AVANT, Vol. XIII, No. 1
ISSN: 2082-6710 avant.edu.pl/en
DOI: 10.26913/ava202204

Motor Simulation and Ostensive-Inferential Communication

Angelo D. Delliponti
Academia Copernicana,
Nicolaus Copernicus University
angelo.d.delliponti@gmail.com

Received 9 September 2021; accepted 19 April 2022; published 17 June 2022.

Abstract

The ostensive-inferential model is a model of communication, an alternative to the code model of communication, based on pragmatic competence: it explains human communication in terms of expression and recognition of informative and communicative intentions, founding comprehension on the distinction between literal meaning and the speaker's meaning. Through informative intentions we try to make evident the content of a message to a receiver, or to make evident what we want to communicate to him/her: communicative intentions are used to make evident the very fact that we intend to communicate. One hypothesis is that ostensive-inferential communication is what makes human language possible. Since an extensive literature has highlighted the role of the Theory of Mind in ostensive-inferential communication, this hypothesis fits with the idea that a mechanism for mentalizing underlies human communication. The aim of the present paper is to stress the role of lower-level mechanisms, specifically of motor simulation, in the recognition of informative and communicative intentions, in order to outline an embodied account of ostensive communication. Specifically, the hypothesis is that this process is involved in language acquisition during development, and that it plays a role in the associative learning process involved in language acquisition during childhood. To this aim, in future research it may be useful to test the involvement of motor simulation (specifically, phono-articulatory and semantic) in the recognition of informative and communicative intentions in toddlers. Since some models of language evolution focus on the role of motor simulation, a supplementary goal is to deepen its role in the biological evolution of language, focusing on the specific link between motor simulation and intentions in the framework of ostensive-inferential model.

Keywords: ostensive communication; motor simulation; communicative intention; informative intention; embodiment; language acquisition

1. Introduction

In this article I argue in favour of the hypothesis that ostensive-inferential communication has an embodied basis, stressing in particular the importance of lower-level mecha-nisms such as motor simulation in the recognition of informative and communicative intentions. I begin by explaining what ostensive-inferential communication is and what are the theoretical principles and the experimental evidence that would make it possible to claim that it has an embodied basis. In doing this, I present a theory according to which the origin of human communication is anchored to gestural communication, which may have guided vocal communication throughout the evolution of language: it is the Mirror System Hypothesis (MSH) (Arbib, 2012). This allows me to show the existing link between ostensive-inferential communication and language: in fact, mirror neurons are involved in the recognition of intentions (Gallese, 2007) and in the processing of words with action content (Pulvermuller, 2005). I then review studies that show how communication and language, in production and understanding, involve areas of the brain dedi-cated to motor processing (Buccino et al., 2001; Fadiga et al., 2002; Hauk et al., 2004; Martin et al., 1996).

Building on those foundations, I hypothesize that there are two subsets of motor simulation—i.e., the reactivation of sensorimotor patterns, extrapolated from their motor functions and exploited in cognitive processes different from those for which they evolved or during which they formed (Borghi & Caruana, 2016)—involved in recognizing informative and communicative intentions during language acquisition: they are phono-articulatory simulation and semantic simulation. Phono-articulatory simulation, which occurs with activation of motor cortex areas involved in speech production, is involved in the recognition of communicative intentions, while semantic simulation, which occurs with activation of motor cortex areas involved in processing action content words, has a role in the recognition of informative intentions. What I want to emphasize here is that these mechanisms play an important role in the acquisition of language during development. Consequently, it is also possible to hypothesize (in a completely speculative way) that motor simulation has a more general role in the recognition of communicative and informative intentions in linguistic communication.

2. Pragmatic competence and ostensive-inferential communication

With pragmatic competence we mean the ability to understand the message conveyed by the utterances in the course of communication; it refers not only to the literal meaning, but mainly to the meaning linked to the context, i.e. the knowledge of the rules of optimal adaptation of a language to the linguistic and extra-linguistic context within which communication takes place (Bambini, 2017). Pragmatic competence is, therefore, the ability to integrate linguistic information with contextual information, in order to understand the meaning of communication beyond the strictly literal level. Within a pragmatic approach based on the distinction between literal meaning and speaker's meaning, the former is a hint of the message: it is therefore possible to transmit messages whose meaning is not to be found only in the code, because the content is often implicit, indirect or non-literal.

