

Michael A. Arbib and  
Derek Bickerton (eds.)

# The Emergence of Protolanguage

Holophrasis vs compositionality

BENJAMINS CURRENT TOPICS

24

# The Emergence of Protolanguage

Holophrasis vs compositionality

*Edited by*

Michael A. Arbib

University of Southern California

Derek Bickerton



John Benjamins Publishing Company

Amsterdam / Philadelphia



The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

**Library of Congress Cataloging-in-Publication Data**

The emergence of protolanguage : holophrasis vs compositionality / edited by Michael A. Arbib and Derek Bickerton.

p. cm. (Benjamins Current Topics, ISSN 1874-0081 ; v. 24)

Includes bibliographical references and index.

1. Language and languages--Etymology. 2. Language acquisition. 3. Human evolution. 4. Historical linguistics. I. Arbib, Michael A. II. Bickerton, Derek.

P321.E46 2010

417'.7--dc22

2010025972

ISBN 978 90 272 2254 1 (Hb ; alk. paper)

ISBN 978 90 272 8782 3 (Eb)

© 2010 – John Benjamins B.V.

No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Co. · P.O. Box 36224 · 1020 ME Amsterdam · The Netherlands  
John Benjamins North America · P.O. Box 27519 · Philadelphia PA 19118-0519 · USA

## The Emergence of Protolanguage

## *Benjamins Current Topics*

Special issues of established journals tend to circulate within the orbit of the subscribers of those journals. For the Benjamins Current Topics series a number of special issues have been selected containing salient topics of research with the aim to widen the readership and to give this interesting material an additional lease of life in book format.

### **Volume 24**

The Emergence of Protolanguage. Holophrasis vs compositionality

Edited by Michael A. Arbib and Derek Bickerton

These materials were previously published in *Interaction Studies* 9:1 (2008), under the general editorship of James R. Hurford.

## Preface

Somewhere and somehow, in the 5 to 7 million years since the last common ancestors of humans and the great apes, our ancestors “got” language. The authors of this volume all agree that there was no single mutation or cultural innovation that took our ancestors directly from a limited system of a few vocalizations (primarily innate) and gestures (some learned) to language. They further agree to use the term “protolanguage” for the beginnings of an open system of symbolic communication that provided the bridge to the use of fully expressive languages, rich in both lexicon and grammar. But here consensus ends, and the theories presented here range from the *compositional view* that protolanguage was based primarily on words akin to the nouns and verbs, etc., we know today with only syntax lacking to the *holophrastic view* that protolanguage used protowords which had no meaningful subunits which might nonetheless refer to complex but significantly recurrent events.

The present volume does not decide the matter but it does advance our understanding. The lack of any direct archaeological record of protolanguage might seem to raise insuperable difficulties. However, this volume exhibits the diversity of methodologies that can be brought to bear in developing datasets that can be used to advance the debate. Here is a quick tour:

Kenny Smith: *Is a holistic protolanguage a plausible precursor to language? A test case for a modern evolutionary linguistics.* As Smith observes, if protolanguage were a holistic system with complex meanings conveyed using unanalyzed forms, there must be some process (*analysis*) which developed the elements of modern language from this system. He draws on evidence from computational modelling, developmental and historical linguistics and comparative psychology to evaluate the plausibility of the analysis process. He concludes that some of the claims for the holophrastic view can be refuted using such evidence, but highlights other key issues where further evidence is required.

Jill Bowie: *Proto-discourse and the emergence of compositionality.* Bowie offers a discourse-oriented perspective on the protolanguage debate – focusing not on isolated utterances but on their use in larger contexts. She argues that discourse should not be viewed as a level of language structure ‘beyond the sentence’ so

much as sequenced communicative behaviour, typically but not uniquely involving language. This provides for continuity from exchanges making use of simple communicative resources such as protolanguage might have offered (and that single words and gestures still offer today) to those making use of complex grammatical conventions. She then marshals evidence from child language and her own experiments with adults using constrained language systems to conclude that the utility of emerging compositional language is not dependent on some critical level of complexity.

