Aproximándose al “Link perdido”: pointing y lenguaje como intencionalidad compartida.

Ernesto E. Guerra1 and Allison F. Smith1

1. EMCL - Università degli Studio di Milano-Bicocca, Milán, Italia.

Resumen

Los humanos se comunican a través de un sistema distinto a otros seres vivos. Este artículo tiene como objetivo responder dos preguntas fundamentales sobre este sistema, el lenguaje humano; 1) ¿qué es lo que hace al lenguaje humano tan diferente de otras formas de comunicación? y 2) ¿qué nos ha llevado a desarrollar un sistema tan particular? Usando evidencia teórica y empírica, mostraremos como el lenguaje humano es único, ya que es una parte intrínseca de las habilidades cognitivas, y ya que la motivación comunicativa es fundamentalmente distinta a otras formas de comunicación entre seres vivos. Este sistema particular se ha desarrollado como el resultado de la aparición de habilidades cognitivo-sociales y de motivaciones sociales de comunicación, ambas manifiestas originalmente a través del acto de pointing (señalar), el cual aceptamos como un precursor del sistema de lenguaje humano más complejo. Esta proposición encuentra soporte tanto filo - como ontogenéticamente. Sobre el final, presentamos una hipótesis comprobable, sobre la evolución del lenguaje, basada en la intencionalidad compartida la cual de ser acertada daría aún mayor soporte a las ideas planteadas en este artículo.

Palabras Clave: Comunicación Humana, Pointing, Intencionalidad Compartida.

Abstract

Humans communicate in a system unlike any other living things. This article aims to answer two fundamental questions about this system, human language; 1) what is it that makes human language so different from other forms of communication? and 2) what has led us to develop such a particular system? Using theoretical and empirical evidence, we show that human language is unique in that it is intrinsically part of cognitive skills and motivation for communication, unlike other forms of communication between living things. This particular system has been developed as a result of the appearance of social cognitive skills and social motivations for communication, both manifested originally through the act of pointing, which we accept as a precursor to the more complex system of human language. This proposal is supported both phylo - and ontogenetically. At the end, we present a testable hypothesis for the evolution of language on the basis of shared intentionality, that given positive evidence for this hypothesis would further support the ideas presented in this article.

Keywords: Human Communication, Pointing, Shared Intentionality.

Contacto E. Guerra: eguerra@integracion-psi.cl
Introduction

In everyday life, human beings (most of us) put in use one exclusive, compared to other species, and complex way of communicating, in such a natural and effortless manner, that we are not even aware of it or its properties. It is very common and necessary for humans to communicate with other humans through language; in fact it is actually hard to think about another way to do it (Tomasello, 2008). This process helps us to coordinate with each other and to use available information from the most simple interactions, like the utterance – Can I use your pen, please?, to very complex tasks, as, for instance, performing an emergency spinal cord surgery. Both complex and simple linguistic exchanges are mediated by the same rules and features.

Even though it seems to be easy and natural for us to communicate, we do it in very particular manner highly different from other animals, even of those which follow us in the evolutive scale. It is true that every living organism has a particular way to communicate with conspecifics, but what is it that makes human language so different? What has led us to develop such a particular system? These two questions will be the directives of the present review, and we will attempt to give a response to both.

Human language is one of the most interesting objects of study for many scientific fields like theoretical linguistics, cognitive psychology, neuropsychology, neurolinguistics and psycholinguistics just to mention a few. The way to access its nature depends on the conceptual ground from which one starts. Nowadays, the question about the nature of human language is one central issue for the cognitive sciences, as Michael Tomasello states “the most useful descriptions for developmental researchers [in the study of language development and evolution] come from Functional and Cognitive Linguistics”1. Cognitive Linguistics can be described as “an approach to the analysis of natural language that focuses on language as an instrument for organizing, processing, and conveying information”2.

