in evolutionary study. In 1896, James Mark Baldwin proposed "a new factor in evolution" (Baldwin 1896, Morgan 1896, Osborn 1896, Waddington 1942, Turney, Whitely, and Anderson 1996). He assumed that if an individual is capable of acquiring an adaptive feature postnatally, addition of such a learning process in the context of evolutionary search potentially changes the profile of populational evolution. In a nutshell, the Baldwin effect is an interaction between evolution and learning, where "a behavior that was once learned may eventually become instinctive" (Turney, Whitley, and Anderson 1996). The possibility of this learning-guided evolution has been repeatedly attested in computer simulations by a number of researchers (Hinton & Nowlan 1987, Maley 1996a, 1996b, 1997, Nolfi, Elman, and Parisi 1994, French & Messinger 1994 *and more*).

It has been a popular idea that the Baldwin effect is a crucial factor of the evolution of language (*e.g.* Pinker & Bloom, 1990, Briscoe 1997). The learning-guided evolution scenario possibly provides a strikingly attractive solution to a longstanding problem. Preliminary studies suggest that language evolution is out of the scope of natural selection mainly because of its dysfunctional nature. For those researchers, language evolution is a consequence of exaptation or a big leap in evolution (Newmeyer in preparation, Piatelli-Palmarini 1989). This no-intermediate scenario would be, however, explicable by natural selection when it is guided by learning since learning can bridge the gap as Hinton & Nowlan showed (Hinton & Nowlan 1987). There is a further advantage of the Baldwin effect in the evolutionary study of the LAD. Its combination of genetically hardwired features and postnatal learning processes are perfectly compatible with Chomsky's P&P theory. Together with its "genetic assimilation" process (Waddington 1975), the Baldwin effect may shed a light on the nature of the current relationship between innateness and postnatal learning in language acquisition.

However, it has been a matter of concern for long time that a straightforward representation of the relationship between genotype and phenotype is unrealistic, especially when the expressed phenotype is beyond basic biological expression. It is unlikely that properties of a certain higher cognitive ability are partitioned by independent genes. It might be more plausible to consider that multi-contribution of genes to their expressions or cascade-like reactions of single gene's expression shape the actual LAD. (Waddington 1942, Deacon 1997, Newmeyer in preparation).

My study examined the real possibility of the effect on the evolution of the LAD under more genetically realistic circumstances based on Turkel's simulation (Turkel, to appear)

First, a complete replication of Turkel's simulation (To appear) was tested. A population of agents with genetically represented principles and parameters was evolved on the basis of a basic genetic algorithm. Thus the evolution of such agents reflects the evolution of the LAD itself. At the end of the simulation, as Turkel revealed, all agents were converged into a single genotype. The unified genotype is randomly determined in each run. Regarding the dysfunctional aspect of language, this randomeness is significant. Each configuration was attained regardless of any external factors; only the dynamic aspect of communication within the community contributes to this arbitrariness. This has implications against the "anti natural selection" theories in language evolution (the "no external reality" theory; Fodor 1989). More importantly, however, the Baldwin effect was observed in this simulation; the number of plastic genes is notably decreased at the end of each run (Fig. 1).

Therefore, agents were successful to establish communication with relatively small consumption of learning trials. This result might be compatible with rapid acquisition of a natural language in real world.

Then, to test the effect of a more complex relationship between genotypes and their phenotypic expressions in Turkel's simulation, I incorporated Stuart Kauffman's NK-Landscape model (Kauffman 1989). In this model, two or more genes express one feature of the phenotype based on a randomly generated look-up table. This means that our model embodies epistasis and pleiotropy in genes. The modification also puts genotypes and phenotypes on completely independent strata while Turkel's original simulation partially conflates these two representations. Effectively, this modification blurs the relationship between genes and their phenotypic expressions. The degree of abstractness is controlled by the value of K that specifies how many genes are needed to express one feature of the phenotype. The results were remarkable; although convergence toward a single genotype was still observed, subsequent emergence of the Baldwin effect was severely suppressed under these circumstances. In the worst case, maximum value of K, no plastic genes were replaced by fixed genes even after 200 generations (Fig. 2).

