

SERENA NICCHIARELLI  
Roma Tre University, Italy

## Formulaic Language: A Living Linguistic Fossil for a Holistic Protolanguage

### Abstract

Humans today have the ability to use language. The common ancestor of chimpanzees and humans probably did not. During recent decades evolutionary linguists have attempted to explain how the gap between a non-linguistic ancestor and our linguistic species was bridged. In this direction, it has become common to invoke the notion of a *protolanguage* as a stable intermediary stage in the evolution of language. A key dispute among the currently-available hypotheses of protolanguage is represented by the distinction between holistic and synthetic accounts: did human protolanguage consist of holistic utterances – later segmented into single words – or did it start with simple units that were added together into more complex structures? The synthetic account is generally recognized as “the standard model,” thus assuming that the earliest forms of a presumed protolanguage were compositional, that is built up from single words, where one word corresponds to one concept. However, recent years have seen the consolidation of the alternative idea: each element of a protolanguage would have been linguistically unanalyzable and referred to a whole situation. This paper presents the case of *formulaic language* as evidence – a living linguistic fossil – which corroborates arguments in support of a holistic protolanguage account.

*Keywords:* language origins, holistic protolanguage, formulaicity.

### Holistic vs. Synthetic: an ongoing debate

With Arbib (2005) let us define a *protolanguage* as a system of utterances which we may recognize as a precursor to human language, but which is not itself a human language in the modern sense.

Until relatively recently, the dominant theory of the evolution of complex language from protolanguage being the “synthetic” approach, which is most clearly articulated in the work of Derek

Bickerton (1990, 1996, 1998, 2000, 2003) and Ray Jackendoff (1999, 2002). This account posits that modern language evolved from an earlier stage involving only single words, or simple concatenations of words, without any additional phrasal or grammatical structure. By such models, it is hypothesized a stage with a vocabulary, which is formed on the basis of pre-existing cognitive concepts (*i.e.* the prototypes of nouns and verbs), and a simple protolanguage, which allows these proto-words to be combined, but which does not constitute an actual grammar (Tallerman 2005, 2006, 2007).

In theories of this nature, words come first, and syntax emerges later, creating structure which utilizes these words; thus, in Bickerton's (1996: 51) view "syntax could not have come into existence until there was a sizeable vocabulary whose units could be organized into complex structures."

Protolanguage, in the synthetic account, exhibits the following properties:

- the order of elements is relatively random;
- no hierarchical syntactic structure constrains surface order;
- different word orders have no link to information structure (Tallerman & Gibson 2012: 482).

Although it is unquestionable that we did not inherit fossilized language output from our hominin ancestors, nevertheless we have the possibility to refer to *living fossils*, namely "types of communication used by modern humans that are close to, but do not share all the features of, fully-modern language" (Kirby 2009: 674). In this perspective, Bickerton (1995, 2009, 2010) proposes three subtypes of this living linguistic fossils:

- pidgins;
- the language of children under the age of two;
- the output of language-trained apes.

What he notes is that all these basic systems of communication, just as protolanguage, are characterized by simple words which can be combined into longer strings, with a minimal structure, a limited inflectional morphology and reduced stylistic or socio-linguistic variability (Dessalles 2000, 2006).

A very different perspective characterizes other recent works, mainly by Alison Wray (1998, 2000, 2008a, 2008b, 2010 (ed.)) and Michael Arbib (2002, 2003, 2005), proposing that words emerge from longer, completely arbitrary strings of sounds – holophrastic utterances – via a process of fractionation (Arbib 2003, 2005). In this "holistic" approach, protolanguage consists of a "small inventory" of holistic utterances, each representing "a complete communicative act" (Arbib *et al.* 2008: 154), exclusively associated with one proposition in which the meaning is taken from the whole and not from the sum of its parts, and in which there are no component units that could be recombined to create a new message. In fact, the parts are not considerable as words at all but just components of the utterance, similarly as the items *by*, *and* and *large* are not in any useful sense separate words in the expression *by and large* (Wray 1998, 2009). It is very significant to emphasize this point, because it is a demonstration that having no grammar implies having no component words and no novelty that can be spontaneously understood and shared by others. In a grammarless system, the only way to accomplish an efficient communication is to establish that a specific sequence of words is related to a certain meaning, and no other meanings should be taken into account for that particular sequence. Thus, if we are reasoning about a language without grammar, we should not start with words, but with the holistic utterances with an agreed meaning.