To understand the meaning of the speaker, it is thus necessary to take into account the linguistic clue and to integrate linguistic material with the context. The notion of context is not uncontroversial. In my perspective, it can be described as the set of space-time, and cognitive and socio-cultural coordinates in which communicative exchanges take place, including the linguistic material of the discourse (Bambini, 2017). The ostensive-inferential model (or ostensive communication), a communication theory alternative to the code model (Shannon & Weaver, 1949), explains human communication in terms of expression and recognition of informative and communicative intentions (Sperber & Wilson, 1986). But what is a communicative intention and what is an informative intention? With informative intentions we try to make evident the content of a message to a receiver, or what we want to communicate to her: the content of an informative intention is the information provided to the interlocutor, and this information corresponds to the changes that the sender intends to produce in the mental representations of the receiver. With the communicative intentions we want to make evident the very fact that we intend to communicate; furthermore, if the expression of an informative intention is not accompanied by the expression of a communicative intention, communication itself fails (Scott-Phillips, 2015). According to the theoretical framework of ostensive-inferential model, the sender provides hints of his intentions and the receiver interprets them: in fact, the meaning of "ostensive-inferential" is precisely this, that is, ostension as an offer of clues and inference as an interpretation of the clues (Scott-Phillips, 2015). What makes the process possible is the fact that whoever communicates can reason about the intentions and mental states of the interlocutor: the intention of the person who produces ostensive stimuli is, in fact, to modify the mental states of the receiver, and not simply, as in the code model, to send a message to be decoded. So this depends on contextual factors, that is, on the beliefs and knowledge that a speaker has of a listener's beliefs and knowledge, and vice versa (Scott-Phillips, 2015).

To better clarify the difference between communicative intention and informative intention, I will take an example directly from Scott-Phillips (2015):

> I am in a coffee shop, I catch the eye of the waiter, and I tilt my coffee cup in a parti-cular, somewhat stylized way. The waiter then comes over and refills my cup. Here, I have an informative intention that the waiter understands

> that I would like a refill. And so on. The content of an informative intention is, in colloquial terms, the information that it provides. More specifically, it is the changes that the signaler wants to make to the receiver's mental representations. […] The tilt of my coffee cup expresses my informative intention, but it also expresses something just as important: the very fact that I wish to communicate with the waiter at all. […] How does the tilt reveal to the waiter that it is a signal? […] I must also make it clear to the waiter that I am trying to communicate with him at all. My intention to do this—that is, my intention to create in my audience a representation of the fact that I have an informative intention—is called a *communicative intention*. This intention is expressed when I establish eye contact with the waiter and tilt my cup in a particular, *ostensive*, way (pp. 35-36).

3. Pragmatics and the Mirror system hypothesis (MSH)

Scott-Phillips (2015) states that ostensive communication comes before language and that transition to language became possible only after the birth of ostensive communication. „What made ostensive communication possible ultimately made language possible too" (p. 134). Said {thus?} this, the mechanism that makes ostensive communication possible is the cognitive module of Theory of Mind, that is the ability to represent others' mental states and to reason about their thoughts (Premack & Woodruff, 1978; Apperly, 2011; Byom & Mutlu, 2013): the reason why mindreading is considered fundamental to pragmatic competence is that, in the act of communicating, it is important to know certain aspects of the mental dimension of our interlocutor in order to understand his intentions (Scott-Phillips, 2015). The theory of Mind is therefore an essential starting point and, despite the attempts to replace it with various theoretical proposals (see for example Gallese, 2007), it would seem difficult to think of a social cognition or, specifically, a pragmatic competence without a Theory of Mind.

Nevertheless, here—as I will show later—I argue that it is not possible to understand ostensive communication without reference to any embodied foundation, or without integrating it with the mechanism of motor simulation (Borghi, Caruana, 2016; Gallese, 2007), which I assume to be involved in the recognition of communicative and informative intentions. The idea is that motor simulation, through mirror neurons, has made the biological evolution of human language possible, and that at its base there is the expression and recognition of informative and communicative intentions.

