

Embodied Pragmatics and the Evolution of Language

Erica Cosentino[†]
ericacosentino@libero.it

ABSTRACT

In the evolutionary theory, a central tenet is that complex forms evolved from simpler ones, according to a bottom-up process. When it comes to the evolution of language, however, a bottom-up approach is problematic. In this case, such an approach often assumes that minimal units that are inflexibly associated to their meaning come first, where the wider discourse is only a later product. In the present paper, I shall argue that we need to assume a top-down perspective on language evolution, which claims that the wider discourse is the evolutionary starting point rather than the final achievement. This approach involves the necessity to focus on the pragmatic abilities of our ancestors and on the biological mechanisms underlying them. Combining a top-down model of language evolution with an embodied account of cognition, I shall argue that a basic mechanism of *affordance* perception supports core pragmatic processes by enabling the individual to determine not only her own action possibilities in the physical environment but also the action possibilities of others and, thereby, enabling her to determine other people's intentions. As a result, I shall introduce the notion of an embodied pragmatics as a key to account for the evolution of language.

Keywords: embodied theory, affordances, language evolution, pragmatics, discourse.

1. Introduction

The present paper is based on a double methodological claim. First, I assume that one's own theory of language has to be constrained by evolutionary considerations. It follows from this assumption that a model of language which is not consistent with the theory of evolution is not a good model and has to be

[†] University of Calabria, Rende, Italy.

abandoned (“your theory of language depends on your theory of language evolution”; cf. Ferretti, 2010). Second, I maintain that also the opposite statement is true, namely one’s own theory of language evolution has to be constrained by an empirically informed theory of language, given that the type of questions for a theory of language evolution to answer depends on what we think language is (“your theory of language evolution depends on your theory of language”; cf. Jackendoff, 2010). The structure of the paper follows from these methodological considerations.

In the first paragraph, I shall introduce two alternative models of language which, on the basis of different predictions concerning the role of discourse-level contextual information in semantic processing, characterize the interpretative process as a two-step or a single-step procedure. I shall discuss the results of an experimental investigation that supports a single-step model, showing that the context of discourse has pervasive top-down effects on the processing of individual words and sentences. In the second paragraph, I shall address the question of how the empirical investigation of modern language can shed light on the issue of language evolution. I shall argue that a two-step model of language processing is incompatible with evolution and that evolutionary considerations lead to assume a top-down perspective according to which pragmatic processes enabling inferential communication were necessary to the evolution of language. In the third and last paragraph, I shall show that a top-down model of language evolution is consistent with an embodied – bottom-up – account of cognition. Focusing on the mechanism of *affordance perception*, I shall finally argue in favor of the notion of an *embodied pragmatics* as a key to account for language evolution.

2. Two models of language

Since the Chomskian revolution in linguistics (Chomsky, 1957), the study of human language in cognitive sciences has been influenced by the idea that the sentence is not only the core unit of syntactic analysis but also the core unit of language interpretation. This notion is often been conjoined with the minimalist semantic claim that sentences encode meaning by means of a context-free rule-based combination of lexical-semantic features of the words within a sentence (Grice, 1975; Borg, 2004; Cappelen & Lepore, 2005).

According to the minimalist or literalist, the process of semantic composition, which consists in putting together the semantic values of the parts to establish the semantic value of the whole, determines what is explicitly said by the speaker or “sentence’s meaning”. This step of the comprehension process is considered necessary, and it corresponds to the level of sentences’ truth conditions. The role of the context at this level is limited to cases of indexicality and anaphors resolution and should be traceable to syntactic elements in the logical form of sentences. That is, pragmatic processes of saturation and disambiguation suffice to determine the literal interpretation of the sentence. All other pragmatic processes involved in the interpretation of the utterance, which contribute to determine “speaker’s meaning”, are secondary, and presuppose the identification of the meaning literally expressed.

From this point of view, then, linguistic interpretation is construed as a two-step procedure. First, the literal interpretation of the sentence is computed by combining fixed word meanings in ways specified by the syntax, and second, information from prior discourse, world knowledge and other sources of extra-linguistic information are used to integrate sentence meaning. According to this perspective, language processing proceeds in a bottom-up fashion, incorporating contextual information only after establishing phrase or sentence local meaning.

