LINKING THE PROCESSES OF LANGUAGE EVOLUTION AND LANGUAGE CHANGE: A FIVE-LEVEL HIERARCHY

JEROME MICHAUD

ICMCS, School of Physics and Astronomy, University of Edinburgh,
Peter Guthrie Tait Road, Edinburgh EH9 3FD UK
v1jmicha@staffmail.ed.ac.uk

The class of problems that can be categorized under *language evolution and change* is very heterogeneous and involves many different timescales and spatial/social scales. In order to better understand the underlying evolutionary processes and the link between language evolution and language change, we introduce a five-system hierarchy based on social structures of the population and apply it to the study of the transition from protolanguage to language. This problem is composed of a biological subproblem and a cultural subproblem. Using our hierarchy, we argue that Bickerton's "living fossils" of language can give some insight only on the cultural subproblem, but not on the biological one or on the coevolution of the two.

1. Introduction

The language faculty is at the heart of what makes us human. It has permitted us to develop increasingly complex societies and cultures, but the way this faculty emerged and subsequently evolved is still to a large extent a mystery, see Hauser et al. (2014). This is due to the complexity of the problem: language evolution and change is not a single and well-defined process. It covers a number of phenomena ranging from language's first appearance in mankind to language acquisition by children, through language use, cultural evolution of languages, language death, new dialects, creole and pidgin formation, to name but a few. In the literature, one usually refers to the language emergence problem by language evolution and to its subsequent evolution by language change. The heterogeneity of language evolution and change problems is such that a single mechanism of evolution is unlikely to be sufficient to explain them all; one therefore has to rely on many different evolution processes. In Kirby (2007), Christiansen and Kirby (2003) it is shown that the biological, cultural and individual evolution processes are the three main processes that drive the evolution of languages. In order to understand the emergence and evolution of a language one has to understand the interplay of these three kinds of processes. As a result, language is a *complex adaptive system*, (Cornish, Tamariz, & Kirby, 2009; Beckner et al., 2009; Steels, 2000).

Given the complexity of language evolution and change problems, it is sometimes difficult to identify which processes are the most relevant to study a specific

question. The main aim of this paper is to provide a systematic approach to analyse a language evolution problem and to identify the most relevant evolutionary mechanisms. Some attempts to achieve this goal have been for example discussed in Enfield (2014) in the broader context of causal relations in language. Our approach is based on an evolutionary point of view, which is best described when adopting a usage-based definition of language. Starting from the three types of evolutionary processes, we then identify five different levels of description characterized by their timescale and their population structure. Joining these two components provides a systematic approach to analyse a given question in language evolution.

In order to illustrate how this new approach can be applied, we consider the problem of the transition from protolanguage to language and discuss in the light of our framework the relevance of different *living fossils* of language first introduced by Bickerton (1995, appendix A). In this paper, the existence of some kind of protolanguage is taken for granted and we will not discuss it in detail. The purpose of this example is to shed light on the relations between language evolution and language change problems.

2. Evolutionary processes and the five-level hierarchy

2.1. *Usage-based definition of language and social structures*

In this paper, we assume that language is a conventionalised semiotic system composed of LINGUEMES (Croft, 2000). The representation that every individual have of the language is conditioned on her history and is usually called an IDIOLECT. The conventionalisation process mediated by language usage occurs in a complex population. One can identify the homogeneous subgroups of this population as SPEECH COMMUNITIES and their averaged language is called a CODE (Croft, 2000). These subgroups have the property to be both well-connected and socially homogeneous. The next level of social structure is when the population is still well-mixed, but socially structured, that is, when there are many speech communities. In this case, the corresponding subgroups form LINGUISTIC COMMUNITIES. The linguistic system used at this level corresponds to what we call a LANGUAGE. If different subgroups are only loosely connected, many linguistic communities emerge and languages are in contact. This is for example the case in creole and pidgins. We assume that the different subgroups have similar language capacities. This is no longer the case if we consider different interacting species. The level of development of their LANGUAGE FACULTY is different and it is not even clear if other species have such a faculty.