Motor simulation has been used in the evolutionary literature for the definition of di-fferent models of language, one of which was proposed by Arbib (2012) through the Mirror System Hypothesis (MSH), an approach that attempts to outline the evolution of language by comparing the systems of praxis and communication of human and non-human primates (Arbib & Rizzolatti, 1997; Rizzolatti & Arbib, 1998), later developed into Cognitive Neuroprimatology (CNP)—(Arbib et al., 2018). In summary, the hypothesis draws on the findings by Poizner et al. (1987), according to which deaf people's lesions in the Broca's area induce, with respect to sign language, a form of aphasia similar, in its outcome, to that of spoken language in subjects with intact hearing. Hence, the hypothesis is that mirror neurons could be the basis for language parity, namely the fact that listeners are able to grasp the speaker's meaning thanks to a system that has a mirror mechanism for gestures at its base, with manual gestures that may have guided vocal gestures throughout the evolution of language. MSH postulates recognition and imitation of complex action as a foundation of the emergence of the language-ready brain (Arbib, 2013): this is compatible with the idea that ostensive communication comes before language and that the transition to language became possible only after the appearance of ostensive communication. Language parity and motor simulation are connected with the neural exploitation hypothesis: the main assumption is that the key aspects of human social cognition are supported by neural exploitation, i.e. an adaptation of the brain mechanisms of sensorimotor integration in order to use them for new purposes concerning thinking and language, and at the same time retain their original functions (Gallese, 2003; Gallese & Lakoff, 2005).

MSH is a gestural hypothesis on the origin of language. There is an extensive literature that places gestural communication as a starting point in the evolution that led to vocal language (Hewes, 1973; Arbib, 2012; Armstrong et al., 1995; Corballis, 2002; Stokoe, 2002; Tomasello, 2008), although, as highlighted by Zywiczynski et al. (2017), in recent years the hypothesis of the original multimodality of proto-language has made its way (Kendon, 2011; McNeill, 2012; Sandler, 2013). There is therefore a vast literature in support of MSH, albeit I do not wish to deny here the possibility of a multimodal origin of human communication and language. I only emphasize that, since the two communication systems—vocal and gestural—for some researchers (McNeill, 2012), are integrated to the point of being part of a single cognitive system (Zywiczynski et al., 2017), I here support the possibility of the compatibility of MSH with the multimodal scenario. However, investigating this aspect goes beyond the scope of this article.

By the way, the present proposal differs from previous models of language evolution because it focuses on the link between motor simulation and intentions, in the framework of ostensive-inferential communication. At the same time, through the latter I intend to distance myself from the syntactic approach of the Universal Grammar (UG): I actually believe that UG (Chomsky, 1957), unlike ostensive communication (which rests on pragmatics), is unable to explain the huge creativity and flexibility of human communication. Indeed UG, which is fully compatible with the code model, states that the basis of language is a set of innate structural rules that evolved completely independent of any pragmatic competence.