Patricia M. Greenfield, Heidi Lyn and E. Sue Savage-Rumbaugh: *Protolanguage in ontogeny and phylogeny. Combining deixis and representation*. Greenfield et al. compare data from chimpanzees and bonobos learning a human-designed visual communication system (lexigrams) with that from children acquiring a first language. They find that the potential to combine two different kinds of semiotic element – deictic (a pointing gesture) and representational (a word) – is fundamental to the “protolanguage” exhibited by the apes and young children, and then argue that this form of compositionality also provided the foundation for the earliest human languages. They use this argument from ontogeny to conclude that human protolanguage was exclusively holophrastic only if one considers the symbol alone without taking into account the gesture as a second element comprising the total message.

Jean-Louis Dessalles: *From metonymy to syntax in the communication of events*. Dessalles sees the key to protolanguage as being *metonymy*, where a signifier (sound, gesture, word) is used to refer to a whole event even though it is normally associated with only one aspect of it. He employs a modular analysis of spontaneous language to support the view that protolanguage first evolved as a grammarless form of expression to signal events which were unexpected yet recurring, extending the capability seen in monkey alarm calls, for example. At this stage, protowords would not be so specific as to refer to the particularities of whole, non-recurring, situations. They referred to elements such as objects or locations, and the communicated event was inferred metonymically. Dessalles argues that compositionality was first achieved without syntax, through multi-metonymy, as words referring to aspects of the same situation were concatenated into proto-utterances.

Markus Werning: *The “complex first” paradox. Why do semantically thick concepts so early lexicalize as nouns?* For a cortically implemented syntax-semantics interface, the more widely distributed a concept’s neural realization is, the more effort it takes – so it would seem – to establish a link between the concept and its expression. Werning grounds his analysis in the finding that, in modern humans, the concepts expressed by concrete nouns are more complex (“thicker”) and

their neural realizations more widely distributed in cortex than those expressed by other word classes. If one assumes the principle that in ontogeny capabilities demanding more effort develop later than those demanding less effort, it seems to be a paradox, the “complex first” paradox, that the meanings of concrete nouns are acquired by children earlier than those of other word classes. Werning then looks at the implications of positing that the paradox also applied in phylogeny and uses this to raise questions for research in computational neuroscience. How could a mechanism evolve that enables certain regions of cortex that are involved in representing a word (phonologically, syntactically, etc.) to address those regions of the sensori-motor cortices that represent the word’s meaning? And could it be that distributive neural states are actually more easily addressable than local ones?

Maggie Tallerman: *Holophrastic protolanguage. Planning, processing, storage, and retrieval*. Tallerman challenges recent assumptions that holophrastic utterances could be planned, processed, stored and retrieved from storage, rooting her arguments in modern data on psycholinguistics. She notes the complexity of conceptual planning of multi-proposition utterances. She questions whether the “holistic” mode posited by holophrastic protolanguage could serve as the precursor of an “idiom mode” in modern language. On her analysis, the production and comprehension of idioms in modern languages does not involve a putative “holistic” mode of language processing. Finally, Tallerman claims that innate constraints on learning lexical items preclude the types of protoword meanings proposed by proponents of holophrastic protolanguage.

Andrew D. M. Smith: *Protolanguage reconstructed*. By contrast, Smith supports a holophrastic view by arguing that protolanguage utterances had varying degrees of semantic complexity, and developed into complex language gradually, through the processes of re-analysis and analogy which still underpin continual change in modern languages. He presents evidence about the nature of linguistic communication to assess the plausibility of different assumptions concerning the semantic complexity of protolinguistic utterances. He shows that communication is fundamentally inferential and characterised by semantic uncertainty. This not only allows individuals to maintain variation in linguistic representation, but also imposes a selection pressure that meanings be reconstructable from context.

