Human language is the result of a cascade of consequences from an initial mutation which provided a new “representational” capacity to some mirror neurons. This initial mutation has high evolvability. The mutation coincidentally allowed representations of the two substances of signs to meet in human brains, thus accounting directly for signs. Recursivity is a result of the self-organization triggered by the chaotic system that emerged from this system of signs.

1. Introduction

Inquiry into the origin of language must explain why it took the form that it did. There are numerous structural properties which are attributed to language. It would be a formidable task to look at hundreds of properties in the light of the origin of language. I investigate two properties for which there is a very broad consensus among scholars: saussurean signs and type-recursion (the embedding of a phrase of type X into a phrase of type X). If we can explain why language evolved with these two basic properties, we are heading in the right direction.

2. Language as a neurological side-effect

Language seems to have evolved to fit some kind of communication or thought process, because that is how we often use language today. So the fit corresponds to what we do with language, an unsurprising panglossian conclusion. However, the human brain did not evolve for these functions. The emergence of language is a side-effect of a new kind of neuron.

There is a crucial distinction between humans and other primates which has recently been discovered and which informs us on what evolution did to make the human brain language ready. The difference is not in the brain plan (Deacon 2003): it is in the way some human mirror neurons function. Recent comparative studies on mirror neurons show that these “elementary particles” differ in an important way in the brain of humans and in the brain of other primates.
Some human mirror neurons get activated in absentia: humans do not have to see or hear an action for these neurons to react. Thus, Meltzoff et Moore (1997) observed that infants spontaneously correct their erroneous imitations of facial expressions. Moreover, infants imitate absent actions, that is, actions previously, but not currently, observed. Monkeys, on the other hand, merely replicate what is immediately perceptible. Also, Gallese (2003) argues that the simulation processes occurring during motor imagery in humans are different: they are triggered by an internal event. This indicates that the “resonance” mechanism has gained sophistication in humans compared to monkeys. Brain imaging reveals that the human brain has intra-representational systems (IRS) of neurons which can “represent” internal events, i.e., brain activations which are triggered by representations of events instead of events themselves, and which produce representations of events with no brain-external realization (Iacoboni et al. 2004; Jacob & Jeannerod 2004). This intra-representation is fundamental to Hurley’s simulationist view that the same neuronal system one uses in one’s own actions is also used off line to understand similar observed actions. She shows “how layered mechanisms of control, mirroring, and simulation can enable distinctively human cognitive capacities for imitation, deliberation, and mindreading” (Hurley 2008: 1). Crucially, these IRSs which encode non-realized goals are unique to humans (Rizzolatti 2005, Hurley 2008). This sets the stage for language, more precisely, for the core elements of language: signs.

### 3. A sign is born

IRSi s are structures and mechanisms for active perception (Hurley 2008). This is a typical instance where a small mutation brought about complex phenotypic changes. I hypothesize that these neurological systems are what makes language possible: they provide the mechanisms for the saussurean sign to emerge.

In order to have signs, there has to be a way for elements of the conceptual and perceptual substances to meet in the brain. IRSs provided our ancestors with stable and productive mental representations that enable this linking. A linguistic sign is a relation between a representation of a perceptual element (the signifiant) and a representation of a conceptual element (the signifié). The key to the emergence of signs is therefore to have a way for this relation to take place in the brain. The central question is in which way exclusive to humans the vocal phenomenon enters into the brain. The answer is provided by the recently uncovered biological properties of the human brain pertaining to the IRSs.

IRSi s explain how a vocal element could take on the function of signifiant. A vocal element activates an IRS. This IRS creates a representation of that vocal
element in the brain. Perceptive and motoric elements are already brain-internal, but as categorial systems linked with brain-external elements. IRSs crucially provide the additional possibility to process these perceptual and motoric elements off-line in the brain, detached from the external world. Once a vocal element is “mentalized” in this way, it can be subjected to meta-representational processes. In particular, the vocal element can be linked to a meaning because IRSs give these dissimilar elements a similar nature which makes the linking possible. I claim that it is this particular property of human brains that makes signs possible, and predict that tests will reveal that the storage, production and understanding of signs depend on the activation of particular IRSs.

IRSs provide a plausible and testable explanation for the emergence of the linguistic sign. The human specificity of these IRSs explains why we have signs and other animals don’t. We do not have a brain that “has” language (in the sense of innate principles and parameters of a Universal Grammar) but instead a language-ready brain with neurons unique to our species (IRS) that provide us with a qualitatively different memory. This system took on this other function of linking percepts and concepts after it was in place, as a side-effect.