But what is the other evidence in favour of the neural exploitation hypothesis and of language parity? Some researchers (Masataka, 2001; Gentilucci et al., 2004a; Bernardis & Gentilucci, 2006) showed that there is a close relationship between the development of both oral and manual motor skills. Among the proposals supporting this thesis is the idea that speech production and manual gestures related to speech can be considered as results of the same process (Goldin-Meadow, 1999); the fact that babbling in 6-8 month old babies is accompanied by rhythmic hand movements (Masataka, 2001); or that children born to deaf parents show hand movement with a rhythm similar to that of babbling (Gallese, 2007). There is also a close relationship between linguistic articulation and manual gestures linked to oral language even in adulthood: in a study by Gentilucci et al. (2004a), participants had to either grasp and bring to the mouth fruits of different sizes such as a cherry or an apple, or observe the same actions performed by someone else, while simultaneously pronouncing the syllable /ba/. What was highlighted is that the second formant of the vowel a, linked to the position of the tongue, increased when they performed or observed the act of bringing the apple to the mouth (or its pantomime), which was the largest object, compared to the case in which the same operations were done with the cherry: this means that the execution/observation influenced the speech production, and that the system involved shares the premotor neural circuits involved in the control of arm/hand actions. Furthermore, it is possible that language production comes precisely from those same mechanisms (Gallese, 2007). Another study is the one of Bernardis and Gentilucci (2006), in which participants had to pronounce words (such as "hello" or "stop"), make communicative arm gestures with the same meaning, or emit the two signals at the same time: results showed how the vocal spectrum of the words was reinforced by the simultaneous execution of the gesture with the same meaning (the second formant) compared to when the words were pronounced alone. The same thing did not happen when the words were meaningless. Saying the words tended rather to inhibit a simultaneous execution of the gesture, and even in this case the effect was not visible with pseudo-words. Subsequently it was found that the reinforcement effect was also present when words were pronounced in response to listening to them and to the simultaneous observation of corresponding gesture by a third person. These results show that „spoken words and symbolic communicative gestures are coded as a single signal by a single communication system within the premotor cortex” (Gallese & Glenberg, 2012, p. 36). Other studies confirm the involvement of Broca area (Gentilucci et al., 2006): since the region contains mirror neurons, it is very likely that the communicative meaning of gestures is merged, through motor simulation, with the articulation of the sounds required to express them in words.

4. Understanding intentions

I claimed that ostensive-inferential communication explains human communication in terms of expression and recognition of informative and communicative intentions (Sperber & Wilson, 1986). But what does it mean to understand the intentions behind someone else's actions? According to Gallese (2007), understanding the reason for performing a certain act, such as grabbing a cup, means detecting the goal of the next imminent and not yet completed act, for example, bringing the cup to the mouth. At the basis of this theorization there was an experiment carried out with functional magnetic resonance imaging (fMRI) (Iacoboni et al., 2005). Volunteers observed three types of stimuli: actions such as a grasping hand without context; the context, like a scene with objects; and a grasping hand inserted in some context. The observation of motor acts within a context, compared with the other two experimental conditions, produced a significant increase in the signal in the posterior part of the inferior frontal gyrus and the ventral premotor cortex, correlated with the actions of the hands. Therefore, according to Gallese (2007), the premotor mirror areas, active both during the execution and during the observation of the movements, are not only involved in the recognition of the action, but also in understanding the reason for an action, or rather the intention of its underlying motives. It would thus be the mirror system to make possible this mechanism through the automatic activation of motor simulation.

Another study (Fogassi et al., 2005) found a class of mirror neurons in the parietal area whose activation during the observation of an act, such as grasping an object, is conditioned by the kind of subsequent act not yet detected, for example bringing the object to the mouth, thus specifying the overall intention of the action; these neurons are activated only in reference to the execution/observation of motor acts linked to a specific action, but aimed to a more distal goal: this neuronal activation occurs in a monkey before the execution/observation of the movement linked to the distal goal. According to Gallese (2007), this means that in addition to target recognition, mirror neurons allow the observing monkey to perform a targeted act (for example, bringing an object to the mouth rather than placing it in a container) to predict what the agent is about to do, thus understanding the overall intention of the action. This mechanism found in non-human primates could be the basis of the most sophisticated forms of understanding intentions typical of our species.

Mirror neurons could therefore play an important role in the recognition of intentions. It could be the recognition of the speaker's intentions—through mirror neurons—at the basis of our communicative ability, and specifically, the recognition of communicative and informative intentions as the basis for the evolution of language. In fact, we will see how mirror neurons could have made it possible to move from the recognition of intentions for primordial communicative purposes (at the beginning presumably, as already seen, in the form of manual gestures) to language, and we will see it by showing that mirror neurons are also involved in recognizing intentions in language.