The following question will be – why is language so important for sciences? There are two main issues that remain unknown in the study of human natural language, one lesser and one greater. The lesser is the question about the origin and evolution, and the greater is the nature of subjective experience (Feldman, 2006). The comprehension of the first issue is strongly relevant for the better understanding of many phenomena that remind mysteries (like the second issue) and are top priority in science’s schedule. One good example of this is the question for the Mind, which has been described as the current most challenging issue for sciences (Damasio, 1999, Kandel, 2005).

In words of Feldman (2006) “there is no such thing as language in isolation from thought”3, which expresses the first assumption of Cognitive Linguistics: which is that language is not an autonomous cognitive faculty, it is rather part of the general intelligence, thus is connected very intimately with other mental processes. The second assumption is that our knowledge of language emerges from language use (Gries, 2006). Based in this second statement, a general theoretical account to the study of language is call “usage-based approach” to emphasize the assumption common to all functional and cognitive approaches that, in this case linguistic structure, emerges from use both phylogenetically and ontogenetically (Tomasello, 2007).

In this essay, we will ascribe to this approach’s perspective, including its two main assumptions, because we believe it is a reliable way to describe and explain the human language system and its evolution.

Origins of Human Communication

The inquiry for the origin of human language as a “uniquely co-operative form of communication is still something of a

1 Tomasello (2007), page 1093.
Guerra, E. Smith, A.

mystery"⁴, but to begin to understand it, it is necessary to find a starting point of analysis. On this respect, one very interesting proposition is made by Michael Tomasello who argues that in order to understand the nature of human language and its evolution, “[…]we must first understand how humans communicate with one another using natural gestures”⁵, and we need to do so because a previous communicative system is needed to develop and learn a second one. Following Wittgenstein (1953) any linguistic system has a non-linguistic infrastructure on its base, this non-linguistic system is a compound of intentional understanding and common conceptual ground and it is by definition primary.

Tomasello explains that most of the researchers that have studied human gestures, have focused either on sign language or gestures that accompany verbal language, and he highlights why these ways of gesturing are not really informative about language origins. The sign language in deaf people contains all the characteristics of modern languages, and the gestures that accompany verbal language have its own qualities serving a supportive role in the communicative process, therefore neither of the two can be a consistent non-linguistic precursor of modern human communicative systems. But, on the other hand, there are two types of gestures, described in terms of how they are used to communicate, that appear as good candidates, they are **Pointing** and **Pantomimic**. In this essay we will mainly discuss the role that pointing plays on the development and evolution of human language, because, assuming that it is the successor of attention-getter apes’ gestures, and being these the only behavior in apes that display a triadic form of communication involving **social and referential**⁶ intention, it “may be considered the closest thing we have to a “missing link” between nonhuman primate communication”⁷.

---

⁴ Tomasello et al. (2007), page 720.
⁵ Tomasello (2008), page 2.
⁶ Referential in the case of apes have not the same meaning that human referential communication. In humans this referential indication, induce social intention inferring which is what the communicator wants the recipient to do, know or feel.
⁷ Tomasello (2008), page 29

Pointing as the first Co-operative Act of Communication

It is widely agreed that **pointing**, just like verbal language, is a complete collaborative, socially shared, communicative act itself and this is because it involves an intention about the communication specifically (Grice, 1957; Tomasello, 2004). They are adjustable and flexible, depending on the recipient’s response of comprehension or incomprehension, being the agent able to change the communicative act to one that is better understood by the recipient (Clark, 1996).

One special functional feature of pointing is that it serves to direct someone’s attention to something, and even this having some variation on its form, it is likely a human universal (Kita, 2003; Butterworth, 2003; Tomasello et al, 2007). But this basic function of capturing the attention of a conspecific is also observed in apes, even though they use other type of gestures like ground-slapping or throwing-things. So what is it that makes human pointing’s message special?

