LONG-DISTANCE DEPENDENCIES ARE NOT UNIQUELY HUMAN

RAMON FERRER I CANCHO

Dpt. de Llenguatges i Sistemes Informàtics, Universitat Politècnica de Catalunya. Campus Nord, Edifici Omega. Jordi Girona Salgado 1-3. 08034, Barcelona, Spain.

VÍCTOR M. LONGA

Dpto. de Literatura Española, Tª da Literatura e Lingüística Xeral, Universidade de Santiago de Compostela, Pr. Isabel a Católica 2, 2ºE, 36204, Vigo, Spain

GUILLERMO LORENZO

Dpto. de Filología Española, Universidad de Oviedo, Teniente A. Martínez s/n, 33011, Oviedo, Spain

It is widely assumed that long-distance dependencies between elements are a unique feature of human language. Here we review recent evidence of long-distance correlations in sequences produced by non-human species and discuss two evolutionary scenarios for the evolution of human language in the light of these findings. Though applying their methodological framework, we conclude that some of Hauser, Chomsky and Fitch's central claims on language evolution are put into question to a different degree within each of those scenarios.

1. Introduction

Hauser, Chomsky and Fitch's (2002) picture of the faculty of language (FL) and their distinction between a broad (FLB) and a narrow (FLN) sense of the faculty is basically a useful tool for guiding research projects on language evolution, even if not always understood as such (Pinker & Jackendoff 2005 and Jackendoff & Pinker 2005). Their main contention is that explaining the evolutionary origins of language requires, as a first step, to tell apart the formal features of languages which can be thought of either as (1) inherited unchanged from a common ancestor, (2) subjected to minor modifications, or (3) qualitatively new (Hauser, Chomsky & Fitch 2002: 1570). Features reasonably put under categories (1) and (2) are said to be part of the faculty of language in the broad sense (FLB), while features suspected of pertaining to category (3) are

said to be part of the faculty of language in the narrow sense (FLN). But the import of the distinction is first and foremost methodological: it is a useful criterion for deciding which aspects of language can be explained in relation with the evolutionary history of other non-human species and which aspects of language can not be illuminated in this way and thus demand somehow special explanations (Fitch, Hauser & Chomsky 2005: 181). A second and in some way stronger contention of Hauser, Chomsky and Fitch has to do with the contents of FLN. They believe that the only serious candidate to be included within this category is the recursive procedure that the computational system of human language makes use of, with its open-ended generativity based on the structural embedding of hierarchically organized phrases (Hauser, Chomsky & Fitch 2002: 1573). But, again, it is important to note that Hauser, Chomsky and Fitch's proposal is mostly put forward as a guideline to direct testable hypothesis in order to refute or validate the claim that recursion is the only real novelty that language evolution has introduced into the natural world (Hauser, Chomsky & Fitch 2002: 1578). In this sense, it is worth remembering that they also advance an alternative hypothesis, according to which FLN is nothing but a rich set of interconnected mechanisms, all shared with other non-human species but only put together in the course of human evolution (Hauser, Chomsky & Fitch 2002: 1578 and Fitch, Hauser & Chomsky 2005: 181), a possibility also exposed to empirical refutation.

This article is not directly aimed at recursion, but at another highly distinctive feature of human language, actually also dubbed by some authors as language specific: long-distance dependencies (henceforth, LDDs) among the items of a sequence. The article is organized as follows. Section 2 offers what we consider the linguists' consensus view on LDDs as a unique feature of human language. Section 3 discusses the link between LDDs and the complexity of human language. Section 4 is devoted to review some recent evidence of long-distance correlations in sequences produced by non-human organisms, which point to the conclusion that this feature of human language is to be classified as pertaining to the domain of FLB. Section 5 is devoted to discuss some consequences of this conclusion for the evolutionary understanding of the faculty of language. These consequences, we contend, are far reaching, attending to the fact that LDDs are usually connected with the existence of non-trivial recursion. This final section depicts two scenarios for the evolution of the faculty of language, both challenging, though to a different degree, in relation to Hauser, Chomsky and Fitch's distinctions and contentions.