Finally the results of the simulation require reconsideration of the evolutionary scenario of the LAD. Either it has traced a completely different evolutionary path or the LAD is equipped with an extremely robust learning mechanism so that even with high plasticity it can learn a language without failure.

References

Baldwin, J. M. (1896) "A New Factor in Evolution" The American Naturalist 30.

Briscoe, T. (1997) "Language acquisition" Unpublished manuscript. Submitted.

Chomsky, N. (1965) "Aspects of the Theory of Syntax" MIT Press.

Chomsky, N. (1981) Lectures on Government and Binding: The Pisa Lectures. Foris, Dordrecht.Deacon, T. (1997) "The Symbolic Species" Penguin Press.

Fodor, J. (1989) "Learning the Periphery. In R. Matthew and W. Demopoulos, ed. Learnability and Linguistic Theory. Kluwer.

French, R., and Messinger, A. (1994). Genes, phenes and the Baldwin effect. In Rodney Brooks and Patricia Maes (eds.), Artificial Life IV. MIT Press.

Hinton, G. C. & S. J. Nowlan (1987) "How Learning Can Guide Evolution" Complex Systems 1.

Kauffman, S. (1989) "Adaptation on rugged fitness landscapes" In D. L. Stein, ed. Lectures in the Sciences of Complexity,1. Addison-Wesley.

Mayley, G. (1996) "Landscapes, learning costs and genetic assimilation", Evolutionary Computation, 4 (3).

Mayley, G. (1996) "The evolutionary cost of learning" In Maes, P., Mataric, M.,

Mayley, G. (1997) "Guiding or hiding" In the Proceedings of the Fourth European Conference on Artificial Life (ecal97). P. Husbands and I. Harvey (eds.).

Morgan, C.L. (1896). On modification and variation. Science, 4.

Newmeyer, F. (in preparation) "On the Reconstruction of 'Proto-World' Word Order"

Nolfi, S., Elman, J.L., and Parisi, D. (1994). Learning and evolution in neural networks, Adaptive Behavior, 3.

Osborn, H.F. (1896). Ontogenic and phylogenic variation. Science, 4.

Piatelli-Palmarini (1989) "Evolution, Selection and Cognition" Cognition 31.

Pinker, S. & P. Bloom (1990) "Natural Language and Natural Selection" Behavioural and Brain Science 13.

Turkel, W. J. (To appear) "The learning guided evolution of natural language" in T Briscoe (ed.). Linguistic Evolution through Language Acquisition: Formal and Computational Models. New York: Cambridge.

Turney, P. (1997) "Evolution, Learning, and Instinct: 100 Years of the Baldwin Effect" Evolutionary Computation, Volume 4, Number 3.

Turney, P., Whitley, D., and Anderson, R.W. (1996). Evolution, learning, and instinct: 100 years of the Baldwin effect, Evolutionary Computation, 4.

Waddington, C. H. (1942) "Canalization of Development and the Inheritance of Acquired Characters" Nature 150.

Waddington, C. H. (1975) "The evolution of an evolutionist" Edinburgh University Press.

Culture *vs.* propositional thought as 'missing link' in the evolution of language

Tadeusz Wieslaw ZAWIDZKI

Philosophy-Neuroscience-Psychology Program
Washington University, St. Louis, MO, USA
tad@twinearth.wustl.edu

Overview

This paper examines the relative merits of two competing paradigms for explaining the evolution of natural language. According to the standard account, competence in natural language evolved when a grammar, capable of mapping propositional thought structures onto a serial medium of communication, was selected for in human pre-history.(Pinker & Bloom 1990) According to the 'neo-Vygotskyan' alternative, competence in natural language co-evolved with a capacity for propositional thought, as the result of “cumulative cultural evolution” (Boyd and Richerson 1996) in populations capable of a non-propositional form of cultural learning.(see especially Tomasello et. al. 1993; Tomasello & Call 1997) The discussion proceeds in two parts. In the first part, I summarize the standard account, and motivate the neo-Vygotskyan alternative by discussing two problems with it. In the second part, I attempt to make the neo-Vygotskyan alternative palatable by defusing some obvious objections to it.