David McNeill, Susan D. Duncan, Jonathan Cole, Shaun Gallagher and Bennett Bertenthal: *Growth points from the very beginning*. This is the first of three papers that assess the possible role of gesture in the emergence of language. The primary methodology is to look at how modern humans use gesture when they are speaking (these co-speech gestures are to be distinguished from the conventionalized forms of a modern sign language) or the use of gesture alone, with further support



from data for a speaker whose sensory fibre neuropathy prevents instrumental actions from occurring normally, yet who still gestures with speech. McNeill et al. hypothesize that early humans formed language units which they call "growth points" consisting of global and discrete dimensions of semiosis in dynamic opposition. Eventually, gestures gained the power to orchestrate actions, manual and vocal, with significances other than those of the actions themselves, giving rise to cognition framed in dual terms. However, their proposal emphasizes natural selection of joint gesture-speech, not "gesture-first," in language origin.

Mark Aronoff, Irit Meir, Carol A. Padden and Wendy Sandler: *The roots of linguistic organization in a new language*. Aronoff et al. approach the emergence of *protolanguage* through comparison with their own data on the emergence of a new language with no direct linguistic history. Al-Sayyid Bedouin Sign Language (ABSL) arose about 70 years ago in a small, isolated community with a high incidence of profound prelingual neurosensory deafness. In ABSL, they were able to identify the beginnings of phonology, morphology, syntax, and prosody. The linguistic elements they find in ABSL are not exclusively holistic, nor are they all compositional, but a combination of both. ABSL has a highly regular syntax as well as word-internal compounding, also highly regular but quite distinct from syntax in its patterns. However, ABSL lacks certain features that have been posited as essential even for a proto-language. It lacks both the spatially organized morphology and the evident duality of phonological patterning observed in more mature sign languages.

Michael A. Arbib: *Holophrasis and the protolanguage spectrum*. Much of the debate concerning the question "Was Protolanguage Holophrastic?" assumes that protolanguage existed as a single, stable transitional form between communication systems akin to those of modern primates and human languages as we know them today. Arbib argues for a spectrum of protolanguages preceding modern languages emphasizing that (i) protospeech was intertwined with protosign (a conventionalized gestural system abstracted from pantomime) and gesture; (ii) grammar emerged from a growing population of constructions; and (iii) an increasing protollexicon drove the emergence of phonological structure. This framework weakens arguments for the view that the earliest protolanguages were not holophrastic while advancing the claim that protolanguages became increasingly compositional over time en route to the emergence of true languages.

Derek Bickerton *But how did protolanguage actually start?* To conclude the volume, Bickerton asks us to consider an issue that would surely have influenced the nature of protolanguage but has too often been neglected: the precise circumstances under which protolanguage arose. Three factors are involved in this neglect: a

failure to appreciate radical differences between the functions of language and animal communication, a failure to relate protolanguage development to the overall course of human evolution, and the supposition that protolanguage represents a package, rather than a series of separate developments that sequentially impacted the communication of pre-humans. Bickerton emphasizes the need for a paleoanthropological approach that takes into account the ecological niches occupied by pre-human species using the methodology of niche construction theory. While agreeing that much has been learned from studies of apes in the wild and attempts to teach them language, he stresses that the environments, ecologies, and means of subsistence of species in the immediate pre-human line differed enormously from those that characterized the chimp/bonobo line.

# Table of contents

Preface	vii
Is a holistic protolanguage a plausible precursor to language? A test case for a modern evolutionary linguistics <i>Kenny Smith</i>	1
Proto-discourse and the emergence of compositionality <i>Jill Bowie</i>	19
Protolanguage in ontogeny and phylogeny: Combining deixis and representation <i>Patricia M. Greenfield, Heidi Lyn and E. Sue Savage-Rumbaugh</i>	35
From metonymy to syntax in the communication of events <i>Jean-Louis Dessalles</i>	51
The “complex first” paradox: Why do semantically thick concepts so early lexicalize as nouns? <i>Markus Werning</i>	67
Holophrastic protolanguage: Planning, processing, storage, and retrieval <i>Maggie Tallerman</i>	83
Protolanguage reconstructed <i>Andrew D. M. Smith</i>	99
Growth points from the very beginning <i>David McNeill, Susan D. Duncan, Jonathan Cole, Shaun Gallagher and Bennett Bertenthal</i>	117
The roots of linguistic organization in a new language <i>Mark Aronoff, Irit Meir, Carol A. Padden and Wendy Sandler</i>	133
Holophrasis and the protolanguage spectrum <i>Michael A. Arbib</i>	153
But how did protolanguage actually <i>start</i> ? <i>Derek Bickerton</i>	167
Name index	175
Subject index	179