2. LDDs and human language: the consensus view

LDDs are a specially pervasive feature of human language, which adopts many different faces: agreement (*John$_i$ often drinks$_i$ wine*), binding (*John$_i$ wonders which picture of himself$_i$ has been stolen*), control (*Mary$_i$ never promised PRO$_i$ to kiss John*), displacement (*which students$_i$ did the president say the police arrested t_i last week?*), among others. What they all have in common is that they imply a relation between two items, one of them to be valuated by the other within a certain search space or domain non-linearly but structurally defined (see Chomsky 2000, where the unifying notion of *Agree* is proposed). Chomsky contends that this property of language seems rather unexpected in that it is "never built into special-purpose systems and apparently without significant analogue elsewhere" (Chomsky 2000: 101). Chomsky's claim is by no means exceptional. Hauser, Chomsky and Fitch also defend that "natural language go beyond purely local structure by including a capacity for recursive embedding of phrases within phrases, which can lead to statistical regularities that are separated by an arbitrary number of words or phrases. Such long-distance, hierarchical relationships are found in all natural languages for which, at a minimum, a phrase-structure grammar is necessary" (Hauser, Chomsky & Fitch 2002: 1577), thus establishing a causal connection between LDDs and the sort of complex syntax characteristic of human language.

In the same vein, Anderson relates the existence of LDDs with that of grammars with recursive phrase structure (Anderson 2004: 203) and contends that this kind of grammars are language-specific (Anderson 2004: 217-218). Not very different are Berwick's opinions, who thinks that *Merge* is to be considered as the ultimate responsible of many of the distinguishing properties of human syntax, among others recursion and displacement, one of the varieties of LDDs (Berwick 1998: 322). In addition, considering LDDs as language-specific is one of the uncommon points of agreement between Hauser, Chomsky and Fitch, on the one hand, and Pinker and Jackendoff, on the other hand. In their own words: "a final important device is long-distance dependency, which can relate a question word or relative pronoun to a distant verb, as in *Which theory did you expect Fred to think Melvin had disproven last week?*, where *which theory* is understood as the object of *disprove*. Is all this specific to language? It seems likely, given that it is specialized machinery for regulating the relation of sound and meaning. What other human or non-human ability could it serve?" (Pinker & Jackendoff 2005: 216).

All these pronouncements speak clearly of a consensus among linguists (yet rarely declared) concerning the language specificity of LDDs and its

inclusion (even if implicitly) within FLN, at least by those who accept the notion.

3. LDDs and the complexity of human language

Our understanding of the complexity of human language is largely influenced by Chomsky's hierarchy of grammatical complexity (Hopcroft & Ullman 1979). Chomsky's hierarchy consists of the following levels (in decreasing order of potential complexity):

- Type-0 grammars (unrestricted grammars) include all formal grammars.
- Type-1 grammars (context-sensitive grammars) generate the context-sensitive languages.
- Type-2 grammars (context-free grammars) generate the context-free languages.
- Type-3 grammars (regular grammars) generate the regular languages.

A key difference between type-3 and the other types is that type-3 grammars do not require memory about the past elements of the sequence in order to be produced or processed. Thus, although all four levels involve recursion, it is important to distinguish the recursion of grammars that are not of type-3 (hereafter non-trivial recursion) from that of type-3 grammars.

Interestingly, the presence of LDDs indicates that memory of the past is needed and that this memory may need to be maintained for a long time. Can this be interpreted as LDDs implying that the complexity of the language is above type-3? Standard automata theory takes a very radical point of view on the issue (Hopcroft & Ullman 1979): any finite language is of type-3, regardless of the internal dependencies within the sequences that it can generate. Moreover, the fact that a language is infinite and shows LDDs does not imply that the language has a higher complexity than type-3 either. Theoretically, one could design artificially a finite language able to produce long but finite length "sentences" in which LDDs between the "words" of a sentence are present. In sum, LDDs are not a sufficient but a *necessary condition* for non-trivial recursion, which means that the existence of LDDs dissociated from complex recursive syntax is a possibility not to be discarded.

4. LDDs are not uniquely human

The study of LDDs using information theory and statistical physics techniques has a long tradition. These techniques operate on sequences of units (e.g., words in texts or nucleotides in genomes). For instance, Ebeling and Pöschel (1994) used transinformation (a two-point correlation function) to determine if pairs of