Signs have very low evolvability. As Bickerton (2009) indicates, it is very unlikely that they evolved from an animal communication system, because the system would have to go through three concurrent transitions: it would have to be decoupled from situations, from current occurrence, and from the limbic system. Since syntax and semantics depend on words, and morphology and phonology “build” words, they may all be not evolvable. We face a paradox. On the one hand, the faculty of language seems to be an implausible trait for a living organism if we apply evolvability conditions directly to language as a complex system or as having certain functions for communication or organizing thought. On the other hand, language exists with all these components.

The solution to the paradox is to look for evolvability not at the functional level, but at the subpersonal level of the neurological substance. Only at that level can we determine the specifications that a biological organism has to meet to have a system with the properties that we observe in human language. IRSs have high evolvability because they are in continuity with the mirror neuron systems which are found in the brain of other primates: IRSs are off-line activations of such systems in essentially the same parts of the brain.

Arbib (2004) proposes the following evolutionary progression towards intra-representational neuronal systems. First, precursor systems for grasping evolved to provide the subtle visual feedback needed for the control of dexterous hand movements, allowing the organism to evaluate its progress towards some goal and correct its movement accordingly. The second stage was a mirror
system resting on the organism’s ability to map its body on (or from) that of another, and recognize which action that organism is conducting. Finally came a capacity for imitation, i.e., to go from recognizing a novel action by another to adding the action to its repertoire by means of a representation of it: the organism can achieve the observed goal by some coordinated control program abstracted from the observed pattern of movement. Each of these steps has high evolvability. Moreover, as Deacon (2008) observes, the size and the concomitant increase in regional differentiation of cerebral cortex are evolvability features of human brains which result in more afferent axons projecting to a greater number of regions in the cortex. This allows more intra-brain interactions, and so enhances the evolvability of IRSs which are triggered by other “brain events” instead of external events.

This continuity at the subpersonal level contrasts with the discontinuity which IRSs produce at the functional level: their new representational capacity had the side-effect that percepts and concepts could meet through their representations, and this got language started by allowing the formation of signs/words. The innovation of language greatly increased functional complexity, but (contra Deacon 2008) the transition to words did not necessitate a relaxation of selection in order to decouple ACSs from situations, current occurrence, and the limbic hard-wiring. Words are not the result of a decoupling but of an override. Vocalizations did not get decoupled from the limbic system so that they could be used symbolically. Instead, some vocalizations tied to the limbic system (and others if there were any) became represented in the IRS and it is these representations that took on the role of signifiant, leaving the limbic system generally intact.

4. The epigenetic factor: self-organization due to building materials

Once signs emerge, they quickly proliferate and self-organization kicks in, deriving the apparent complexity of language. As in the case of other biological systems, IRSs are complemented by epigenetic self-organizing constraints which emerge from interactions of properties of building materials which limit adaptive scope and channel evolutionary patterns (Jacob 1982). Since the representation of any percept can potentially be linked to the representation of any concept, there are innumerable possibilities. Therefore, IRSs introduce a new chaotic system in the brain. As in other chaotic situations, chance meetings are progressively amplified by material properties and result in clusterings, in order out of chaos. For instance, out of the chaotic passages of termites,
concentration points are created by chance droppings: these points become more important as they attract more passages, and the outcome is termite mounds with complex structures (Prigogine & Stengers 1984).

In language, the potential chaotic dispersions of arbitrary signs are constrained by material constraints which restrict the linguistic sign system in a way that maximizes contrastive dispersion.

4.1. Contrastive dispersion of percepts and combinatorial phonology

Vocal percepts cluster in a few particular “hot spots” among the innumerable, chaotic possibilities we can produce and perceive. As in other chaotic systems, the clusterings depend on frequency and accumulation, which result from ease of production and distinctness of perception. Thus certain vocalizations are easier to pronounce and require less energy: this is likely to favor their use and increase their frequency (Lindblom 1992). Also, the human perceptual systems set upper bounds on the distinctions which we can perceive or produce as signifants. Nowak et al. (1999) found that the demands of discriminability (as well as memory and time to learn) constrain the system to a fairly small set of signals. Oudeyer (2005) shows in simulations that canalisation by the vocal tract and general acoustic theory define eight discrete regions of a tube like the vocal tract which afford greatest acoustic contrast for least articulatory effort, and these correspond to places of articulation in natural languages.