Before continuing, however, it is right to make a clarification on the role that mirror neurons have in the recognition of intentions: this in the light of the various criticisms that have emerged, especially in the last ten-twelve years

(Cook et al., 2014; Hickok, 2009), on the importance attributed to mirror neurons regarding their role in the aforementioned process, or even more important, in that of understanding what is meant by the expression "understanding actions", which presupposes understanding intentions. Identifying goals and intentions requires a generalization on the perceptual characteristics of the observed actions (Thompson et al., 2019). This is because a goal (such as "to grab") or an intention ("to drink"), can be achieved using different types of grip, and most importantly, the same type of grip can be used to accomplish a large number of different goals and intentions. Since there is no one-to-one correspondence between body part configurations, goals, and intentions (Jacob & Jeannerod, 2005), the same pattern of mirror neuron activation cannot simultaneously represent the action, goal, and intention of the other (Thompson et al., 2019a). Some researchers claim to have found that mirror neurons allow a distinction between different targets (Hafri et al., 2017); however, other evidence has shown that mirror neuron brain areas encode different types of actions based on their perceptual characteristics (Nicholson et al., 2017), suggesting that mirror neuron areas appear to be able to encode the targets of observed actions only when those targets are perceptually distinguishable. Furthermore, generalization about the perceptual characteristics of observed actions appears to occur in conjunction with activity in other, non-motor brain regions that are thought not to contain mirror neurons (Wurm et al., 2016; Wurm & Lingnau, 2015; Spunt & Adolphs, 2014; Spunt, Lieberman, 2013). What therefore seems important to underline is that the main error in the scientific literature on mirror neurons is when it is attributed to them a homuncular-like functioning (Mikulan et al., 2014), as for example in the hypothesis of direct correspondence, which states that an action is understood when its observation causes a resonance in the motor system of the observer (Rizzolatti et al., 2001); this is a case in which the mirror system is given an automatic and mandatory mechanism for understanding (Csibra, 2007). It is therefore possible that mirror neurons alone are not sufficient to explain the encoding of the intentions of others, that is, of the mental states underlying the observed actions. However, there is evidence to support the thesis that mirror neurons are involved in identifying the configuration of body parts when we observe an action (Thompson et al., 2019a). Moreover, it's possible that „the information encoded by mirror neurons is then used by different brain areas in order to identify the mental state underlying an observed action" (Thompson et al., 2019b, p. 110). The most recent approaches to the interpretation of the functioning of the mirror neuron system (MNS) see mirror neurons as part of a system or network that goes beyond the motor cortex and extends to other parts of the brain, including those involved in high-level cognitive processes such as mentalization (Salo et al., 2019). This is also due to the evidence found in laboratory on the increase in connectivity between the areas of mirror neurons and those involved in the processing of others' mental states, when participants are asked to infer the intentions underlying an observed action, with respect to the condition in which they have to judge only how an action is performed (Thompson et al., 2019b; Cole et al., 2019; Libero et al., 2014; Cavallo et al., 2015).

5. Motor simulation and language understanding

As anticipated, motor simulation—a process made possible by neural exploitation—is the reactivation of sensorimotor patterns, detached from their motor functions and exploited in cognitive mechanisms different from those for which they evolved (Borghi & Caruana, 2016). From the perspective of embodied cognition, motor simulation is usually understood as an automatic mechanism and is made possible by mirror neurons. Several studies conclude that it is involved in understanding others' intentions (Binkofski & Buccino, 2006; Gallese, 2007): as said, mirror areas, active both during the execution and during the observation of movements, are not only involved in the reco-gnition of an action, but also in understanding the underlying reasons for the action or its intention.