letters at a certain distance within a literary text are related or not. Using this and more elaborated techniques, LDDs between letters (Ebeling & Pöschel 1994) or words (Montemurro & Puri 2002) have been reported. Applying these methods to other species is providing growing evidence that LDDs are not a unique feature of human language. This kind of dependencies are found in the sequence of units of humpback whale songs (Suzuki *et al.* 2006) and the sequence of dolphin surface behavioural patterns (Ferrer i Cancho & Lusseau 2006). To illustrate this kind of research, we show a piece of a sequence of dolphin surface behavioral patterns borrowed from Ferrer i Cancho & Lusseau (2006): *TSD*, *TO*, *AS*, *FB*, *TSD*, *TSD*, *AS*, *TSD*, *AS*, *LT*. In this piece of ten patterns, there are seven different behavioral patterns: *TSD* (tail-stock dive), *TO* (tail out), *AS* (active surfacing), *FB* (fart blow), *LT* (lob tail). Metaphorically speaking, these are the "words" of dolphin surface behavioural sequences. Using transinformation as Ebeling & Pöschel (1994), Ferrer i Cancho & Lusseau (2006) studied collections of dolphins behavioral sequences (metaphorically speaking, each sequence could be seen as a "sentence" and the whole collection of sequences could be seen as a "corpus") and found that the dependency of a dolphin surface behavioral with previous patterns extends back at lead to the 7^{th} past behavioral pattern. Notice that the analysis based on transinformation cannot determine if two concrete occurrences of a pattern of a specific sequence are dependant or not. It just provides global information about the span of dependencies in the whole collection of sequences. Besides, this kind of analysis is based on the concept of mutual dependency: if a pattern depends on a second pattern, then the second pattern depends on the former pattern. In other terms, the possibility of subordination of one pattern to another dependant pattern is not considered as in the case of human words in syntactic theory.

It is important to notice that the LDDs that are found in human language do not have an exact correspondence in other species in all cases. In human language, the LDDs between syntactic constituents involve meaningful units. In contrast, the units on which LDDs are found in other species, i.e. humpback whale song units (Suzuki *et al.* 2006), dolphin surface behavioural patterns (Ferrer i Cancho & Lusseau 2006) or bird song types (Ferrer i Cancho 2007) are meaningless except for dolphins surface behavioural patterns, for which there is evidence of a rudimentary meaning (Ferrer i Cancho & Lusseau 2006). Besides, strong evidence of LDDs needs that the units are produced by a single individual. In Ferrer i Cancho and Lusseau 2006 the evidence of LDDs is limited to the sequence produced by a population of dolphins, not an individual

speaker as in human language. However, whales and birds songs exhibit LDDs at an individual level (Suzuki *et al.* 2006) and it can not be discarded that a deeper study on dolphins distinguishing the identity of the performer of a behavioural pattern could provide evidence of LDDs in the production of single individuals.

In sum, an exact correspondence between LDDs in syntactic constituents and in the sequences produced by another species has not been clearly found at the level of meaning and their maintenance within the behaviour of a single organism has not been studied in all cases. Anyhow, we think that the evidence so far reported suffices to start thinking about language LDDs as a modified version of this more rudimentary class of long-term correlations.

It could be tempting to think that LDDs are a unique feature of complex organisms (i.e. organisms with a brain). However, this hypothesis does not stand. LDDs are found, for instance, in DNA sequences (Li & Kaneko 1992). In this case, the units are nucleotides and, as far as we know, single nucleotides do not have biological meaning. Even more challenging for proponents of long-distance dependencies as a uniquely human property is the evidence of LDDs in the atmosphere (Harrison 2004) or in earthquakes (Telesca *et al.* 2007). All this means that the presence of LDDs is pervasive, not just in the specific case of language, but in nature at large. The remainder of this article is aimed at showing the implications this finding raises for the evolution of language.

5. Some conclusions: LDDs and the evolution of human language

The data referred to in the previous section seems completely at odds with the consensus view that LDDs are uniquely associated with human language and rather points to the following conclusion, that we are going to assume:

LDDs-as-FLB Thesis:

"The establishment and maintenance of LDDs falls into FLB".

The evolutionary consequences of this statement are far reaching, given the causal link also customarily assumed among linguists between the existence of LDDs and the existence of complex syntax with non-trivial recursion. As a matter of fact, the statement above seems to open two alternative scenarios concerning the evolution of the faculty of language:

Scenario 1. LDDs can be dissociated from complex recursive syntax (or, for the sake of clarity, syntax with structural embedding of hierarchically organized units). Within this scenario, the contention could be made that LDDs have acted as an evolutionary driving force pointing to the advent of phrase

structure grammars, possibly as a device for the resolution of LDDs within optimally restricted search spaces or domains.

Scenario 2. LDDs can not be dissociated from complex recursive syntax. Actually, they serve as a diagnostic cue for the existence of this kind of syntax: any system of non-human behavior exhibiting LDDs would automatically be deemed as being in possession of recursion. Within this scenario, the FLB vs. FLN distinction would cease to be operative, demonstrated the existence of some other system of this type aside from language. As a matter of fact, FL would not be other thing that a collection of capacities repeatedly evolved, with their incorporation into a single system as the only distinguishing feature of the faculty as such (Lorenzo 2006: ch. 3).