The two-step model is challenged by the idea of free pragmatic enrichment according to which contextual information can be immediately incorporated into the truth-evaluable sentence meaning such that global context and lexical content contribute to sentence meaning at once, leading to a single-step model. Contextual information may be then used in a top-down fashion, such that the local contribution of individual words or sentences is a function of the construction of a situational interpretation at the global meaning level (Cosentino et al., 2013).

In order to adjudicate between these two models, in a recent experimental study (Cosentino et al., 2014) we tested their predictions regarding the semantic integration of a sentence in a discourse. A crucial aspect of a two-step perspective on interpretation is that local semantic constraints have precedence over global contextual factors and, consequently, local semantics cannot initially be overruled by the wider context. Single-step models on the contrary, assume that there is no such priority of local constraints over contextual factors such that, in principle, the wider context of discourse has an

immediate effect on the interpretation of the unfolding linguistic information. More to the point, we used the electroencephalography (EEG) to record event-related brain potentials (ERPs) while people read short stories. We focused on the N400 component, a negative ERP deflection peaking around 400 ms after stimulus onset and larger over centro-parietal electrodes. The N400 has become particularly relevant in language studies given its close relation to the processing of word meanings in context (see for a review Kutas & Federmeier, 2011).

In our experiment, subjects were presented with stories in which a human character selected an object to accomplish a specific goal. The combination between object and action could be either appropriate or not with respect to (1) the local telic component of the meaning of nouns (i.e., the function or purpose of an object coded in the lexical entry), and (2) the context-driven affordances of the object (i.e., its action possibilities). Two-step models predict that local semantics cannot initially be overruled by the wider context so that a telic violation such as “She uses the funnel to *hang* her coat” will always be considered inappropriate, regardless of the wider discourse. In our study we found that when this anomalous combination is preceded by a conventional context, it elicits indeed a typical N400 effect, indicating that the subject is experiencing interpretative problems. However, if the very same combination is preceded by a modified/unconventional context, the latter has the power to neutralize the violation, as reflected by the reduction of the N400. Additionally, an unconventional context can also make a locally acceptable combination (e.g. using a funnel to *pour* water into a container) globally incongruent with reference to the context-driven affordances of the object. These findings challenge the two-step model, showing that contextual information has top-down effects on linguistic processing as predicted by the single-step model. The latter, then, can better account for our results.

This conclusion is consistent with recent work in the field of lexical pragmatics (Wilson & Carston, 2007), which emphasizes that the distinction between semantics and pragmatics can be applied also at the level of individual words or phrases rather than whole sentences. In such a view, the meanings of words are often pragmatically adjusted and fine-tuned in context, in accordance with speakers’ needs and gleaned opportunistically to what they know about the world, their interlocutors, and previous discourse. Understanding a word in context may involve the construction of an ‘ad hoc’ concept or occasion-specific sense, which is based on encoded concepts,

contextual information and pragmatic expectations (Barsalou, 1987, 1993; Wilson & Carston, 2007; see also Recanati, 2004; Glucksberg, 2001). The results of the construction of an occasion-specific sense is the modulation of the lexically encoded meaning via narrowing or broadening, that is, the linguistically-specified meaning can be made either more specific (e.g. *drink* used to mean ‘alcoholic drink’) or more general (e.g. *square* used to mean ‘squarish’).

Importantly, most current approaches to lexical pragmatics maintain that the occasion-specific senses created by the pragmatic interpretation of individual words and phrases are components of the proposition explicitly expressed by the speaker. According to this contextualist thesis, there is no level of semantic content that is independent of pragmatic processes (Recanati, 2004; Carston, 2002; Wilson & Carston, 2007; Wilson & Sperber, 2004). Sentences’ truth conditions are pragmatically constructed: the output of syntactic processes is subpropositional, and it provides only a template or schema for building propositions by means of pragmatic processes. These pragmatic processes are not restricted to saturation, which is mandatory and bottom-up (in order to understand what is said in “She is John’s sister”, it is necessary, at least, to determine to whom the speaker refers by the pronoun “she”), but include also top-down processes driven by the context such as narrowing, broadening and “free enrichment” (“She took out the key and opened the door” can be freely enriched, or specified, by “She opened the door with the key”). The latter processes are not triggered by any linguistic property of the utterance and are entirely pragmatically motivated. Pragmatics then is not only concerned with what is implicitly meant by the speaker, but it also heavily shapes what the speaker explicitly says. Whereas two-step models do not allow for the immediate integration of contextual information, single-step models are more flexible and allow for the priority of either local or contextual factors to be established case by case, depending on the context. In the next paragraph I will suggest that the idea that meaning is immediately contextualized has significant implications not only for the study of actual language processing but, even more importantly to our current aims, for the question of the evolution of language.