2.2. *Evolutionary account*

With respect to language evolution, the biological evolutionary processes are concerned with brain evolution and the biological component of our language faculty;

cultural processes are responsible for the propagation and the conventionalization of the language through a population and depend on the social and/or cultural structures of the population; individual processes are responsible for language learning and language use. They are influenced by the cognitive architecture and by the psychology of the individuals. Following Kirby (2007), Christiansen and

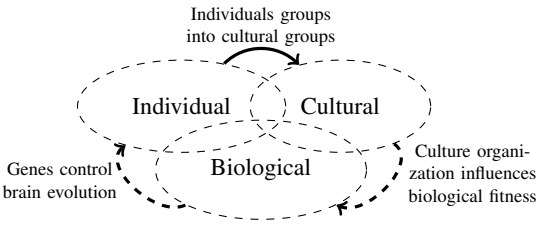

Figure 1. Languages evolve and change by three main processes: individual evolution, cultural evolution and biological evolution. The dashed arrows mean that links are weaker than the others. This figure has been adapted from Kirby (2007), Christiansen and Kirby (2003)

Kirby (2003), Figure 1 sketches the relation between the different types of evolutionary processes. The link between cultural and biological evolution is not really transparent and deserves some comments. In fact, it is not clear how cultural evolution affects the biological fitness and therefore the better survival of the population. Furthermore, the link between biological and individual evolution is weak because of the huge timescale difference between the two types of processes.

2.3. *The five systems of language evolution and change*

Comparing the language definition and the corresponding social structures with the three evolutionary processes mentioned above, one can assign to every social structure the main corresponding evolutionary mechanism together with the language type and corresponding system as shown in Table 1. An alternative

Table 1. Correspondence between social structures, language type, systems and evolutionary processes.

Level	**Social struct.**	**Lang. type**	**System**	**Evo. process**
1	Ind.	Idiolect	Ling. System	Individual
2	Homo.	Code	Multi-agent	Cultural
3	Struct.	Language	Syst. of Speech Com.	Cultural
4	Hetero.	Set of lang.	Syst. of Ling. Com.	Cultural
5	Homo.	Lang. faculty	Ecosystem	Biological

representation of this table is given in Figure 2 that represents the five levels of language evolution and change. The solid (red) arrows means "groups into" and each level can be considered as a system where the parts are elements of the previous level, see Table 1. With this association, one can study the dynamics of a

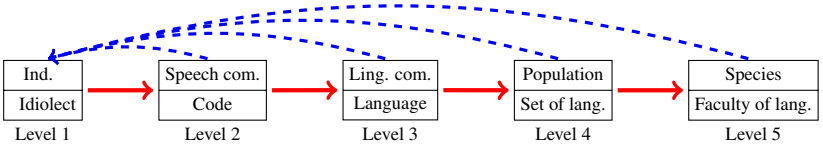

Figure 2. Illustration of the structures underlying the existence of a language in complex populations. The solid (red) arrows means "groups into" and the dashed (blue) arrows symbolize the influence of higher level structures on the smallest one as an individual is aware of the population structures and this influences her behavior.

system by studying the interaction between its parts. Consider for example level 4, a population is a system of linguistic communities (level 3) that interact through the creation and evolution of inter-linguistic speech communities (level 2). The dashed (blue) arrows in Figure 2 underline the fact that an individual is aware of the population structures and acts accordingly. There is therefore not only a micro-to-macro connection between the systems, but also a macro-to-micro relationship. Note that Table 1 corresponds to the state of a fully developed language. At earlier stages, the distinction between levels 2-4 only relies on the social structures of the population.

3. Case study: Protolanguage to language transition

In order to illustrate how this hierarchy of systems can be applied in practice, we consider the case of the transition from protolanguage to language. Since this is a difficult problem, some assumptions are discussed below.

3.1. *Assumptions*

We assume the existence of a non-holistic protolanguage (Bickerton, 1995; Jackendoff, 1999), that is, a low-structured linguistic system composed of words that can be juxtaposed, but lacking any kind of higher-level grammatical structures. Another kind of protolanguage has been proposed by Wray (1998). She proposed a holistic view of protolanguage, where each utterance is unstructured, that is, it can not be decomposed into words. In the literature, the question of the nature of protolanguage, holistic or not, has been discussed extensively, see for example Hurford (2000), Arbib (2005), Tallerman (2007).