# Is a holistic protolanguage a plausible precursor to language?

## A test case for a modern evolutionary linguistics

Kenny Smith

Northumbria University

If protolanguage was a holistic system where complex meanings were conveyed using unanalysed forms, there must be some process (*analysis*) which delivered up the elements of modern language from this system. This paper draws on evidence from computational modelling, developmental and historical linguistics and comparative psychology to evaluate the plausibility of the analysis process. While some of the criticisms levelled at analysis can be refuted using such evidence, several areas are highlighted where further evidence is required to decide key issues. More generally, the debate over the nature of protolanguage offers a framework for developing and showcasing a modern, evidence-based evolutionary linguistics.

### 1. Introduction

Humans have language. It is hypothesised that the common ancestor of chimpanzees and humans did not. Evolutionary linguists therefore have to explain how the gap between a non-linguistic ancestor and our linguistic species was bridged. It has become common to invoke the concept of a *protolanguage* as a stable intermediary stage in the evolution of language: “[t]he hypothesis of a protolanguage helps to bridge the otherwise threatening evolutionary gap between a wholly alingual state and the full possession of language as we know it” (Bickerton, 1995, pp. 51).<sup>1</sup>

What was protolanguage like? Under the *holistic* account, (see, e.g., Wray, 1998), protolanguage was a system in which individual signals, lacking in internal morphological structure, conveyed entire complex propositions. The transition from a holistic protolanguage to language occurred when holistic utterances were broken down to yield words and constraints on their combination. This process is known as *analysis*, also sometimes referred to as segmentation (Wray, 1998) or

fractionation (Arbib, 2005).<sup>2</sup> In order for holistic protolanguage to be a plausible precursor to modern language, it must be possible to get from such a protolanguage to language: the analysis process must be shown to be plausible.

In the context of a broader assault on holistic protolanguage, Tallerman (2007) provides a thought-provoking critique of the analysis process. According to Tallerman, analysis suffers from the following defects:<sup>3</sup>

1. Analysis requires cognitive resources greater than we can expect early hominids to exhibit.
2. Analysis would be blocked by counter-examples to any nascent generalisations.
3. Analysis forces us to posit fundamental discontinuities between prehistoric and contemporary processes of language change.

In Sections 3–5 I consider the evidence available to evaluate each of these criticisms. This process provides a useful framework in which to explore the kinds of evidence we can use to evaluate theories of protolanguage in particular, and the evolution of language more generally. It is often claimed that evolutionary linguistics suffers from a paucity of evidence:

“To enter [the field] costs little: you can’t do experiments, so no expensive equipment is required (...). It’s still a pencil-and-paper field” (Bickerton, 2007, pp. 524)

I will argue that, on the contrary, there is a wealth of empirical evidence which evolutionary linguists can draw on to constrain and inform theory: many relevant experiments have already been done and, importantly, any serious attempt to evaluate any theory of the evolution of language is likely to suggest further experiments which remain to be done. In this paper I will use evidence from comparative psychology, developmental and historical linguistics and computational modelling to evaluate the plausibility of a transition from a holistic protolanguage to language via analysis, and identify several key areas where further evidence is needed to discriminate between competing claims. This provides an illustration of the more general process of evaluating theories of the origins and evolution of language. A modern evolutionary linguistics should draw on existing data from all these areas, but more importantly, use methodologies from these fields to actually go out and test hypotheses from the literature.

## 2. Learning by segmentation and the analysis process

Analysis is the process by which holistic utterances are broken down over historical time into component words plus rules which govern their combination. Wray (1998) describes a scenario under which chance cooccurrences of meaning and form between holistic utterances lead protolanguage learner/users to segment out words, leaving behind a residual template. Wray (1998, pp. 55–56) illustrates this process with a hypothetical example of segmentation, based around the following two signs of a holistic protolanguage — as in Wray (1998), signals are given as sequences of phonemes, semantics are given in English.