3. A top-down model of language evolution

Although human languages are very rich codes, they are not optimal. While in an optimal code a signal is strictly associated with only one meaning, in human languages the same linguistic expression can take on different meanings in different contexts. In the previous section, we have argued that meaning is immediately contextualized such that a single-step model of language processing, which allows for top-down effects of contextual information, can better account for language understanding. This claim has striking implications for the issue of language evolution given that empirically-informed models of modern language functioning should provide constraints to build plausible models of the evolution of language (see Jackendoff, 2010).

A preliminary trivial remark is that a full-fledged code cannot be assumed to be the evolutionary starting point. Therefore, if contextual factors play a major role in the interpretation of completely developed current linguistic codes, they must have been even more crucial in the evolutionary history of language, when those codes were not yet evolved. I am suggesting that the rising verbal communication could not work unless initial linguistic expressions were contextually constrained, just like current linguistic expressions are. It means that the issue of the evolution of language must be addressed assuming a top-down approach. In order to better appreciate the repercussions of this statement and help situating our proposal within other recent theories of language evolution, it might be useful to frame the notion of a top-down model of language evolution with reference to the debate on the passage from protolanguage to language.

The ability to process complex languages thanks to the existence of syntactic abilities is considered by many to be a uniquely human feature. It is hypothesized that the common ancestor of chimpanzees and humans was not in possess of this capacity. Therefore, evolutionary linguists face the problem of explaining the gap between a non-linguistic ancestor and our linguistic (i.e., syntactical) species. The notion of protolanguage has been commonly invoked as a stable intermediary stage in the evolution of language: “[t]he hypothesis of a protolanguage helps to bridge the otherwise threatening evolutionary gap between a wholly ailingual state and the full possession of language as we know it” (Bickerton, 1995, p. 51).

Protolanguage has been defined under two competing accounts: the synthetic account and the holistic account. Under the *synthetic* account

(Bickerton, 1990, 1995; Tallerman, 2007), protolanguage had a limited set of word-like symbols which could be used to convey simple, atomic meanings, effectively the ancestors of modern nouns and verbs. Under the *holistic* account (Wray, 1998, 2000; Arbib, 2005), protolanguage units represented complex propositions, similar to modern sentences but lacking in internal morphological structure. Both accounts assume that protolanguage had distinguishable units, but their disagreement over the level of complexity of those units leads to different ideas of how protolanguage could have developed into modern language. Clearly, from this point of view, the origin and evolution of syntactic language is guided by evolutionary pressures to evolve more and more efficient systems of communication. A detailed review of the current positions concerning the nature of protolanguage is beyond the scope of this paper. However, it might be useful to bear in mind that my argument relies on a notion of protolanguage that seems to be more consistent with the holistic account. In fact, I argue that a linguistic unit conveys a far more complex meaning than assumed by supporters of the synthetic account and, as we will see, this perspective raises a relevant issue concerning which cognitive abilities are involved in inferring such a meaning.

What is even more relevant to the present paper is that this complex meaning is not fixed but rather pretty flexible and needs to be interpreted according to broader contextual constraints. From this perspective then, instead of focusing on the precursors of human syntactic abilities, I adopt a pragmatic perspective on the evolution of language and argue that the origin of modern syntactic language can only be explained by considering the protolanguage in terms of protodiscourse, namely in terms of the pragmatic processes that determine the contextually appropriate interpretation of linguistic expressions (see Ferretti, 2013 for a related perspective). This statement raises the question of how contextual constraints can lead to the appropriate interpretation of an utterance in the evolutionary scenario. What abilities were involved? Also, since we are interested in biological as opposed to cultural evolution, what were the core biological mechanisms that enabled these abilities? In order to answer these questions we need to specify, first of all, what a pragmatic model of language evolution is meant to explain and why it should be a better framework for analyzing the evolutionary issue.