The Standard account

According to Pinker and Bloom (1990), natural language is a clear example of complex, adaptive design, and, as such, its evolution is best explained by the mechanism of natural selection. They argue that the complex design apparent in the grammar of natural language is ideally suited for the function of mapping propositional thought structures onto a serial medium, for the purposes of communication: grammars for spoken languages must map propositional structures onto a serial channel, minimizing ambiguity in context, under the further constraints that the encoding and decoding be done rapidly, by creatures with limited short-term memories, according to a code that is shared by an entire community of potential communicants . . . language is a complex system of many parts, each tailored to mapping a characteristic kind of semantic or pragmatic function onto a characteristic kind of symbol sequence.(Pinker & Bloom 1990, 713)

Clearly, this explanation of the evolution of natural language implies that some non-language-using, evolutionary ancestor of humans had propositional thought structures to communicate, without the capacity to communicate them. If the function that natural language competence was selected for was the communication of propositional thought structures, then propositional thought must have pre-dated natural language.

There are two major problems with this picture. First, it is far from established that our non-language-using precursors wielded a language of thought. If the cognitive accomplishments of our nearest primate relatives can be accounted for without attributing a system of propositional thought to them, then this is *prima facie* reason not to attribute propositional thought to our common evolutionary ancestor. Currently, there exist numerous research programs based on the premise that we do not even need to attribute propositional thought to humans, in order to explain most human behavior, let alone other primates.(for a nice survey, see Clark 1997) Thus, it is far from clear that explanations of the evolution of natural language can help themselves to the assumption that our non-language-using precursors wielded a language of thought.

The second problem is apparent in an uncontroversial empirical fact. Our closest primate cousins do not use natural languages. But our most recent, non-language-using precursors were descendants of an ancestor that we share with our closest primate cousins. Thus, if our most recent, non-language-using precursors wielded a system of propositional thought, then it is likely that the ancestor that we share with our closest primate cousins did. If this is granted, then the standard account seems to imply the following: whereas, in the 5 to 7 million years since proto-chimpanzees and proto-hominids diverged from the common evolutionary ancestor (Pinker & Bloom 1990, 726), hominids have managed to develop a system to communicate their propositional thought, chimpanzees have not. Pinker & Bloom (Ibid., 724-725), as well as other defenders of the standard account, try to come up with ecological scenarios to explain this, often focusing on the unique challenges of a hunter-gatherer ecology. However, it is obvious to any student of chimpanzee ecology, that the adaptive benefits of a system for communicating propositional thought, if chimpanzees had it, would be considerable. Therefore, the question remains, if chimpanzees and their precursors have had propositional thought for just as long as humans and their precursors, why did a system for communicating propositional thought only evolve in humans?

There are two possible responses to this question. First, one might question the assumption that the evolution of natural language required the pre-existence of a system of propositional thought. However, if the function that competence in natural language was initially selected for consisted in the communication of thought, then surely we must follow Pinker & Bloom (1990) in defending this assumption. Natural language is not a good tool for communicating other kinds of cognitive state, like emotions, or representations encoding Euclidean relations.(Ibid., 715) Thus, if one questions the assumption that the evolution of natural language required the pre-existence of a system of propositional thought, one must also reject the assumption that natural language was initially selected to communicate thought. As I argue in the longer version of this paper, this option is most consistent with the neo-Vygotskyan alternative.

Second, one might argue that the brief period of evolutionary time, separating humans from the ancestor we share with chimpanzees, saw the evolution of propositional thought first, *followed by* the evolution of language. However, given the relatively brief period of time involved, it is exceedingly hard to construct ecological scenarios that can explain this progression. The only plausible scenario seems to call for some kind of self-amplifying, co-evolutionary dynamic between language and propositional thought, in which more language-like forms of communication led to more proposition-like forms of thought, and *vice versa*. However, as I argue in the longer version of this paper, this is also more compatible with the neo-Vygotskyan alternative, than with the standard account.