The small number of clusters “automatically brings it about that targets are systematically re-used to build the complex sounds that agents produce: their vocalizations are now compositional” (Oudeyer 2005: 444). How could that be? Where does the compositional percept come from? The answer is again in the material properties of the elements. Vocalizations occur in time. So our perceptual system captures the linear properties of vocalizations when they are produced, in particular the linear relationship between two vocal clusters, the most salient one being linear adjacency. The linear adjacency of two vocal percepts is therefore a percept, and it can be represented by an IRS, just like any percept. The relational percept of juxtaposition is already in the stock of our perceptual system, hence it is available for IRSs that link concepts and percepts. Another material property of vocalizations is intonation: therefore, another perceptual element represented by IRSs is the tone superimposed on a phoneme, of which there are only a few distinctive values due to contrastive dispersion.

Crucially, in an arbitrary system, the percept represented by an IRS and linked to a concept can be any element among those recognized by the perceptual system: a phoneme, a juxtaposition, or an intonation. Under
arbitrariness, it makes no difference whether the represented element is simple or complex. The acoustic image can be a single phoneme, or the relational percept of juxtaposition applying any number of times to phonemes, as well as any of the available intonations on these elements. These complex elements remain within the limits of what humans can distinctively perceive or produce because their parts have the appropriate qualities. Phonological compositionality therefore comes from two properties of vocalizations: they are temporally linearized, which provides the percept of juxtaposition as a significant; additionally, the material property of intonation provides tone as a significant.

As a result of all these constraints, a limited number of discrete phonemes and their combinations emerged from self-organization in the chaotic system of the speech sounds represented by IRSs: they constitute the potential signifiants.

4.2. Contrastive dispersion of meanings

Material properties and self-organization also affect the conceptual interface of the linguistic system. Here too order arises out of chaos and clusters are formed in the mass of the conceptual substance as a result of frequency and accumulation. The more a situation has some importance and/or is encountered frequently by the organism, the more likely concepts associated with it will be activated. The accumulations self-organize on the concepts most used by the organism. It is this “usefulness” which makes the signifiés tend to correspond to fairly broad and/or usual categories of things, actions, qualities, etc. (Locke 1690:15), basic-level concepts (Rosch et al. 1976, Murphy & Lassaline 1997) whose names are among the first common nouns learned by children (Brown 1958: 20). Discrete meanings are formed in the mass of the conceptual substance (Saussure 1916) as a result of maximizing contrastive dispersion across the space for signifiés. These clusters have to be relatively few in number due to storage limitations of the brain, hence they tend to have fairly broad meanings. This does not adversely affect the communicative or thinking functions of language because linguistic signs reside in organisms that independently have an inferential system that supplies the required complementary information.

4.3. Self-organization and syntax

I hypothesize that combinatorial syntax also emerged out of chaos due to self-organization. The element of the conceptual system which IRSs represent the most frequently is the relation of predication, since it is common to all the attributions of properties. Predication is the broad meaning *par excellence*, it is
the meaning which creates the strongest concentration point in the chaos of IRS representations of meanings. Moreover, on the perceptual side, the most frequent elements are temporal sequencing and intonation. These happen to be relational percepts. So the hottest accumulation point in the mass of the conceptual substance is the relation of predication, and the hottest accumulation points in the mass of the perceptual substance are the two relational percepts juxtaposition and intonation. These accumulation points in their respective domains are so overwhelmingly dominant that the probability of links between predication and juxtaposition/intonation reaches a point where these links inevitably accumulate and crystallize. Therefore, given these prior properties, I predict that computational simulations will show that human languages inescapably develop combinatorial signs involving predication as a signifié, and juxtaposition and intonation as signifiants. Thus compositional syntax is also a consequence of self-organization arising out of the chaos created by IRSs.

Combining signs by means of a combinatorial sign produces a complex sign, i.e. a phrase with a compositional meaning and a compositional signifiant. Assuming that uni-signs fall into one of the lexical categories/types like noun, verb, adjective, preposition, and also that the category of a complex sign is the category of one of its immediate constituents (this endocentricity following from the parsimonious assumption that the only primitives are lexical and combinatorial signs), type-recursion occurs when a restraining sign or one of its elements happens to be of the same type as the restrained sign whose category projects and determines the category of the complex sign. Therefore, I predict that, given these additional prior properties, computational simulations will also show that type-recursivity is a side-effect of the combinatorial capacity of “syntactic” signs. In short, the presence of IRSs in human brains explains both saussurean signs and type-recursion. They no longer pose insoluble challenges for natural selection.

References