According to the classical model of communication, or “code model” (Shannon and Weaver, 1949), communication is described by the twin acts of encoding and decoding. Speakers encode the meaning into a succession of

sounds and then transmit the signal. Receivers then decode the message in order to be able to share the thoughts of the speaker. From this point of view, the key question is to look for a specific mechanism that can transform the sounds uttered by the speaker into meaning. In opposition to this view, and exploring the implications of a pragmatic perspective for the evolution of language, Origggi and Sperber (2000) and Sperber and Origggi (2010) have argued that the function of language is not to encode the speaker's meaning but rather to offer some hints that can help to infer that meaning. Their claim is that the inferential model is necessary for a structured code to develop. They make the convincing claim that for coded communication to work, speaker and listener have to share exactly the same code. Any difference that affects an individual's code is likely to produce a mismatch between his signals and those of his conspecifics compromising the success of communication. In this situation then, a linguistic mutant who introduces modifications of the code, including the addition of new signals, would have counter-adaptive effects because these modifications would only be advantageous once they are shared by the population.

In the context of the inferential communication though, the situation would be very different. The inferential model in fact does not assume that the code needs to be shared. The code, actually, plays a little role here because the inferential model allows communication to precede grammar such that a fragmentary and ambiguous coding is generally sufficient, in the context, to unequivocally indicate the speaker's meaning. Then, not only the inferential model is necessary for the code to change but, what is more, if we imagine a protolanguage that consists only of words associating sound and meaning without any syntactic structure, such a poor and fragmentary code could be of use only to beings capable of inferential communication. For such individuals, the role of linguistic expressions in the communication process is just that of providing evidence of the intended message which is sufficient for reconstructing a full-fledged meaning, the speaker's meaning. This is a crucial notion within an inferential model of communication. Following Grice (1957), speaker's meaning is defined as a complex communicative intention aimed to achieve a certain effect upon the mind of the hearer by means of the hearer's recognition of the very intention to achieve this effect. In these terms, then, pragmatic interpretation is ultimately an exercise in mind-reading, the ability to infer the mental states of others. In the next section I will focus on this ability providing a new insight on its evolutionary foundation and its pragmatic

function. In particular, while Sperber and Wilson (2002) argue that the interpretation process involves a dedicated “metacommunicative” system, with its own specific principles and mechanisms, I will suggest that the interpretation process is strongly linked to other mechanisms rather than being encapsulated. More to the point, I will argue that motor activity takes part in the construction of meaning.

4. Embodied evolutionary pragmatics

In the last decades there has been much debate about the nature of the ability to ascribe mental states to others and the putative mechanism underlying it. Most of the research in this area has focused primarily on high-level cognitive processing, such as understanding that others can have different desires and beliefs from one’s own (Baron-Cohen et al., 1985). However, from an evolutionary perspective, it is relevant to ask how these complex abilities have originated from more basic capacities, and what the impact of the latter was on human socio-communicative abilities and language evolution. While I adopt a top-down model of language evolution, I argue that the pragmatic processes that allow for the immediate contextualization of meaning are based on very basic cognitive abilities and biological mechanisms. It means that top-down effects are supported by processes and biological mechanisms that can be described assuming a bottom-up perspective. Specifically, I suggest that the notion of the pragmatic origins of language has to be combined with an embodied account of cognition which claims that “high-level” cognitive processes are rooted in “low-level” processes such as action and perception (Pecher & Zwaan, 2005). From this point of view, pragmatic processes are based on sensory-motor processes that enable individuals’ physical interactions with their environment.

In the last twenty-five years, a growing number of studies have emphasized that language and cognition are deeply rooted in the experience that results from possessing a body with certain physical features and a specific sensory-motor system. Most important to our current aims, humans’ ability to attribute mental states to others may also be interpreted in these terms. In particular, “motor theories of social cognition” aim to derive human social cognition from human motor cognition showing that the ability to understand other people’s minds crucially involves the capacity to understand other people’s intentions

by observing their actions. An action is a goal-directed sequence of bodily movements initiated and monitored by a ‘motor intention’. Thus, understanding a perceived action requires at least representing the agent’s motor intention and, possibly, also her social and communicative intentions (Jacoboni et al., 2005; but see Jacob & Jeannerod, 2005 for a critique). Given that human adults readily explain actions ascribing certain beliefs to the agent, it is possible that this lower-level of “theory of mind” is a prerequisite for high-level processing such as representing others’ desires and beliefs (Blackemore & Decety, 2001; Gallese, 2003; Wolpert et al., 2003).