The neo-Vygotskyan alternative

According to the neo-Vygotskyan alternative, the evolution of language should be understood in the context of cultural evolution. For example, Tomasello, et. al. (1993), and Tomasello & Call (1997), suggest that the key cognitive divergence between the evolutionary precursors of hominids and the ancestor we share with chimpanzees, consisted in a capacity to create, transmit, and elaborate cultural practices. This led to a kind of 'cultural evolution,' among the products of which was natural language. By learning to use this product, humans learn to wield a propositional cognitive system.

I want to address two obvious problems with this alternative. First, on some understandings of what 'culture' and 'cultural learning' consist in, this is not really an alternative to the standard account. Many theorists would argue that the capacity to create, transmit, and elaborate cultural practices requires the kind of propositional thought that Pinker and Bloom claim natural language was selected to communicate. Therefore, in order to constitute a credible alternative to the standard account, the neo-Vygotskyan account must provide a definition of 'culture' and 'cultural learning' that does not presuppose the kind of propositional thought presupposed by the standard account.

Second, even if an acceptable definition of 'culture' and 'cultural learning' that did not presuppose propositional thought were provided, the connection between such phenomena and natural language would still appear mysterious. The great advantage of the standard account is that it explains why natural language has precisely the structure that it does: its main function is to communicate thought with analogous structure. Why would culture and cultural learning based on non-propositional forms of thought lead to the evolution of a system of communication with the structure of natural language?

In the long version of this paper, I suggest some strategies for defusing these two serious problems. With regard to the first problem, I turn to the animal culture literature, in search of definitions of 'culture' and 'cultural learning' that do not presuppose a capacity for propositional thought. Following Tomasello, et. al. (1993), I suggest that any population that displays the “ratchet-effect” (495), or what Boyd & Richerson (1996) call “cumulative cultural evolution” (79), should count as displaying culture and cultural learning. Such phenomena do not require propositional thought. They merely require mechanisms of “social canalization” (Boesch 1996, 257), like fairly rudimentary capacities to imitate models. There is ethological evidence that certain chimpanzee populations display such phenomena.(Ibid., 255-265) Furthermore, there is neurobiological evidence that many primates have neural mechanisms capable of implementing imitative learning.(Arbib &Rizzolatti 1996)

Given this understanding of 'culture' and 'cultural learning', the neo-Vygotskyan alternative seems threatened by the same sorts of objections as the standard account. If some chimpanzee populations display evidence of culture and cultural learning, and if many primates have neural mechanisms capable of implementing cognitive capacities necessary for culture and cultural learning, then why is there no evidence for the cultural evolution of natural language in non-human, primate species? In response to this worry, I draw on Boyd & Richerson's (1996) formal, evolutionary argument for the claim that, while culture may be common, cumulative cultural evolution is inevitably rare.(82-88)

Finally, I conclude the paper by offering some speculative suggestions for defusing the second serious problem with the neo-Vygotskyan alternative: if language is the product of cumulative cultural evolution based on non-propositional forms of cultural cognition, then why does it have the structure that it has? I suggest that the phylogenetically earliest function of language-like systems of communication consisted in supplementing imitation, in the transmission of ecologically crucial, cultural practices from parents to offspring. Proto-language may have consisted in a gestural form with a mimetic function: by miming hierarchically organized sequences of gestures involved in tool use, parents could enhance the transmission of ecologically crucial, tool-using practices to offspring. Such communicative behaviors would inherit the combinatorial structure of the tool use that they mimicked, and would thereby constitute an early form of a combinatorial, communicative system, like natural language.

References

Arbib, Michael A. & Rizzolatti, Giacomo (1996). "Neural Expectations: A Possible Evolutionary Path from Manual Skills to Language," *Communication and Cognition 29*, 393-424.

Boesch, Christophe (1996). "The Emergence of Cultures among Wild Chimpanzees," *Proceedings of the British Academy*, *88*, 251-268.

Boyd, Robert & Richerson, Peter J. (1996). "Why Culture is Common but Cultural Evolution is Rare," *Proceedings of the British Academy*, *88*, 77-93.

Clark, Andy (1997). *Being There*. Cambridge, MA: A Bradford Book/MIT Press.

Pinker, Steven & Bloom, Paul (1990). "Natural Language and Natural Selection," *Behavioral and Brain Sciences*, *13*, 707-784.