In contrast to the synthetic model, the holistic account "posit that the initial pairing between form and meaning mapped entire propositions to complex but undecomposable forms" (Fitch 2005: 218).

Complex form preceded words in the modern sense (Arbib 2003) and each utterance would be phonetically arbitrary, without any relation in sound to even those utterances with a similar meaning.

The utterance that we might translate as *give her the food* could be, say, /mèbita/ and that for *give me the food* /ikatubè/. Each utterance would be kept separate from others by contrasts that we might term ‘phonemic’, so, for example, /mabu/ might carry the meaning *keep away* while /madu/ meant *take the stick*. Protolanguage would, then, be a phonetically sophisticated set of formulaic utterances, with agreed function-specific meanings, that were a direct development from the earlier noises and gestures, and which had, like them, no internal structure. (Wray 1998: 51)

A language with no grammar may have no need of words, but it does need utterances and these, therefore, must have come first. Individual holistic signals, lacking in internal morphological structure, conveyed entire complex propositions, rather than semantic atoms; they are intended to have an effect upon the world of the speaker, by manipulation of other individuals through commands and threats, greetings and requests (Wray 1998: 52). Since a complete utterance, with an agreed meaning, is processed as if it were a single irreducible item, a relevant effect of the holistic strategy is that certain ways of expressing an idea become fixed as the preferred ones in the speech community.

Despite being non-intuitive, there is increasing evidence that such models should be taken seriously (Wray 2000), e.g. children’s one-word utterances operate as whole propositions (consider for example the infant’s “up:” it means “please pick me up,” and is certainly different from common adult’s meaning of the proposition “up”), or adult language: it is full of holophrasis (Jackendoff 1999; Wray 2000), and contains many phrases with the un-decomposable form/meaning pairings (consider “abracadabra” for the revealing of a magic trick, or “kick the bucket” as “to die”) (Fitch 2005).

### **Formulaic language: living clue for holistic protolanguage**

As seen above, Wray argues that the holistic processing strategy may predate the emergence of single referential words and the analytic grammar that combines them. In this perspective, first communications of our ancestors were completely holistic, and this fact limited them to using only a fixed set of routine manipulative utterances. Turning first to living fossils, Wray’s work draws attention to a common characteristic of language, namely the existence of formulaic utterances: although they may seem to be composed of individual words in the conventional way, these constructions cannot be changed internally, but operate as a whole, inseparable chunk.

Formulaic language plays a central role in our linguistic behaviour today. Much of what we say is prefabricated in multiword units for a swift retrieval, without any necessity to activate grammar rules (e.g. Wray 1998, 1999, 2000; Wray & Perkins 2000). In order to define formulaic sequences, Wray (2000) describes a range of types of pre-stored word string, including ones which:

- are undoubtedly holistic in nature, because they can not be generated by grammar, e.g. *by and large; to go the whole hog;*
- are grammatically sound but semantically holistic, e.g. *to pull someone’s leg; the oldest profession;*
- appear both grammatically and semantically ordinary but have transformational restrictions, e.g. *I didn’t sleep a wink, \*I slept a wink; He was fed up, \*The encounter had fed him up* (Wray & Perkins 2000);

- consist of a specific form of words with an agreed social function, e.g. *happy birthday; I now pronounce you man and wife;*
- are indistinguishable from novel utterances, except that, within a certain speech community or individual's idiolect, they are preferred over other equally possible formulations, e.g. *put the kettle on, will you?* vs. *please would you make me a hot drink?*