In a general way, we can say that the two types of human communicative gestures, pointing and pantomimic, work in parallel to the two types of great ape gesture, attention-getters and intention-movements respectively, in the sense that all of them are nonconventional, action-based, naturally meaningful and cooperative (Tomasello, 2008). Nevertheless, there are very significant differences. In the particular case of pointing gestures, that are similar with the attention-getter acts in great apes in function; they are both oriented to direct attention to something in the immediate context, based in the natural tendency of primates to follow gaze. But following gaze is not the same thing as understanding communicative intentions (Tomasello, 2004). There is evidence that shows that the fundamental differences between the way in which humans and apes direct the attention of someone else to a specific target in the environment are strongly connected with *shared intentionality*. In this sense, humans have the necessary skills and motives to generate *joint goals* and *joint attention*, and what provides the common conceptual ground is the basic cognitive skill of *recursive mindreading* (we will specify
through the paper what we intend by these concepts, but so far we needed to introduce them); something that apes do not share with us.

In fact, there are many human language evolutionary precursors that we do share with apes. Both humans and apes exhibit communicative motives, intentionality in communication and, of course, communication devices. But there is an evolutionary-behavioral distinction in their manifestations; while apes only request as motive, e.g. throwing a piece of wood to catch others’ attention, humans also help and share with one another. Furthermore, from these two motives humans generate culturally arranged norms of communication, being examples of all humans institutions, from schools to banks, but also merely any specific language, like English or Spanish are culturally arranged norms of communication. In terms of intentionality, apes seem to understand goals and perception. This is based in experiments made with chimpanzee in non-natural environment and in interaction with humans, e.g. chimpanzees act frustrated when a human fail in passing food for no good reason (like if she is not willing to share the food), but when human fail for good reasons (if she is not able or had an accident along the way) apes wait patiently showing understanding of others’ goals. At the same level humans share goals and communicative intentions and show joint attention and common ground. On regard with the communicative devices, there is no actual evidence for any apes’ vocalizations to be referential and/or flexible, (Cheney & Wrangham, 1987; Crockford & Boesch, 2003; Pika & Liebal, 2006) thus their connection with human language is loose (to be generous). Therefore, apes’ vocalizations are not good “precursor candidates” and we have to look to other communicative devices. Gestures are good candidates, since they appear to be ritualized signals, ontogenetically learned and flexible used communicational devices both in human and non-human primates. But human gestures go further in complexity through the ability to imitate, due to the ability of share intentions, which leads us to develop communicative “arbitrary” conventions: verbal language.8

Pointing itself does not communicate a message; it just (as mentioned above) directs someone’s attention to a target. Thus, the true message is given by the fact that the communicator and the recipient both know that there is something relevant on that specific target. This is what Clark (1996) has called common ground or, sometimes joint attentional frame, when the shared perceptual context is emphasized. Therefore, the following question arises – Where is this common conceptual ground coming from?

Following Tomasello (2008), the answer of the latter question is naturally “context”, but specifically human context, a communicative context, which does not just mean everything (in the physical-world sense) that surrounds us. Instead, this human communicative context is more connected with “what is relevant to the social interaction, that is, what each participant sees as relevant and knows that the other sees as relevant as well—and knows that the other knows this as well, and so on, potentially ad infinitum”9. Consequently, this human ability to create common ground, which involves joint attention, shared experiences, cultural common knowledge and so on, is a totally relevant dimension of all human communication.

Hence, when we participate in human everyday-life interactions, and someone points us something, then we immediately (more or less unconsciously) ask ourselves - what is she directing my attention to? (which addresses the referential intention) and why is she directing me to it? (which addresses the social intention). There is a close relationship between this two questions or intentions and this relation is very well expressed by Tomasello (2008) in following paragraph:

8 In the conclusions, we will extend on the argument about the notion of human language as a device that has necessarily emerged from a co-evolutionary process, between cognitive skills, social motives and communicational gestures.

9 Tomasello (2008), page 74.
“Given my signal of a communicative intention, I draw your attention to some referential situation in the external world—my referential intention—which is designed (along with some expression of motive) to lead you to infer my social intention via processes of cooperative reasoning, since you are naturally motivated to find out why I want to communicate with you (based on mutual assumptions or norms of cooperation)” (Chapter 3, page 97).