The idea is that the mechanism for inferring other people’s intentions from observed actions might depend on the same mechanism that allows interpreting the consequences of one’s own actions as being produced by one’s own intentions: the intentions of an actor are estimated via a process of *simulation*, in which the actions of the actor are interpreted as if they were one’s own actions, and thus relating them to one’s own intentions (Goldman, 2006). The recently discovered “mirror neurons” may be the cortical mechanism supporting the simulation process, as they are involved in coupling observation and execution of goal-related motor actions responding both when a particular action is performed and when the same action, performed by another individual, is observed. By automatically matching the observed movements of an agent onto one’s own motor repertoire without executing them, the firing of mirror neurons in the observer’s brain simulates the agent’s observed movements and thereby contributes to understand the perceived action (Gallese & Goldman, 1998; Gallese et al., 2004). Because mirror neurons provide motor, not purely perceptual, representations of actions, they must be crucially involved in establishing whether represented actions are executable, that is, consistent with the rules of the motor system. This step – establishing whether the action is possible or not – is crucial for an individual to understand actions. Here I would like to focus on this specific component and on the mechanism of *affordance perception*. Given that this is a very basic mechanism, the attempt to link it to more sophisticated mental processes deserves great attention if we are interested in the foundations of social cognition and, ultimately, of communication.

Affordances are commonly defined as qualities of an object or an environment that, in combination with a particular bodily structure, determine possibilities for actions (Gibson, 1979; Glenberg et al., 2013). That is, for an individual to detect an affordance is to perceive an opportunity for action which

eventually depends on many factors such as the perceptual characteristics of the object or environment, the features of one's body, and the situation at hand. In opposition to more traditional accounts that see perception purely as an input system and action as a separate output system, the notion of affordance incorporates motor aspects in perception, thus emphasizing the close connection between action and perception.

It is interesting to note that the role of affordance perception in language evolution has been previously discussed in the context of theories of the gestural origins of language with reference to the hypothesis of the involvement of the mirror system in language evolution ("the mirror system hypothesis"; see Arbib et al., 2014 for a recent review). According to this hypothesis, a "mirror for actions" system, concerned with both generating an action which is appropriate to object's affordances and recognizing the action being performed by another individual, provided the evolutionary basis for the emergence of a "mirror for words and constructions" system. The hypothesis builds on the distinction between a dorsal path concerned with the *how* of converting affordances of an object into motor parameters for interacting with it and a ventral path involved in determining which actions are to be executed by taking into account not only *what* the object is but also context, task and more. The crucial argument is that word recognition based on how the word is articulated must be differentiated from interpreting the meaning of the word via a ventral pathway. In particular, the dorsal "mirror for actions" system played a particularly relevant role in language evolution supporting attempts to reproduce the articulatory form of words that were not in one's motor repertoire. According to this hypothesis then, a mirror and production system for praxic actions provided the evolutionary basis for the emergence of a mirror and production system for words and larger utterances as articulatory structures. In this way, the mirror system allowed for complex imitation to be transferred from manual skills to a new communicative domain. Although supporters of this hypothesis state that the mirror system involved in affordance perception has little to say about the notion of intentional communication as an exercise in mindreading, I would like to emphasize that the role of affordance perception in the passage from a manual to a vocal form of communication is consistent with the role of this mechanism within a motor account of social cognition.

Even if affordances have been generally studied with respect to the ability of an individual to determine his own action possibilities in the environment,

there is also a wealth of research supporting the notion that affordances perception is fundamental to the success of social interactions. Humans' ability to understand and predict the movements of others in social contexts (Marsh, Richardson, & Schmidt, 2009; Ramenzoni et al., 2008a, 2008b) and to engage in ongoing interactions (Davis et al., 2010; Richardson, Marsh, & Baron, 2007) depends to large extent on it. Perceiving affordances is a key step in the process of building representations of both one's own intended actions and the potential actions of the agent with whom one is interacting. These representations are used to make predictions and estimates of the social consequences of the represented actions and, when these actions come to execution, they are perceived as social signals which can confirm or not the agent's predictions and, possibly, modify the agent's beliefs and desires. What is relevant to our current aims is that in the conceptual framework of the simulation hypothesis, the same sensory-motor processes that allow individuals to determine their own action possibilities in their environment are also recruited to estimate the action possibilities of other people and, thereby, to understand their actions. Affordance perception, then, may play a crucial role in the evolution of socio-communicative abilities.