- (1) /mebita/  $\longleftrightarrow$  “give her the food”
- (2) /kameti/  $\longleftrightarrow$  “give her the stone”

Wray suggests a scenario where a segmenting learner notes and exploits the partial regularity in (1) and (2), namely that a common element of signal (/me/) corresponds to a common element of meaning (“her”).<sup>4</sup> This coincidence of meaning and form occurs by chance in the holistic protolanguage providing signs (1–2). The segmenting learner notes this regularity, and segments out a morpheme which captures it, leaving behind a residual unanalysed template. Schematically, the internal representation of the partiallysegmented protolanguage would be:

- X/bita/  $\longleftrightarrow$  “give X the food”
- /ka/X/ti/  $\longleftrightarrow$  “give X the stone”
- /me/  $\longleftrightarrow$  “her” (in contexts where it substitutes into position X)

Subsequently, the individual who has discovered this structure may produce novel utterances which exploit this regularity, systematically using /me/ to convey the meaning “her”. The accumulation of these segmentations, and their exploitation by segmenting learners, leads to the historical process of analysis, whereby an initially unstructured holistic system comes to exhibit structure based on words and constraints on their combination.

Have we any concrete reason to believe that a holistic protolanguage will evolve into a system with words and rules in a population of individuals learning in this way? Human intuitions on these kinds of complex historical processes tend to be poor. In cases such as these, computational and mathematical models provide a valuable tool for conducting “opaque thought experiments” (Di Paolo, Noble, & Bullock, 2000), or mechanically working through the macroscopic consequences of a well-specified set of microscopic assumptions (e.g. assumptions about the processes of learning).

Kirby (2002) provides a model which demonstrates that cultural transmission in a population of segmenting learners can result, under certain plausible

transmission conditions, in a transition from holism to a compositionally-structured linguistic system.<sup>5</sup> A number of such models demonstrating this process exist (see Kirby, Smith, & Brighton, 2004, for review): similar results have been shown for different models of language learning (e.g. a heuristic grammar inducer in Kirby, 2002; an associative network model in Smith, Brighton, & Kirby, 2003), different treatments of population (purely vertical transmission in e.g. Kirby, 2002; purely horizontal transmission in Batali, 2002), and different treatments of the grounding of language in use (no grounding of meaning in e.g. Kirby, 2002; grounding in Vogt, 2005).

The wealth of formal modelling in this area serves two functions. Firstly, each model provides proof of concept for the analysis process. Secondly, the diversity of modelling approaches suggests that the analysis process is at least somewhat robust, having been demonstrated under a wide range of assumptions about how learning works, how populations are structured, and how meaning is constructed. This breadth of approaches is significant — while we can debate the relevance of the assumptions made in one model, repeated demonstrations of the same phenomenon in a range of models provides converging evidence that the process we are dealing with is not completely dependent on certain key assumptions. Without this diversity of coverage, we need to either be more cautious in extrapolating from modelling results, or have greater confidence in the key assumptions made in our models.

### 3. Criticism 1: Can *Homo* analyse?

Computational models show that analysis can in principle deliver up words and rules from a holistic protolanguage. But how cognitively demanding is the type of learning underpinning analysis? Can modern humans do it? If so, we might accept that analysis could have begun with the advent of *Homo sapiens*. Could earlier hominids (e.g. *Homo erectus*, tied to the inception of the analysis process by Tallerman, 2007) do it?

#### 3.1 Can modern humans analyse?

There is strong developmental and historical evidence that modern humans do segment and analyse. Tallerman herself points towards a contemporary example of segmentation in action: language acquisition. Children successfully segment out words and constraints on their combination from instances of language use which must be treated, at least initially, as unanalysed meaning-form pairs

(see, e.g., Tomasello, 2003). Indeed, Wray's initial account of segmentation was explicitly motivated by Peters's (1983) account of language learning. The historical literature also suggests that structure can be introduced into words where none was previously present, through processes of *back-formation* and *reanalysis* (e.g. back-formation of the verb "peddle" from the noun "peddler" due to the coincidence between the "er" ending of the noun and the derivational affix "-er", productive elsewhere in the language: Simpson & Weiner, 1989).