Tomasello, Michael & Call, Josep (1997). *Primate Cognition*. New York: Oxford University Press.

Tomasello, Michael, et. al. (1993). "Cultural Learning," *Behavioral and Brain Sciences*, *16*, 495-552.

Social patterns guide evolving grammars

Willem H. ZUIDEMA[*] **& Paulien HOGEWEG**[**]

[*] *Sony CSL, 6, Rue Amyot, 75005, Paris, France*
jelle@csl.sony.fr
www-binf.bio.uu.nl/~jelle

[**] *Theoretical Biology, Utrecht University*
Padualaan 8, 3584 CH Utrecht, The Netherlands
p.hogeweg@bio.uu.nl

We study the evolution of syntax in a simple multi-agent model. The fitness of agents in the model is not a fixed function of the individual languages, but depends on their communicative success in the group and thus on the composition of the population. This fact significantly alters the evolutionary dynamics, and can both facilitate and hinder the development of syntactic language. The results challenge the traditional picture of the transition towards syntactical language.

Introduction

The transition from short, finite communication systems found in many animal species, to the open ended language system of humans, is considered to be one of the major transitions in evolution (Maynard-Smith & Szathmáry 1995). There is large agreement that the main qualitative difference is the syntax of human language: the syntactic nature allows for a systematic production and interpretation of an unbounded amount of different messages. Syntax therefore reconciles the need for a large expressiveness with the limitations in human learning and memory. This aspect is, in the traditional view, what makes syntax *selectively advantageous*, and caused the transition from an extensive non-syntactical "protolanguage" to a more efficient, syntactical language system (Pinker & Bloom 1990; Nowak & Krakauer 1999).

We study this transition in a computational model of an evolving population of communicating agents. The main advantages of computational and mathematical models such as (Hurford 1989; Steels 1997; Hashimoto & Ikegami 1996; Nowak & Krakauer 1999), are that they are *relatively precise* and *productive*, in the sense that they generate new concepts and hypotheses. The main contribution so far is that they have shown the plausibility of *cultural evolution* as a mechanism in the development of more complex languages (De Jong 1998; De Boer & Vogt 1999; Batali 1997; Kirby 2000).

Fewer studies exist that model genetic transmission of language capabilities. Following (Hashimoto & Ikegami 1996), the model reported in this paper studies the dynamics of genetic transmission of language. Our model takes an extreme position, as it ignores learning mechanisms and semantics, and models genetic adaptation of *particular* grammars. Language capabilities are described with "context free grammars", that make compositional and recursive structures very easy to obtain. However, unlike some other studies of genetic transmission (*e.g.* Batali 1994), no static fitness function is defined; the grammars of all individuals in a group determine the environment in which an agent must survive.

Under these simplified conditions, the interaction between evolutionary dynamics and group dynamics can be studied. We will show that even without learning and cultural transmission, "social patterns" can influence the evolutionary dynamics. We observe that the multi-agent set-up can yield powerful, recursive grammars, but it can also prevent a population from obtaining them. Interestingly, because of these group effects, rules in one agent's grammar can influence the persistence of rules in other grammars, even though the mechanism of cultural evolution is excluded. We will show, that the results in some sense challenge the traditional picture of the transition towards syntactical language.

Model description

The model consists of a small set of agents that play a language game. They communicate in a language of short sequences ($max_l = 6$) of 0's and 1's. Agents speak ("derive") and understand ("parse") these strings using a Chomskyan rewriting grammar, which they inherited – with some random mutations – from their parent. In each language game, all agents can speak once and try to understand each of the spoken strings. Agents receive scores depending on their success in speaking, understanding and (not) being understood. After a number of language games, scores are evaluated and offspring is produced. Successful agents have a higher chance of survival and reproduction.

The grammars of the agents are context free grammars, with a small terminal ($V_{te} = \{0, 1\}$) and non-terminal alphabet ($V_{nt} = \{S, A, B\}$). As an extra restriction, the start symbol is not allowed on the right-hand side of rules. At the start of most simulations, grammars are randomly initialized with either $S \rightarrow 1$ or $S \rightarrow 0$.