The evidence is still sparse in order to decide which of the advanced scenarios is the most promising for the evolutionary understanding of language. Payne and McVay (1971) defend that the songs of humpback whales are hierarchically structured and Suzuki, Buck and Tyack (2006) contend that they also exhibit long-term correlations of a sort not capable of being represented by a Markovian model. This findings strongly point to the second scenario, but more research is still needed to adhere to the idea. Ferrer i Cancho and Lusseau's (2006) detection of long-term correlations in the behaviors of dolphins points at a minimum to the fist scenario, but still there could be the case that those behaviors are hierarchically structured, a detail not considered in that paper. Birds, whose patterns of singing behaviour have also been defended as hierarchically organized (Todt & Hultsch 1998 and Gentner *et al*. 2006), are also a promising organism for extending the search of LDDs.

It seems not too risky to assert that a deeper understanding of the behavior of organisms like those mentioned above will shortly serve to straighten the evolutionary gulf between language and the rest of nature, a gulf perhaps not so wide as customarily thought of by most linguists.

Acknowledgements
This work was funded by a Juan de la Cierva contract from the Spanish Ministry of Education and Science under the project BFM2003-08258-C02-02 (RFC) and by the Spanish Ministry of Education and Science and FEDER under the project HUM2007-60427/FILO (VML and GL).

References

Anderson, S. R. (2004). *Dr. Dolittle's delusion: animal communication, linguistics, and the uniqueness of human language*. New Haven: Yale University Press.

Berwick, R. C. (1998). Language evolution and the Minimalist Program: the origins of syntax. In J. Hurford, M. Studdert-Kennedy, & C. Knight (Eds.), *Approaches to the evolution of language. Social and cognitive bases* (pp. 320-340). Cambridge: Cambridge University Press.

Chomsky, N. (2000). Minimalist inquiries: the framework. In R. Martin, D. Michaels & J. Uriagereka (Eds.), *Step by step. Essays on minimalist syntax in honor of Howard Lasnik* (pp. 89-155). Cambridge (MA): The MIT Press.

Ebeling, W. & Pöschel, T. (1994). Entropy and long-range correlations in literary English. *Europhysics Letters*, 26 (4), 241-246.

Ferrer i Cancho, R. & Lusseau, D. (2006). Long-term correlations in the surface behavior of dolphins. *Europhysics Letters,* 74, 1095-1101.

Fitch, W. T., Hauser, M. D., & Chomsky, N. (2005). The evolution of language: clarifications and implications. *Cognition,* 97, 179-210.

Gentner, T. Q., Fenn, K. M., Margoliash, D. & Nusbaum, H. C. (2006). Recursive syntactic pattern learning by songbirds. *Nature 440*, 1204-1207.

Harrison, R. G. (2004). Long-range correlations in measurements of the global atmospheric electric circuit. *Journal of Atmospheric and Solar-Terrestrial Physics* 66 (13-14): 1127-1133.

Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: what is it, who has it, and how did it evolve? *Science,* 298, 1569-1579.

Hopcroft, J. & Ullman, J. (1979), *Introduction to automata theory, languages and computation*. Massachusetts: Addison-Wesley.

Jackendoff, R. & Pinker, S. (2005). The nature of the language faculty and its implications for evolution of language (reply to Fitch, Hauser, and Chomsky). *Cognition*, 95, 211-225.

Li, W. & Kaneko, K. (1992). Long-range correlation and partial $1/f^{\alpha}$ spectrum in a non-coding DNA sequence. Europhysics Letters, 17, 655-660.

Lorenzo, Guillermo (2006). *El vacío sexual, la tautología natural y la promesa minimalista. Ensayos de biolingüística*. Madrid: AMachado Libros.

Montemurro M. A. & Puri, P. A. (2002). Long-range fractal correlations in literary corpora. *Fractals*, 10 (4), 451-461.

Payne, R. S., & McVay, S. (1971). Songs of humback whales. *Science*, 173, 587-597.

Pinker, S., & Jackendoff, R. (2005). The faculty of language: what's special about it? *Cognition*, 95, 201-236.

Suzuki, R., J. Buck, & P. Tyack. (2006). Information entropy of humpback whale songs. *Journal of the Acoustical Society of America,* 119, 1849-1866.

Telesca, L., Lovallo, M., Lapenna, V. & Macchiato, M. (2007). Long-range correlations in two-dimensional spatio-temporal seismic fluctuations. *Physica A*, 377 (1), 279-284.

Todt, D. & Hultsch, H. (1996). How songbirds deal with large amounts of serial information: retrieval rules suggest a hierarchical song memory. *Biological Cybernetics*, 79, 487-500.