From the last paragraph, in addition to these two types of intention, Tomasello introduces the concept of motives, which all human communicators have. There are three basic human motives, and they emerge early in ontogeny, thus they have plausible evolutionary roots in more general human social interaction; also this paragraph introduces the concept of mutual assumptions.

In regard to motives, it is easily observable that in everyday-life we communicate with others in order to make the others do what we want them to do, whether asking for help or to perform a specific action. If we need someone to meet us somewhere, most often we would explicitly say, for example – Let’s meet under the clock on the Main Square. Through this imperative utterance we make others do what we want them or we need them to do for us. Another way to make others do things for us is to be less explicit, is to use and informative utterance, even though the motive is the same; for example, one can say – here in Salamanca have the custom to meet under the clock on the Main Square, assuming that the counterpart would understand this as a suggestion of a place where to meet.

We often communicate by offering help to others as well, either giving valuable or just interesting information in order to let the others know something; this happens even if we have not been requested to do so. For example, a businessman at train station asks to the ticket seller – On what platform is the next train to Rome? The ticket seller, answers him – Platform 6, but it stops at every station. The 1:45 will get you there sooner. The businessman could answer – Thank you. One ticket on the faster train; or – Thank you very much, and then proceeds to platform 6 (adapted from Caplan, 1999). In this example we can see that the ticket seller gives information to the businessman that was not actually requested for him, but which the ticket seller believes could be helpful for him in the case that he wanted to get as soon as possible to Rome. Following Grice’s conversational maxims (Grice, 1975), this is due to the relevance maxim, which is a presupposition of relevance about the utterances that both interlocutors have, given the case the ticket seller presupposed that when he is inquire for the “next train”, the underlying motive is “I need to get as soon as possible to...”

Finally, we communicate with others with the simple aim of sharing feelings or attitudes towards specific things or situations as well, and by doing this we can make the recipients feel an attitude or emotion, therefore sharing it. This communicative exchange is not based in any imperative or informative motives as the two first basic motives mentioned above. Nevertheless, it is base of the generation and expansion of human social common ground. Typically we talk about the weather, or we comment about the sports that we like, or about the school or work and so forth, with the aim of share our opinions and attitudes about those things, and consequently making others feel in a similar emotion.

Synthesizing the three human motives described above, an association can be made between a do and a requesting motive, between know and informing motive, and finally between feel and sharing motive. Mutual assumptions, in the other hand, appear when the cognitive skills underlying mindreading, are applied to the motives, so “we both know together that we are (and should be, from the point of view of the social group) cooperative”

As can be understood from the human mutual assumptions, motives and intentions for communication as described above, in addition with the common conceptual ground, human communication is particularly co-operative (Grice, 1957, 1975; Clark, 1996; Searle, 1999; Tomasello, 2008). And within this co-operative “framework”, we can...
identify *intentional communication* which is understood by both actors in interaction, since the recipient knows that the communicator is trying to communicate, and consequently will pay attention to the communicator’s referential act and try to infer the social intention underlying this act.

So far we have described, mainly based on the research made by Tomasello and collaborators, many of the characteristics of contemporary language using the “simple” act of pointing as starting point. “Pointing may thus represent a key transition, both phylogenetically and ontogenetically, from nonlinguistic to linguistic forms of human communication”11.

**Ontogeny and Phylogeny of Pointing**

Humans have the physical capacity to point from very early stages in development; in fact it has been observed that the pointing gesture comes naturally in 3-month-old babies (Hannan & Fogel, 1987). They have motives to engage with one another in co-operative communicational processes as well, but they do not use pointing until the age of one year old, instead, they use other means (like crying for instance). So the following question would be – What happens at one year of development that makes a difference? In fact, there is empirical evidence that shows the emergence of abilities as the use of common ground and cooperative motives on the infant’s pointing at this stage (Golinkoff, 1986; Liszkowski, 2005; Tomasello et al., 2007), because the meaningful act of pointing depends on those skills and motives. At the same time, pantomimic appears “on the heels of [infant’s] first pointing, requiring a communicative intention to be effective”12.