In support of this hypothesis, it has been recently suggested that an impaired mechanism of affordance perception may be the cause of the known impairments in socio-communicative skills of individuals with Autism Spectrum Disorders (ASDs; Linkenauger et al., 2012). Specifically, subjects with ASDs have impairments in both social and motor skills. Whether they are different deficits or may originate from a common etiology is a highly controversial issue. An interesting recent study has provided data which seems to support the intriguing hypothesis that an impaired ability to perceive affordances is the putative mechanism underlying both impairments. In this study, in order to test the ability to perceive one's own affordances, a well-developed experimental paradigm was used (Ishak, Adolph, & Lin, 2008; Linkenauger et al., 2012; Warren & Whang, 1987) that consisted in asking individuals with ASD and normotypical control subjects to estimate whether they could perform simple actions without overt feedback. Then, subjects were asked to actually perform these actions and their accuracy in perceiving affordances was assessed. In particular, subjects performed three tasks: (1) graspability, (2) reachability, (3) aperture. In the graspability task, they were asked to estimate the extent of their grasping ability with respect to foam board blocks of different sizes; after they finished all estimates, the experiment

assessed whether they could actually grasp the object: a successful grasp was defined as the ability to lift the block completely off the table. For the reachability task, participants observed the experimenter moving a plastic chip away from them or toward them and they were instructed to tell the experimenter when they thought the chip was at the limit of their reach; their actual reaching abilities were subsequently assessed. In the aperture task, subjects had to anticipate if their hand could fit through a hole whose size was gradually manipulated, and after the estimates, the smallest aperture through which they could indeed fit their hand was assessed.

It is relevant to note that these tasks are well-suited to explore affordance perception in individuals with ASDs as they do not involve other skills that might be as well compromised in those individuals. For example, even if autistic people do have general motor problems, they are as accurate as typically-developed children in terms of hand and arm movements and their visual estimation abilities (e.g., size and distance estimation) are even superior. Moreover, deficits in these tasks cannot be attributed to clinical comorbidities since motor planning is intact in other clinical populations such as hyperactivity disorder, and finally affordance perception is not generally associated with imagination or creativity but rather with perceptual-motor integration, so even if individuals with ASDs may have limitations in imagination this cannot be the source of their impaired ability to perceive affordances. Taking into account these remarks, it is particularly interesting that difficulties with affordance perception tasks have been shown to correlate with scores on the Social Communication Questionnaire (SCQ) (Rutter et al., 2007), a specific measure of autistic socio-communicative abilities based on parents' report of the lifetime presence of a child's autistic symptoms. This questionnaire is composed of three subscales, communicative ability, reciprocal social interaction, and restrictive/repetitive behaviors, and highly correlates with "gold standard" diagnostic measures (such as the ADOS). A high SCQ score indicates more autistic symptoms. Relevant to our current aims, higher SCQ scores predict poorer results in the affordance perception tasks for ASD participants with the correlation being driven by communication and reciprocal social interaction subscales.

As mentioned above, in the light of the notion of simulation, impairments in perceiving one's own affordances may also affect the capacity to perceive the action possibilities of other individuals and, consequently, to understand their actions, leading to striking limitations in the social and communicative

domains. Of course, the issue of the role of affordance perception in socio-communicative deficits of autistic individuals needs to be further investigated and we should be very careful in drawing conclusions from this data. However, what this data suggests is that the rather counterintuitive notion of the involvement of a mechanism of sensory-motor integration in communication seems at least to have some plausibility. Exploring this relation might have extremely interesting implications for theories of language evolution given that affordance perception is a very basic mechanism but still may be involved in the critical pragmatic function of inferring communicative intentions, then allowing for top-down contextual effects. These effects, in turn, crucially constrain the interpretation of linguistic expressions enabling the construction of appropriate interpretations and, ultimately, guarantying the success of communication.