Even though some of these constructions could be analyzed, they are stored, retrieved and used as “prefabricated strings” referring to a whole situation, or aiming at a specific effect; they are processed quickly, and perceived as if they are single items rather than multi-part units. We use a notable number of these holistic utterances in everyday conversation, and the idea is that they speed up retrieval, production and processing time precisely because they do not require analysis. For this reason, although formulaic sequences are infrequently the only way of expressing a particular idea, they are certainly very often the preferred way. Thus, even if *Do not step on the lawn* and *Please perform an act of kindness for me* are comprehensible, we are much more likely to encounter *Keep off the grass* and *Would you do me a favour?*

However, the function related to formulae seems to be more relevant than form; e.g. consider that the meaning of *Can I help you?* is something more than a simple *Am I able to give an assistance to you?* So, it is possible to recognize three primary socio-functional classes, which indicate the way of using formulae in our language today:

- the incisive manipulation of others;
- group membership;
- fluency and holding the turn.

It seems that if utterances with certain kinds of social functions, such as requesting, ordering and negotiating, have a tendency to take on a preferred form. What they share is that they are used to elicit a particular response in the hearer; in fact we use them to get others to change our world for us, whether physically (*pass me the salt; present arms!*), mentally (*tell me what happened; could you repeat that?*) or emotionally (*say you love me; leave me alone*). As such, it seems plausible that the advantage in their being formulaic is that it makes them easier for the hearer to recognize and decode, something that is clearly entirely in the speaker's interest in this kind of directive utterance. Thus, such formulae seem to survive exactly because they are unanalyzed, even sometimes unanalyzable in term of the speaker's grammar (Tallerman 2007: 15; Kirby 2009).

Therefore, a striking proportion of formulaic expressions are used to manipulate others into physical, emotional, and perceptual reactions. The manipulative functions of our formulaic language correspond closely with those observed in the communicative behaviour of chimpanzees in the wild. As seen above, in holistic protolanguage, each unit would have been a longer, indivisible and arbitrary string of sounds which constituted a whole proposition. Just as a non-human primate call might mean “I want to play with you,” or “Beware, there's a predator!”, so a protolinguistic string would correspond in full to a warning, a greeting, or a request. Chimpanzees use their noise and gesture system to bring out changes in their world, maintaining social structure and express the place of the individual within it. It seems that the functions of holistic utterances in chimp communication represent a subset of those in human language. In both species, in fact, they are used for social interaction, where their purpose is the manipulation of the hearer, either to act in the interest of, or to recognize the identity and status of, the speaker (Wray 2000; Knight *et al.* (eds.) 2000).

On the basis of this parallel, it is possible to argue that the holistic strategy for expressing manipulative messages in phonetic form may be considerably more ancient than the analytic strategy.

The holistic cries and gestures of our pre-human ancestors were transformed, over a long period of time, into a phonetically expressed set of holistic message strings, each with a manipulative function such as greeting, warning, commanding, threatening, requesting, appeasing. The holistic delivery of such messages is something that we still prefer, though, clearly, our holistic forms today are not direct descendants of those original ones. What we have inherited is not the forms themselves, but the *strategy* of using holistic linguistic material to achieve these key interactive functions. We resolve modern problems of interaction (including that of getting the hearer to react in a desired way) in part by using an ancient holistic processing strategy, applied to our modern linguistic formulation. (Wray 2010: 115)

Even if there is the reasonable possibility that there is nothing in common between the holistic language we use today and the holistic communication used by primates, but the fact to be both holistic, when we compare the functions of formulaic sequences today with those for which holistic noise-gesture utterances appear to be used in primates, we find a strikingly close correspondence (Reiss 1989).

On the base of this *cross-species* correlation, it does not indeed seem unreasonable to suppose that in protolanguage too, everyday social interactions were achieved by means of holistic utterances.

If the holistic system is not taken into account when examining linguistic performance, its effectiveness in transmitting semantically complex messages at relatively restricted processing cost may be erroneously interpreted as a clue of a greater analytic competence in the language than actually exists (McMahon & McMahon 2013). This is clarified when we consider the case of our ancestors: despite the absence of grammar, their interaction was both complex and subtle.