We also assume that the transition from protolanguage to language is gradual, see for example (Kirby, 2007; Sandler et al., 2014) and contrary to Davidson and Noble (1994). One comment has to be made here: since the transition is gradual the categorical distinction between a protolanguage and a language is somehow misleading. To avoid misinterpretation, when we speak about protolanguage, we always mean a low-structured language, whereas a language is high-structured. The emergence of complex grammatical structures seems to be in contradiction with the gradual evolution of the language faculty. In fact, this can be understood

since we are dealing with a complex system. One characteristic of such systems is to have phase transitions, that is, some minor changes can fundamentally change the behaviour of the system. In other words, the emergence of the faculty to handle complex structures may be the consequence of small changes in the brain organisation or be an artefact of the timescale difference between cultural and biological evolution.

3.2. *Analyse of the protolanguage to language transition*

We can now analyse the protolanguage to language transition. In the light of our hierarchy of systems, this transition assumes a change in the cognitive abilities of the population. The phylogenetic change therefore belongs to level 5 of Figure 2[a]. This change occurs in the ecosystem to which humans belong. The emergence of the cognitive ability to deal with complex structures and the conditions in which this faculty fixed in the human genotype remains to a large extent a mystery (Hauser et al., 2014). It is really difficult to decide whether this faculty emerges by adaptation, exaptation or simply by random drift.

It is important to note that having the faculty to process complex utterances does not directly imply the existence of a complex linguistic system. This faculty is necessary, but not sufficient. In order for a fully-fledged language to emerge, complex structures have to be innovated and conventionalised by the population. The problem of protolanguage to language transition can therefore be decomposed into two interconnected subproblems: (i) understanding the genetic changes that enable a faculty to deal with complex language structures and (ii) explaining the emergence of a complex language from a protolanguage by a succession of innovation and conventionalisation processes. The first subproblem is biological and the second is cultural. I will argue that "living fossils" of language can be used to gain insight into the second problem, but the first one seems to be out of reach.

Going back to our hierarchy of systems, the first subproblem belongs to level 5, but the second problem belongs to lower levels 2-4, since it is concerned with a cultural process. In fact, we consider that the transition from a protolanguage to a language occured through a series of language contact problems where (proto)languages of different complexity were in contact. For a successful transition to occur, more complex languages should have outcompeted simpler ones. This problem of language competition therefore belongs to level 4 and should be studied using inter-linguistic speech communities and their internal dynamics (level 2, multi-agent system).

Since the biological and the cultural evolution processes coevolved, the capacity to deal with more complex languages is conditioned on the genetic characteristics of the different linguistic communities and many generations are needed to complete a transition from a low-structured to a high-structured language.

[a] From now on, the mention of levels assumes the reference to Figure 2.

3.3. *Living fossils*

We now come to the discussion of "living fossils" of language. We discuss the three original examples given by Bickerton (1995, App. A). These three protolanguages are: (i) child language, (ii) pidgin and creole languages and (iii) trained apes language. These living fossils have been claimed by different authors to shed light on the protolanguage to language transition. Such a parallel is for example suggested by Kirby (2007, p.9). In principle one could choose other examples that are probably better examples of protolanguages such as home sign or developing sign languages, see (Sandler et al., 2014; Kegl, Senghas, & Coppola, 1999). However, this choice has the advantage to span all the different levels proposed in our hierarchy and therefore best demonstrates the power of our approach.

3.3.1. *Child language*

In the case of child speech evolution, the transition from a low-structured to a high-structured language is mediated by the learning process. This process is internal to the family speech community (level 2) that can be further decomposed into the children speech community and the adults' one. That is, the systems which are involved in this process are the linguistic system (level 1) and the multiagent system (level 2). The process analogous to the conventionalisation problem is learning. This analogy is fairly weak, since learning and conventionalisation occur at two different levels of the hierarchy and through different evolutionary processes. Learning is driven by individual evolution with contact with a speech community of speakers of a fully-fledged language, whereas the conventionalisation problem occurs in a relatively well-mixed population of people with similar language abilities. This parallel is therefore pretty weak. However, learning also played a role in the transition from protolanguage to language since each new generation had to learn to communicate with the previous generations. The process analogous to the biological process of language emergence (level 5) is brain maturation (level 1). In our hierarchy, level 5 and level 1 are very different. For this analogy to be explanatory, one has to accept Haekel's *recapitulation theory* (Haekel, 1874), which states that: "in developing from embryo to adult, animals go through stages resembling or representing successive stages in the evolution of their remote ancestors", that is, level 1 represents level 5. This is sometimes referred to as the *ontogeny recapitulates phylogeny* assumption (Mufwene, 2008). This theory could justify the biological analogy and it has been shown to be plausible to a certain extent by Kalinka and Tomancak (2012). Note that the recapitulation theory has been disputed in the literature, see for example Payne and Wenger (1998), Løvtrup (1978) or Mufwene (2008). As a result, one can say that the conventionalisation problem between children and adults is relevant for the protolanguage to language transition since during the transition, many generations of learners have been involved. Explanations of biological evolution heavily relies on the recapitulation theory,

but this theory says nothing about the selection mechanisms that have constrained brain evolution. Therefore, child language can not give insight into the biological evolution problem or its coevolution with cultural processes.