Derivation always starts with the start symbol, and applies iteratively random fitting rules for some maximum number of steps ($max_d = 60$; failure), until no fitting rule exists (failure), or until a string of only terminal symbols is reached (success). In *parsing* rules are tried in the order they are stored, and fitting rules are applied recursively until the maximum number of steps ($max_p = 500$) is reached (failure), no other fitting rules exist to any intermediate string (failure), or the start symbol is reached (success).

The model architecture is similar to the model introduced by (Hashimoto & Ikegami 1996). They discuss their results in terms of the Chomsky hierarchy of grammars and languages. In a domain of changing grammars and finite languages, we believe it is much more convenient to use a classification in terms of "routes". A *route* is a sequence of rewriting steps that connects the start symbol *S* to a string of terminal symbols. Routes can be categorized as *lexical* (directly from *S* to a terminal string), *compositional* (via non-terminal symbols from *S* to a terminal string) or *recursive* (leading from a non-terminal symbol via one or more rewriting steps to the same non-terminal symbol). The number of routes, can be divided in three components R_I, R_C, R_R, that depend on each of these categories of routes. Similarly, expressiveness (the number of distinct strings a grammar can parse) can be divided in E_I, E_C, E_R routes. Grammars can be characterized by these values, and classified according to the largest component (Zuidema 2000).

Results

To evaluate some general properties of the model, we studied the behavior with the parameter settings of (Hashimoto & Ikegami 1996), and a number of variations. Similar to their results, we find that evolution can quickly lead to grammars that can parse a large fraction of the 126 possible strings. However, under slightly different parameter settings we also find quite different results. We observe three types of behavior:

i. The most frequent behavior is a quick growth of expressiveness, from 1 at initialization, to over 100 after about a 1000 generations. In the first stage the expressiveness depends only on lexical routes. Soon, however, compositional routes and recursive routes become more important. Eventually, recursive routes dominate the grammar's expressiveness.

ii. Sometimes, it takes much longer to reach the high level of expressiveness, ranging from 2000 to many thousands of generations. In these type of runs, compositional routes quickly become important, but recursive routes are infrequent.

iii. Least frequent are runs that show very little growth in expressiveness. After 3000 generations, only around 20 words can be parsed. In these runs, expressiveness depends almost exclusively on lexical routes.

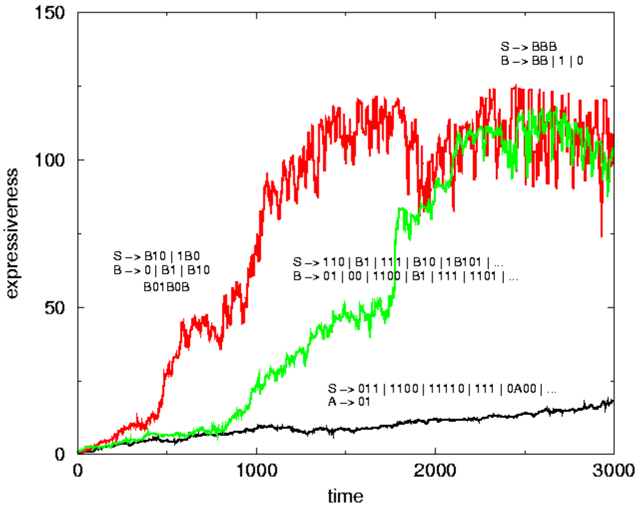


Figure 1: *Three runs, typical for the lexical, compositional and recursive regimes, and some example grammars.*

These types of behavior also differ in their robustness against mutations and generalization abilities. With some particular parameter settings, each of the three types of behavior can occur, solely depending on the "seed" for the random generator. At different generations we restarted runs with original grammars but a different random seed. In early generations, a change of type of behavior occurs frequently. However, in restarts from later generations, the type of behavior seems fixed and a change of type becomes increasingly improbable. The types of behavior thus form *self-enforcing, dynamical regimes*.