This section demonstrates that modern formulaic language plays significant roles that pivot on social interaction. It is reasonable to ask why should we need them, now that we rely upon a grammar-based analytic system. The most convincing answer is that “so that we do not have to go through the labour of generating an utterance all the way out from ‘S’ every time we want to say something” (Becker 1975: 17). Thus, during our usual communications, we seem to have the recourse to holistic processing, not because using an analytic system is impossible, but because it is an expensive strategy (Wray 1998: 63).

## Conclusions

The aim of this paper was to take a position into the ongoing debate about whether the words in the first language spoken by humans expressed single concepts or complex holophrases.

We argued that the holistic processing strategy, used for achieving manipulative interaction, has to be considered as predating the emergence of single referential words and the analytic grammar that combines them. Early interactions of our ancestors were completely holistic, *i.e.* they are restricted to use only a fixed set of routine manipulative utterances. In this perspective, we analyzed the case of a formulaic language as a living linguistic fossil in support of the holistic strategy. In fact, formulaic expressions are geared to reducing the speaker’s processing effort and to facilitate the hearer’s processing and reaction. Thus, the hearer recognizes the string and decodes it holistically; in this way its length and internal structure are no hindrance to swift processing, and that increases the likelihood of successful comprehension.

## References

- Arbib, Michael A. (2002) *The Mirror System, Imitation and the Evolution of Language*. Cambridge, MA: The MIT Press.
- Arbib, Michael A. (2003) *The Handbook of Brain Theory and Neural Networks*. Cambridge, MA: The MIT Press.
- Arbib, Michael A. (2005) "From Monkey-Like Action Recognition to Human Language: An Evolutionary Framework for Neurolinguistics." [In:] *Behavioral and Brain Sciences* 28 (2); 105–167.
- Arbib, Michael A., Katja Liebal, Simone Pika (2008) "Primate Vocalization, Gesture and the Evolution of Human Language." [In:] *Current Anthropology* 49 (6); 1053–1076.
- Becker, Joseph (1975) *The Phrasal Lexicon*. Cambridge, MA: Beranek & Newman.
- Bickerton, Derek (1990) *Language and Species*. Chicago: University of Chicago Press.
- Bickerton, Derek (1995) *Language and Human Behavior*. London: UCL Press.
- Bickerton, Derek (1996) *Language and Human Behavior*. London: UCL Press.
- Bickerton, Derek (1998) "Catastrophic Evolution: The Case for a Single Step from Protolanguage to Full Human Language." [In:] James R. Hurford, Michael Studdert-Kennedy, Chris Knight (eds.) *Approaches to the Evolution of Language: Social and Cognitive Bases*. Cambridge: Cambridge University Press; 341–358.
- Bickerton, Derek (2000) "How Protolanguage Became Language." [In:] Chris Knight, Michael Studdert-Kennedy, James R. Hurford (eds.) *The Evolutionary Emergence of Language. Social Function and the Origins of Linguistic Form*. Cambridge: Cambridge University Press; 264–284.
- Bickerton, Derek (2003) "Symbol and Structure: A Comprehensive Framework for Language Evolution." [In:] Morten H. Christiansen, Simon Kirby (eds.) *Language Evolution*. Oxford: Oxford University Press; 77–93.
- Bickerton, Derek (2009) *Adam's Tongue: How Humans Made Language, How Language Made Humans*. Farrar: Straus and Giroux.
- Bickerton, Derek (2010) "Foraging versus Social Intelligence in the Evolution of Protolanguage." [In:] Alison Wray (ed.) *The Transition to Language. Studies in the Evolution of Language*. Oxford: Oxford University Press; 207–225.
- Calvin, William, Derek Bickerton (2000) *Reconciling Darwin and Chomsky with the Human Brain*. Cambridge, MA: The MIT Press.
- Dessalles, Jean-Louis (2000) *Aux origines du langage*. Paris: Hermès Science Publications.
- Dessalles, Jean-Louis (2006) "Du protolangage au langage: modèle d'une transition." [In:] *Marges linguistiques* 11; 142–152.
- Fitch, William Tecumseh (2005) "The Evolution of Language: A Comparative Review." [In:] *Biology and Philosophy* 20; 193–230.
- Fitch, William Tecumseh (2010) *The Evolution of Language*. Cambridge: Cambridge University Press.
- Jackendoff, Ray (1999) "Possible Stages in the Evolution of the Language Capacity." [In:] *Trends in Cognitive Sciences* 3 (7); 272–279.
- Jackendoff, Ray (2002) *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford: Oxford University Press.
- Kirby, Simon (2009) "The Evolution of Language." [In:] Robin Dunbar, Louise Barrett (eds.) *The Oxford Handbook of Evolutionary Psychology*. Oxford: Oxford University Press; 669–681.
- Knight, Chris, Michael Studdert-Kennedy, James R. Hurford (eds.) (2000) *The Evolutionary Emergence of Language. Social Function and the Origins of Linguistic Form*. Cambridge: Cambridge University Press.
- McMahon, April, Robert McMahon (2013) *Evolutionary Linguistics*. Cambridge: Cambridge University Press.