3.3.2. *Pidgin/Creoles*

In the pidgin case, the transition between a low-structured and a high-structured language (level 5, since it is assumed to be biologically encoded) is arguably in analogy with the pidgin to creole transition assumed by Mühlhäusler (1986). In the five-level hierarchy, the *creolisation* process involves a system of linguistic communities (level 4) that interact through the creation of speech communities (level 2). The structure of this speech community and its evolution is not much discussed in the literature. But in Mühlhäusler (1986), a pidgin is characterized by the fact that it does not have native speakers, whereas creole languages do. This means that a transition between a pidgin and a creole needs to be driven by a change in the inter-linguistic speech community structure. There is no reason for this structural change to be necessary (Mufwene, 2008). Moreover, it is possible that the inter-linguistic speech community always includes children and in this case, no pidgin will ever be formed and a creole or another type of mixed language can directly emerge. As a result, the transition from pidgin to creole is not necessary and might even be wrong (Mufwene, 2008).

On the one hand, pidgins and creoles are formed inside inter-linguistic speech communities through conventionalisation processes that might be close to those involved in the cultural evolutionary processes active in the protolanguage to language transition, at least inside the homogeneous inter-linguistic community. However, the presence of fully-fledged languages in surrounding communities is weakening this analogy. On the other hand, there is clearly no biological process in this case and the biological subproblem of the protolanguage to language transition has no analogue. As a result, this language fossil might be useful to explain the cultural aspect of the protolanguage to language transition, but the biological aspect and the coevolution of the two seems to be out of reach.

3.3.3. *Trained Apes*

The third language living fossil that Bickerton has proposed is the trained animal language, in particular, the trained apes language illustrated by the case of Koko, a female lowland gorilla, see Patterson (1978). This case is fairly close to the pidgin/creole case, since it concerns a contact between two linguistic communities (level 4), the animal one and the human one, and the subsequent evolution of the corresponding code (level 2). The structure of the resulting code is quite close to the structure of a pidgin (Bickerton, 1995). It is interesting to note that when the animal starts to learn a human language when it is still young, which is the case of Koko, then the structures of the emerging code are more complicated, reflecting in

a different situation the same kind of distinction as between a pidgin and a creole.

The trained apes case can be argued to belong to level 5 of the hierarchy since it involves two different species. However, the type of evolutionary processes involved is cultural, not biological, between two linguistic communities with a different language faculty. This situation is quite close to the competition between communities with different language faculties as is the case in the protolanguage to language transition. The study of this fossil is therefore informative on the cultural evolutionary aspect. As for the pidgin/creole case, the existence of a fully-fledged language is weakening the analogy and one has to be careful when interpreting the results. Furthermore, the main problem is the biological evolutionary aspect of the protolanguage to language transition. If one wants to make a useful analogy, the Haekel's recapitulation theory (Haekel, 1874) has to be assumed and the same problems as in language acquisition appear. One can therefore conclude that the cultural aspect of the protolanguage to language transition can be partially studied through experiments with animals, even though a fully-fledged language exists in this case. However, the biological evolutionary aspect of the protolanguage to language transition and the coevolution between biological and cultural processes seems once again to be out of reach.

4. Discussion and concluding remarks

In this paper, we have introduced a five-level hierarchy of systems that serves as a tool to analyse problems in language evolution and change. This tool allows us to quickly and systematically identify the relevant social structures underlying the evolutionary problem at hand. As an illustration, we have considered the protolanguage to language transition and have discussed the relevance of the three living fossils proposed by Bickerton (1995). We have shown that the "living fossils" proposed can give some insight on the cultural evolution subproblem of the protolanguage to language transition, but not on the biological aspect or the coevolution of the cultural and biological problems, since this assumes the acceptance of Haekel's recapitulation theory. This comparison provides an important link between language evolution and language change processes.