5. Conclusions

The aim of this paper was to argue in favor of a top-down model of language evolution in which pragmatic abilities are grounded on sensory-motor processes of action and perception. We have shown that the mechanism of affordance perception that is concerned with the physical interactions between an individual and his environment is also crucially involved in understanding the actions of other individuals and, consequently, in the evolution of socio-communicative abilities that were at the foundation of the pragmatic origins of language. By combining a top-down approach to language evolution with a bottom-up perspective of the cognitive mechanisms underlying language functioning, we have suggested that the notion of an embodied pragmatics may have a key role in understanding the evolutionary origins of language.

REFERENCES

- Arbib, M.A., Casser, B., & Barrès, V. (2014). Language is handy but is it embodied? *Neuropsychologia*, *55*, 57-70.
- Arbib, M. A. (2005). From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences*, *28*(2),105-124.

- Baron-Cohen, S., Leslie, A., & Frith, U. (1985). Does the autistic child have a theory of mind. *Cognition*, 21, 37–46.
- Barsalou, L.W. (1987). *The instability of graded structure: Implications for the nature of concepts*. In U. Neisser (Ed.), *Concepts and conceptual development: Ecological and intellectual factors in categorization*. Cambridge: Cambridge University Press, 101–140.
- Blakemore, S.J., & Decety, J. (2001). From the perception of action to the understanding of intention. *Nature Reviews Neuroscience*, 2, 561–567.
- Barsalou, L.W. (1993). *Flexibility, structure, and linguistic vagary in concepts: Manifestations of a compositional system of perceptual symbols*. In A.F. Collins, S.E. Gathercole, M.A. Conway, & P.E. Morris (Eds.), *Theories of memory*. Hillsdale, NJ: Erlbaum, 29–101.
- Bickerton, D. (1990). *Language and species*. Chicago, IL: University of Chicago Press.
- Bickerton, D. (1995). *Language and human behaviour*. London: University College London Press.
- Borg, E. (2012). *Pursuing Meaning*. Oxford: Oxford University Press.
- Cappelen, H., & Lepore, E. (2005). *Insensitive Semantics: A Defense of Semantic Minimalism and Speech Act Pluralism*. Oxford: Blackwell.
- Carston, R. (2002). *Thoughts and Utterances: The Pragmatics of Explicit Communication*. Oxford: Blackwell.
- Chomsky, N. (1957). *Syntactic Structures*. The Hague: Mouton & Co.
- Cosentino, E., Baggio, G., Kontinen, J., Garwels, T., & Werning, M. (2014). *Lexicon in action: N400 effect on affordances and telicity*. In P. Bello et al. (Eds.), *Proceedings of the 36th Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society, 2079–84.
- Cosentino E., Adornetti I., Ferretti F. (2013). *Processing Narrative Coherence: Towards a top-down model of discourse*. Open Access Series in Informatics (OASICS). Special Issue “2013 Workshop on Computational Models of Narrative”, 32, 61–75.
- Davis, T.J., Riley, M.A., Shockley, K., & Cummins-Sebree, S. (2010). Perceiving affordances for joint actions. *Perception*, 39, 1624–1644.
- Ferretti, F. (2010). *Alle origini del linguaggio umano. Il punto di vista evoluzionistico*. Laterza, Roma-Bari.

- Ferretti F. (2013), Navigation, Discourse and the Origin of Language. In Brambilla P., Marini A. (Eds.), *Brain evolution, language, and psychopathology in schizophrenia*, New York: Routledge, 22-32.
- Gallese, V. (2003). The manifold nature of interpersonal relations: the quest for a common mechanism. *Philosophical Transactions of the Royal Society B*, *358*, 517–528.
- Gallese V., Keysers C. & Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends in Cognitive Sciences*, *8*, 396–403.
- Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, *12*, 493-501.
- Gibson, J.J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.
- Glenberg, A. M., Witt, J. K., & Metcalfe, J. (2013). From the Revolution to Embodiment: 25 Years of Cognitive Psychology. *Perspectives on Psychological Science*, *8*, 573-585.
- Glucksberg, S. (2001). *Understanding Figurative Language: From Metaphors to Idioms*. Oxford: Oxford University Press.
- Goldman, A.I. (2006), *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading*, Oxford: Oxford University Press.
- Grice, P. (1975). *Logic and conversation*. In Cole, P., & J.L. Morgan (Eds.), *Speech Acts*. New York: Academic Press, 41–58.
- Ishak, S., Adolph, K.E., & Lin, G.C. (2008). Perceiving affordances for fitting through apertures. *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 1501–1514.
- Jackendoff, R. (2010). Your theory of language evolution depends on your theory of language. In R. Larson, V. Déprez, & H Yamakido (Eds.), *The Evolution of Human Language: Biolinguistic Perspectives*. Cambridge: Cambridge University Press, 63–72.
- Jacob, P., & Jeannerod, M. (2005). The motor theory of social cognition: a critique. *Trends in Cognitive Sciences*, *9*(1), 21-25.
- Jacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J., & Rizzolatti, G. (2005). Grasping the intentions of others with one's own mirror neuron system. *PLoS Biology* *3*, e79, pp. 1–7.

- Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, *62*, 621-647.
- Linkenauger, S., Lerner, M.D., Ramenzoni, V., & Proffitt, D. (2012). A perceptual-motor deficit predicts social and communicative impairments in individuals with autism spectrum disorders. *Autism Research*, *5*, 352 - 362.
- Marsh, K.L., Richardson, M.J., & Schmidt, R.C. (2009). Social connection through joint action and interpersonal coordination. *Topics in Cognitive Science*, *1*, 320-339.
- Origgi, G., & Sperber, D. (2000). Evolution, communication and the proper function of language. In P. Carruthers & A. Chamberlain (Eds.), *Evolution and the Human Mind: Language, Modularity and Social Cognition*. Cambridge: Cambridge University Press, 140-169.
- Pecher, D., & Zwaan, R. A. (Eds.) (2005). *Grounding cognition: The role of perception and action in memory, language, and thinking*. New York: Cambridge University Press.
- Ramenzoni, V.C., Riley, M.A., Shockley, K., & Davis, T. (2008a). Carrying the height of the world on your ankles: Encumbering observers reduces estimates of how high an actor can jump. *Quarterly Journal of Experimental Psychology*, *61*, 1487-1495.
- Ramenzoni, V.C., Riley, M.A., Shockley, K., & Davis, T. (2008b). An information-based approach to action understanding. *Cognition*, *106*, 1059-1070.
- Recanati, F. (2004). *Literal Meaning*. Cambridge: Cambridge University Press.
- Richardson, M.J., Marsh, K.L., & Baron, R.M. (2007). Judging and actualizing intrapersonal and interpersonal affordances. *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 845-859.
- Rutter, M., Bailey, A., & Lord, C. (2005). *SCQ: The social communication questionnaire manual*. Los Angeles: Western Psychological Services.
- Sperber, D., Origgi, G. (2010). A pragmatic account of the origin of language. In R. K. Larson, V. Déprez, & H. Yamakido, (Eds.) *The Evolution of Human Language: Biolinguistic perspectives*. Cambridge: Cambridge University Press, 124-132.
- Tallerman, M. (2007). Did our ancestors speak a holistic protolanguage?. *Lingua*, *117*, 3, 579-604.

- Warren, W.H., & Whang, S. (1987). Visual guidance of walking through apertures: Body-scaled information for affordances. *Journal of Experimental Psychology: Human Perception and Performance*, 13, 371–383.
- Wilson, D., & Carston, R. (2007). *A unitary approach to lexical pragmatics: relevance, inference and ad hoc concepts*. In N. Burton-Roberts (ed.), *Advances in Pragmatics*. Basingstoke: Palgrave Macmillan, 230-260.
- Wilson, D., & Sperber, D. (2004). *Relevance theory*. In L. Horn, & G. Ward (eds.), *The Handbook of Pragmatics*. Oxford: Blackwell, 607-632.
- Wolpert, D.M., Doya, K., & Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philosophical Transactions of the Royal Society London B*, 358, 593–602.
- Wray, A. (2000). *Holistic utterances in protolanguage: The link from primates to humans*. In C. Knight, M. Studdert-Kennedy, & J. Hurford (Eds.), *The evolutionary emergence of language*. Cambridge: Cambridge University Press, 285-302.
- Wray, A. (1998). Protolanguage as a holistic system for social interaction. *Language and Communication*, 18, 47–67.