Context and group effects

These results crucially depend on the fact that the fitness of an agent is evaluated with respect to its performance in the group, rather than with respect to some static fitness function. In a random population, agents with more expressive grammars speak more novel strings, understand more strings and are less likely to be understood, and thus should receive higher scores. The existence of the dynamical regimes, in a non-random population, can be explained by three mechanisms: a *mutation bias* (mutation tend to produce similar grammars), a *context effect* (rules are generally most successful in a context of similar rules) and a group effect (agents are most successful in a group of similar agents).

The derive-languages of individuals, jointly constitute a group language, that in turn determines the success of agents in parsing. This indirect feedback can best be described as a *social pattern* that emerges from individual behaviors, and in turn restricts individual success. Initial similarities (in terms of our classification) are enforced by these social patterns.

Apparently, the larger an lexical grammar is, the less likely it is that evolution can lead to compositional and recursive grammars. This in some sense contradicts the traditional picture of the evolution of syntax, that states that only when lexical grammars became too large, syntax emerged.

A simple analysis can lead to some qualitative predictions on how, given the existence of these regimes, different variables in the model should relate. One can show, that the number of routes grows linearly with grammar size in the lexical regime. In a compositional regime it grows faster, and in a recursive regime extremely fast[1]. A rough estimate of how expressiveness[2] depends on *R*, gives a qualitative explanation for the trajectories in the phase space in figure 2. If a linear growth of grammar size over time is assumed, the shape of the curves in figure 1 can also be explained.

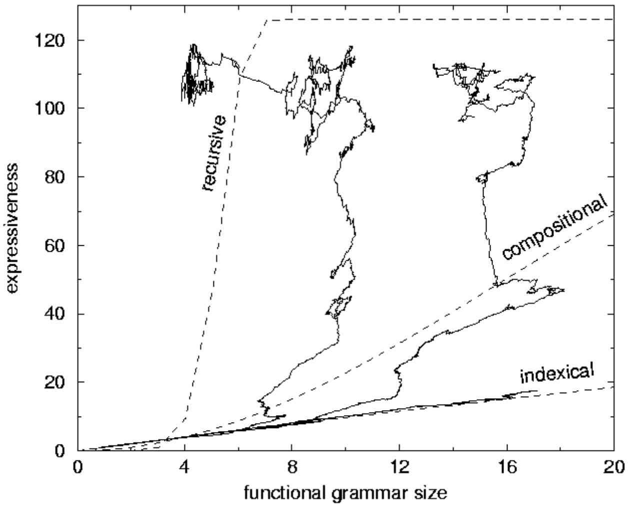


Figure 2: *Trajectories of the same three runs in a phase space of functional grammar size (the number of rules that are actually used in communication) vs. expressiveness*

Selective advantages

With the scope of possible behaviors sketched, we can now turn to the question under what selection pressures the different dynamical behaviors are likely to arise. To study this, we designed several simple scoring schemes. These include: (*i*) communication, where both speaker and receiver benefit from exchanging information; and (*ii*) perception, where only the recognizing agent benefits from the information it receives.

Although recursive structures are always only a few mutations away, the development of recursive and expressive grammars is not trivial at all. With the default initial grammar (randomly $S \rightarrow 1$ or $S \rightarrow 0$), the communication scheme shows no increase in expressiveness, unless an explicit pressure is put on innovation. In that case sometimes recursive structures develop, but slowly and to a limited extent. If grammars are initialized with a longer lexical grammar, even this explicit innovation pressure can not force the simulation out of the lexical regime.

---

[1] Take for example the simple case of grammars with $V_{nt} = \{S,A\}$, and at most one non-terminal and at least one terminal symbol at all right-hand sides of rules. Estimates of *R* in each of the regimes are: $R_I \approx N$, $R_C \approx (½\,N)^2$, $R_R \approx (N/3)^{max_c+2}$, where max_c is the maximum number of cycles.

[2] $E \approx Emax\ \left(1 - (1 - {}^{1}/_{E_{max}})^{R}\right)$, here Emax = 126.

In contrast, the perception scheme leads to recursive grammars under all circumstances considered. However, when the population has been at a high level of expressiveness for some time, agents start to develop grammars that are just as expressive, but have a high probability of failing in derivation. The asymmetry in parsing and derivation makes this possible.