So far, no one really knows where exactly pointing comes from ontogeny (Tomasello, 2008). Even so, it would be unlikely to think that collective intentionality suddenly came out into the human behavioral repertory with all of its complexity. It makes much more sense to think that there was a point in human evolution when some individuals began to understand a simple-recursively relation as “she sees me seeing it” and consequently later in time became a fully recursively understanding. But the natural question is – How? A very interesting proposition is made by Feldman (2006), even though he does not explicitly say that this is “the how”, his theoretical proposal can be taken if one likes as complementary with theoretical and empirical facts coming from Tomasello and colleagues’ work. According with the neural theory of language development by Feldman, “simulation might well be a cornerstone in the evolution of human language and thought”13, and this simulation can be seen as a parallel with the recursive understanding. There are two, at least, involuntary simulation behaviors that mammal exhibit: *dreams* and *play*; first very important for memory consolidation and second relevant for the physical and social development in all kind of species from this class, thus both adaptations. Taking all of this into account, one evolutionary adaptation is needed to reach the particular way in which human being voluntarily control the *simulation* (we can call it imagination). Therefore, *image* that the brain evolution on hominids, and the consequent complexity in brain circuits, would allow some of these subjects to explicitly control what was being imagined. Those individuals would be able to detached simulation from its primary involuntariness could have access, by simulation, to past, to future and eventually to other minds. At this point we got back to the recursive mindreading, the understanding of other minds, and as we already have discussed this is the ground of a richer communication system.

Tomasello proposes, based in empirical and theoretical evidence, the following evolutionary path for human language. To understand this path, an ascription to the evolutionary perspective is needed, which is to say that we assume the Haeckel’s idea about the relationship between phylogeny and ontogeny: ontogeny recapitulate phylogeny. To adhere to this principle means that we assume that evolutionary novelty is added onto old features that already exist. This new evolutionary structures are added by

---

11 Tomasello et al. (2007), page 721.
12 Tomasello (2008), page 323.
terminal addition (Richardson and Keuck 2002), which applies in the case of language evolution from this perspective.

Having as a starting point modern-day Apes we can say that they do have many components of human cooperative communication (see above for details). But it is doubtful to say that they have skills or motives of shared intentionality, thus their communication is not fully intentional and referential and consequently even though they engage in group activities using goal/intention and perception understanding and practical reasoning, they do not get involved in collaborative activities that imply joint goals/intentions, joint attention/common ground and recursive mindreading. These mutualistic collaborative activities, that will be reserved in this evolutionary pathway to creatures that Tomasello (2008)14 just calls Homo, were only possible by the fact that these creatures have a new communicational motive, and began to request help and in addition they began to answer supplying help. In a collaborative context, this kind of behavior is expected on the understanding that – helping another helps me too.

Nevertheless, the most important variation for them was the appearance of recursive intention-reading. This crucial component is responsible for the emergence of joint-goals, thus joint-attentional frames which are going to be the base of common ground. With the development of these cognitive skills, and their conjunction, the foundation of pointing as a fully communicative act is settled. At this point, we can give another step in the evolutionary chain to the Earlier Sapiens, who have consolidated the mentioned skills and develop mutual expectations of cooperation and communicative intentions as part of their everyday-life interactions. In addition, we can extend the logic behind mutualism to an indirect reciprocity, which means that if helping others helps me too, then a good helper would be more sought to be part of a group; subsequently there is another communicative motive that comes out, which is informing. Mutual cooperative expectation and communicative intentions were only possible by the combination of intentional-reading skill, and these two motives, requesting and informing (Tomasello, 2008; Tomasello et al, 2007; Tomasello, 2004).