Tallerman, following Johansson (2005), raises the important caveat that children segment a system of form-meaning pairs which contains abundant evidence of structure ripe for segmentation. Similarly, historical reanalysis is dependent on structure already present elsewhere in the language (e.g. a productive affix). In contrast, the analysis process requires segmentation in situations where apparent structure is sporadic and not generated by any underlying rule. While it is therefore safe to argue that humans are capable of segmentation and analysis under modern conditions, whether a modern human learner would also segment and analyse a holistic protolanguage is less clear.

There is in fact some evidence to support Tallerman's contention that segmentation during acquisition (and the related historical processes) is dependent on the presence of a large body of confirming evidence for this segmentation. Type frequency (the number of verbs participating in a particular inflectional paradigm) is a key determinant of the productivity of inflectional paradigms (Bybee, 1985, 1995). This relationship between type frequency and productivity is readily explicable under a segmenting model of learning: high type frequency provides precisely the circumstances required under the segmenting model for generating structural abstractions, specifically a varied range of fillers capable of slotting in to a particular template. However, the fact that productive abstractions seem to require high type frequency suggests that infrequently-occurring chance correspondences (such as we might expect to find in a holistic protolanguage) would not lead to segmentations which can be productively applied. Child-directed speech also appears to be well-adapted to a segmenting learner who requires abundant evidence for segmentations (Cameron-Faulkner, Lieven, & Tomasello, 2003), being dominated by a small number of templates (e.g. "What's X doing?"; "That's a X", "Are you going to X?") with a wide range of items slotting in to those templates.

While this is rather suggestive, it is unclear what the boundary conditions on segmentation are: how little evidence does a modern human require to make a segmentation? If that evidence must be abundant, as Tallerman suggests, we should be sceptical as to the likelihood that analysis could get off the ground based on a small number of chance correspondences. While the evidence from morphology and child-directed speech is suggestive, a more direct means of addressing



this crucial question is desirable. The most straightforward way of resolving this issue would be to conduct a relatively simple psycholinguistic experiment similar to those described in Gómez (2002). Gómez demonstrates experimentally that templates which apply to highly variable fillers are more readily learned by adults and children than a template which appears in more stereotyped circumstances, applying only to a small number of fillers. A similar methodology could be applied to explore whether fillers for a given template have to be highly variable for the filler-template representation to actually be internalised.

### 3.2 Could earlier hominids analyse?

Although the all-important boundary conditions for segmentation remain mysterious, there is pretty good historical and developmental evidence that modern humans can do segmentation and analysis in at least some conditions. Would earlier *Homo* have had similar capacities to modern humans?

Tallerman is deeply sceptical:

“words will never appear out of formulae unless the hominids using holistic protolanguage have both the necessary motor control and the neural capacity to recognise phonetic strings ... how could these abilities exist prior to the language faculty itself?” (Tallerman, 2007, pp. 595)

How can we know what early *Homo* was capable of in terms of segmentation and analysis? While we might note that the ability to spot co-occurrences of meaning and form across two signs can be realised by fairly rudimentary learning devices (e.g. an associative network, Smith et al., 2003) or learners with fairly limited capacities of attention and memory (e.g. children), the more general question remains of how to evaluate claims about the cognitive capacities of extinct species.

Saffran, Aslin, and Newport (1996) investigate the process whereby language learners break up a continuous stream of sounds into words. This mechanism could be used by a segmenting learner to identify strings of syllables (or phonemes) which tend to co-occur across utterances, such sequences being candidates for segmentation. They found that 8 month old infants were able to use simple statistical properties of the input (syllable transitions within words are relatively predictable relative to syllable transition between words) to segment out words. Crucially, the same capacities have been shown by Hauser, Newport, and Aslin (2001) to be present in cotton-top tamarins (*Saguinus oedipus*, last common ancestor with humans around 40 million years ago: Rosenberger, 1992). Although this doesn't show that the capacity to analyse is within the capabilities of non-human primates, it is at least an indication that some of the capacities (e.g. “the neural capacity to recognise phonetic strings”), are found in non-linguistic species. It also highlights