These results yield an interesting paradox. Under the parameter settings that lead to expressive grammars, the willingness to speak is absent, while in cases where communication is mutually beneficial, no increase in expressiveness occurs.

Conclusions

This study concerns the interaction between group dynamics and evolutionary dynamics. We have seen that social patterns influence the course of evolution. Under some conditions powerful, recursive grammars develop (Hashimoto & Ikegami 1996). This appears to be due to the social embedding that yields a dynamical fitness landscape, because simulations with a fixed fitness evaluation fail to give similar good results.

However, in other circumstances social patterns hinder the development of such grammars. These results are particularly interesting, as these specific circumstances in some sense resemble the situation that is thought to precede the emergence of syntax: large lexical grammars and mutually beneficial communication. In the model we arrive at a paradox, where those selection pressures that lead to syntactical languages, also lead to unwillingness to speak. Preliminary results indicate, that this paradox can be solved if a spatial distribution of agents and local communication is assumed\citep (Zuidema & Hogeweg 2000).

Relaxing the idea of explicit selection pressures for syntax, the analysis of the shapes of the curves in figure 2 points at an alternative mechanism for the development of recursion. The fact that recursive expressiveness (E_R) grows very fast with the number of rules (N), shows that the larger N (i.e. the "storage capacity"), the larger the expected *relative* fraction of recursive expressiveness. Whereas the traditional view emphasizes that *cognitive limitations* create the need for syntax, this observation indicates that larger *cognitive abilities* in fact make recursive expressiveness more likely to dominate. This might explain the apparent paradox, that the species with the most extended cognitive abilities, is the only species that developed "efficient", recursive communication.

References

Batali, J. (1994). "Innate biases and critical periods: Combining evolution and learning in the acquisition of syntax". In *Artifical Life IV*, Brooks & Maes (eds.), 160-171, MIT Press.

Batali, J. (1997). "Computational simulations of the emergence of grammar". In: *Approaches to the evolution of language: social and cognitive bases*, Hurford & Studdert-Kennedy (eds.). Cambridge Univ. Press.

De Boer, B. & Vogt, P. (1999). "Emergence of speech sounds in changing populations". In: ECAL'99 *Advances in artificial life.* Floreano, Nicoud & Mondada (eds.). 664-673, Springer

De Jong, E. (1998). "The development of a lexicon based on behavior". In: *Proceedings of BNAIC'98.* La Poutré & Van den Herik (eds.). 27-36, CWI.

Elman, J. L. (1991). "Distributed representations, simple recurrent networks, and grammatical structure. *Machine Learning 7*, 195-225.

Hashimoto, T. & Ikegami, T. (1996). "The emergence of a net-grammar in communicating agents". *BioSystems 38*, 1-14.

Hurford, J. (1989). “Biological evolution of the saussurean sign as a component of the language acquisition device. *Lingua 77*, 187-222.

Kirby, S. (2000). “Syntax without natural selection”. In: *The evolutionary emergence of language: social function and the origin of linguistic form (to appear)*. Knight, Hurford & Studdert-Kennedy (eds.). Cambridge Univ. Press.

Maynard-Smith, J. & Szathmary, E. (1995). *The major transitions in evolution*. Morgan-Freeman.

Nowak, M. A. & Krakauer, D. C. (1999). The evolution of language. *Proc. Nat. Acad. Sci. USA\/ 96*, 8028-8033.

Pagie, L. & Hogeweg, P. (1997). “Evolutionary consequences of coevolving targets”. *Evol. computation 5*, 401-418.

Pinker, S. & Bloom, P. (1990). “Natural language and natural selection”. *Behavioral and brain sciences*.

Steels, L. (1997). “Synthesising the origins of language and meaning”. In: Approaches to the evolution of language. Hurford & Studdert-Kennedy (eds.). Cambridge Univ. Press.

Zuidema, W. H. (2000). *Evolution of syntax in groups of agents*. *Master's thesis*, Utrecht University.

Zuidema, W. H. & Hogeweg, P. (2000). Selective advantages of syntax. *Submitted to Cognitive Science 2000.*

Authors