- Reiss, Nira (1989) "Speech Act Taxonomy, Chimpanzee Communication, and the Evolutionary Basis of Language." [In:] Jan Wind, Edward G. Pulleyblank, Éric de Grolier, Bernard H. Bichakjian (eds.) *Studies in Language Origins*. Vol. 1. Amsterdam: John Benjamins; 283–304.
- Smith, Kenny (2006) "The Protolanguage Debate: Bridging the Gap?" [In:] *Evolution of Language. Proceeding of the 6th International Conference (Evolang6)*. World Scientific Publishing; 315–322.
- Tallerman, Maggie (2005) *Language Origins. Perspectives on Evolution*. New York: Oxford University Press.
- Tallerman, Maggie (2006) "A Holistic Protolanguage Cannot Be Stored, Cannot Be Retrieved." [In:] *Evolution of Language. Proceeding of the 6th International Conference (Evolang6)*. World Scientific Publishing; 447–449.
- Tallerman, Maggie (2007) "Did Our Ancestors Speak a Holistic Protolanguage?" [In:] *Lingua* 117 (3); 579–604.
- Tallerman, Maggie, Kathleen Gibson (2012) *The Oxford Handbook of Language Evolution*. New York: Oxford University Press.
- Wray, Alison (1998) "Protolanguage as a Holistic System for Social Interactions." [In:] *Language & Communication* 18; 47–67.
- Wray, Alison (1999) "Formulaic Language in Learners and Native Speakers." [In:] *Language Teaching* 32 (4); 213–231.
- Wray, Alison (2000) "Holistic Utterances in Protolanguage: The Link from Primates to Humans." [In:] Chris Knight, Michael Studdert-Kennedy, James R. Hurford (eds.) *The Evolutionary Emergence of Language. Social Function and the Origins of Linguistic Form*. Cambridge: Cambridge University Press; 285–302.
- Wray, Alison (2010) "Dual Processing in Protolanguage: Performance without Competence." [In:] Alison Wray (ed.) *The Transition to Language. Studies in the Evolution of Language*. Oxford: Oxford University Press; 113–137.
- Wray, Alison (2008a) "Genes and the Conceptualization of Language Knowledge." [In:] *Genomics, Society and Policy* 4 (1); 58–73.
- Wray, Alison (2008b) "The Puzzle of Language Learning: From Child's Play to 'Linguaphobia.'" [In:] *Language Teach* 41 (2); 253–271.
- Wray, Alison (2009) "Future Directions in Formulaic Language Research." [In:] *Journal of Foreign Languages* 32 (6); 2–17.
- Wray, Alison (ed.) (2010) *The Transition to Language. Studies in the Evolution of Language*. Oxford: Oxford University Press.
- Wray, Alison, Michael R. Perkins (2000) "The Functions of Formulaic Language: An Integrated Model." [In:] *Language & Communication* 20 (1); 1–28.