Our example corresponds to a critical application of the five-level hierarchy of systems. One can also apply it in a productive way as a guiding principle to develop evolutionary scenarios or mathematical models of language change. For example, the process of creation and propagation of a convention in a speech community should be treated in a cultural evolution framework, not in a biological one as it is for example the case in the work of Nowak and Krakauer (1999).

Acknowledgements

I would like to thank Richard A. Blythe and Edwige Dugas for interesting discussions and comments. This work was support by the Swiss National Science Foundation under grant number P2GEP2_159156.

References

Arbib, M. A. (2005). From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral and brain sciences*, *28*(02), 105–124.

Beckner, C., Blythe, R. A., Bybee, J., Christiansen, M. H., Croft, W., Ellis, N. C., Holland, J., Ke, J., Larsen-Freeman, D., & Schoenemann, T. (2009). Language is a complex adaptibe system: Position paper. *Language Learning*, *59*, Suppl. 1, 1–26.

Bickerton, D. (1995). *Language and human behavior.* University of Washington Press.

Christiansen, M. H., & Kirby, S. (2003). Language evolution: Consensus and controversies. *Trends in cognitive sciences*, *7*(7), 300–307.

Cornish, H., Tamariz, M., & Kirby, S. (2009). Complex adaptive systems and the origins of adaptive structure: What experiments can tell us. *Language Learning*, *59*, Suppl. 1, 187–205.

Croft, W. (2000). *Explaining language change: An evolutionary approach.* Pearson Education.

Davidson, I., & Noble, W. (1994). In K. R. Gibson, K. R. Gibson, & T. Ingold (Eds.), *Tools, language and cognition in human evolution* (pp. 125–155). Cambridge University Press.

Enfield, N. J. (2014). *Natural causes of language: Frames, biases, and cultural transmission.* Language Science Press.

Haekel, E. (1874). *Anthropogenie; oder, entwickelungsgeschichte des menschen, keimes- und stammesgeschichte.* Leipzig,W. Engelmann.

Hauser, M. D., Yang, C., Berwick, R. C., Tattersall, I., Ryan, M. J., Watumull, J., Chomsky, N., & Lewontin, R. C. (2014). The mystery of language evolution. *Frontiers in psychology*, *5*.

Hurford, J. R. (2000). Social transmission favours linguistic generalization. *The evolutionary emergence of language: Social function and the origins of linguistic form*, 324–352.

Jackendoff, R. (1999). Possible stages in the evolution of the language capacity. *Trends in cognitive sciences*, *3*(7), 272–279.

Kalinka, A. T., & Tomancak, P. (2012). The evolution of early animal embryos: conservation or divergence? *Trends in Ecology & Evolution*, *27*(7), 385 - 393.

Kegl, J., Senghas, A., & Coppola, M. (1999). *Creation through contact; sign language emergence and sign language change in nicaragua.* Cambridge, MA: MIT Press.

Kirby, S. (2007). The evolution of language. *Oxford handbook of evolutionary psychology*, 669–681.

Løvtrup, S. (1978). On von baerian and haeckelian recapitulation. *Systematic*

Biology, *27*(3), 348–352.
Mufwene, S. S. (2008). *Language evolution: Contact, competition and change.* Bloomsbury Publishing.
Mühlhäusler, P. (1986). *Pidgin and creole linguistics.* Blackwell Oxford.
Nowak, M. A., & Krakauer, D. C. (1999). The evolution of language. *Proceedings of the National Academy of Sciences*, *96*(14), 8028–8033.
Patterson, F. G. (1978). The gestures of a gorilla: Language acquisition in another pongid. *Brain and Language*, *5*(1), 72–97.
Payne, D. G., & Wenger, M. J. (1998). *Cognitive psychology.* Houghton Mifflin College Division.
Sandler, W., Aronoff, M., Padden, C., Meir, I., Enfield, N., Kockelman, P., & Sidnell, J. (2014). Language emergence: Al-sayyid bedouin sign language. *Cambridge handbook of linguistic anthropology*.
Steels, L. (2000). Language as a complex adaptive system. In M. Schoenauer et al. (Ed.), *Parallel problem solving for nature–ppsn vi* (Vol. 1917, pp. 17–26). Berlin: Springer.
Tallerman, M. (2007). Did our ancestors speak a holistic protolanguage? *Lingua*, *117*(3), 579–604.
Wray, A. (1998). Protolanguage as a holistic system for social interaction. *Language and Communication*, *10*(1), 47–67.