The consecutive step in this path came from the importance gained by the group acceptance. This means that those individuals that show themselves as better helpers, by means of offering help in a regular base, were more likely to be selected as members of a group. In other words, those individuals had developed yet another motive: sharing. We have arrived then to the end of the evolutionary pathway with the Later Sapiens. At this stage, the combination of the last motive and mutual expectations produced, as an emergent property, rules that are present in many human activities. As it was mentioned before, pointing does not appear in the very beginning of the ontogenetical path but only after the child’s first birthday. As we can observe this is completely coherent with the phylogenetical path in which this modality of communication does not appear until skills like joint attention and common ground have appeared as well. Also, we can observe that pantomimic appears later in ontogeny, so it does in phylogeny. Thus pointing as a characteristic of Homo, in union with pantomimic as a characteristic of Early Sapiens, would be the precursors of a characteristic of Later Sapiens, the conventional human language.

The way and the reasons for those gestures to become a vocal-based-modality language are very complex, and therefore difficult to explain. However, we can say that this switch should have happened using these naturally action-based gestures as means of transition, because communicative conventions cannot be founded without other already meaningful acts of communication (Wittgenstein, 1953). So, in synthesis, the evolutionary path had to involve, first collaborative activities, second some source of action-based, perhaps natural, cooperative communication that act as a “bridge-to-the-shift”, and then finally the “arbitrary” conventional communication.

14 The names to these pre-humans creatures were given arbitrarily by Michael Tomasello in his book Origin of Human Communication, Cambridge: MIT Press, 2008.
Conclusions

In the introduction two questions were stated; 1) what is it that makes human language so different from other forms of communication? and 2) what has led us to develop such a particular system?. Throughout this essay, mostly theoretical evidence (although it is based in empirical evidence)\(^{15}\) has been shown in the attempt to answer these two questions.

To answer 1) we have shown how human communication is fundamentally embedded in cognitive skills (as social cognition; joint goal and intentions, recursive mindreading and common conceptual ground) and motivations for communication (such as social motivations; basically helping and sharing, and furthermore requesting and informing as well), that in concurrence are the base for shared intentionality. These characteristics are what make the human cooperative conventional communication so unique.

To answer 2) we have shown a proposal for an evolutionary path based on the ontogeny and phylogeny of language, both in which the materialization of pointing, and its meaningful properties, are embedded in the emergence of the social cognitive skills and social motivations for communication mentioned on the argument for 1), because they are the base for the collaborative nature of human activities.

In addition, as it can be inferred, those truly collaborative activities are the “original home” of human-kind communication, and the only way in which that this could have occurred (from gesture to conventional) is by means of a co-evolutionary process “by which basic cognitive skills evolve phylogenetically, enabling the creation of cultural products historically, which then provide developing children with the biological and cultural tools they need to develop ontogenetically”\(^{16}\) a specific conventional and arbitrary verbal language.

At the end of this essay, we would like to present a testable hypothesis for the study of language evolution based in shared intentionality. If we assume that a) ontogeny replicates phylogeny, plus the fact that b) pointing and pantomimes appear with coherent disparity both in ontogeny and phylogeny, and c) attention-getters and intention-movements are precursor of pointing and pantomiming respectively, it would be plausible to state that: in infants younger than 9 months, we should observe a modality of attention-getters and a modality of intention-movements, with all the components observed in primates (understand goals/intentions, understand perception, practical reasoning) and without the components (joint goals/intentions, joint attention/common ground, recursive mindreading) that change relations qualitatively from non-cooperative group activities to mutualistic interactions.

Experimental research is needed, but according to the presented arguments this hypothesis is likely to be true, and to find these pre-human gestures in very early stages of human development can be taken as evidence supporting the tentative path that has been schematically developed here.

Acknowledgement

We would like to thank Prof. Laura D’Odorico for her comments on the preliminary version of this paper and Prof. Claudio Luzzatti for all his work for EMCL. Also, a special thanks to Camila Brahm for providing the environment that made this article possible.

References


---

\(^{15}\) For a detailed review of the empirical background see Tomasello et al., 2007.

\(^{16}\) Tomasello (2008), page 345.