What could instead be the meeting points between motor simulation, communication, and language? First of all, several studies show how language and action are linked together. One of these has to do with the indexical hypothesis (Glenberg & Robertson, 1999), according to which sentences are understood by creating a simulation of the actions underlying them. In one experiment, Glenberg and Kaschak (2002) created a set in which participants had to judge the meaningfulness of sentences describing the transfer of concrete objects, for example "Andy gave you the pizza/you gave the pizza to Andy", and abstract information such as "Liz told you a story/you told Liz a story": half of the sensible sentences described a transfer to the reader, the other half from the reader to someone else. Participants responded using a box with three buttons held in such a way that the buttons were aligned on the forward/back axis: the sentences were read by holding down the central button with the desired hand. In one condition, the sensible response was made by moving a hand towards the distant button, which then required a movement consisting in simulating a transfer to another person; in the other condition, the response was made by pressing the nearby button, which required a movement similar to a transfer from another person to the reader. As expected, an interaction was found with the time necessary to judge the meaning of a sentence: judgements were faster when the action implied by the sentence matched the action required for the response (approaching or moving away from the body), and this was true for both concrete and abstract transfer sentences. The authors referred to this interaction as the Action-sentence Compatibility Effect (ACE). ACE-type interactions have also been reported from studies employing the use of hypothetical phrases (De Vega, 2008) (for a critique of ACE, see Morey et al., 2021). These results are then confirmed in neuroimaging studies and in the neuropsychological literature, for example, Bak and Hodges (2003) have dealt with how the degeneration of the motor system associated with a motor neuron disorder—in this case referring to amyotrophic lateral sclerosis (ALS)—influences the understanding of action verbs more than nouns; other studies refer instead to the use-induced plasticity of the motor system in influencing the processing of concrete and abstract language (Glenberg et al., 2008), or to the early activation of the motor system following the presentation of a stimulus (Pulvermuller, 2008).

It has been shown that in humans the observation of actions performed with different effectors (hand, foot, mouth) involves the same motor representations that are active during the execution of those same actions (Buccino et al., 2001): this has provided further evidence of the existence of the mirror system in humans, which in our species is not confined only to the Broca area (corresponding to the premotor area F5 of the macaques), but also includes the parietal lobe. Furthermore, an activation of the mirror system is observed, caused by the simple perception of the sound of an action or even when the actions are described verbally (Rizzolatti & Craighero, 2004; Buccino et al., 2004b, 2006); there was also found a somatotopic organization and an overlap between the motor areas activated during the observation of the actions and the motor areas activated during the understanding of the sentences describing those actions (Aziz-Zadeh et al., 2006). These latest studies provide solid evidence for the thesis that motor simulation plays a role in language understanding (we will also see others). The idea is that when individuals listen to words or phrases that imply actions, a modulation of the mirror system should correspond: the effect of this modulation would then influence the excitability of the primary motor cortex and therefore the production of the movements it controls (Buccino et al., 2005; Hauk et al., 2004; Tettamanti et al., 2005).

5.1. Motor simulation and language acquisition

There are some studies and theories in defence of the thesis according to which motor simulation (that is, the resonance mechanism of the motor cortex allowed by mirror neurons) is involved in the process of language acquisition during development. One of these is Gallese's and Glenberg's (2012) Action-based Language (ABL) theory. In summary, it predicts that it should be easier for infants to learn the names of actions and objects with which they have already learned the appropriate modes of interaction (Huttenlocher et al., 1983). Here we report an example made by the authors (Gallese & Glenberg, 2012) concerning how an infant, who already knows the practical ways of drinking, could learn the verb "to drink". At the moment when an infant is drinking from a bottle, the parent could say "good drink!": the child's mirror neurons would be activated by the parent's speech act and a hebbian learning process would begin to establish connections between the control of the action aimed at drinking and motor representation of the vocal signal. Then the parent might say "drink (from) your bottle!": if the child has already learned the name "bottle", then she may direct her attention to the bottle, grab it and start drinking. Suppose instead that the child focuses on the unknown word "drink" and does not engage in the corresponding action. At this point the parent could say "look, this is what drinking means", and then mimic the act of drinking from the bottle: since the child already knows how to drink, her mirror system would activate the controller necessary for drinking, making possible therefore also in this case a hebbian learning between the modules of the word and those of the action.

As already mentioned, the ABL model for verb learning predicts that infants learn verbs more efficiently if they first learned the corresponding actions: Angrave and Glenberg (2007) found data consistent with this prediction using data taken from MacArthur Child Development Inventory. They estimated the average age, in months, for the acquisition of actions such as drinking, scouting, reading, and the average age of the production of the corresponding verbs: the correlation between the two was very strong, thus making it possible to conclude that the development of word went hand in hand with the development of action (although there was a gap between the production of the action and the production of the word. For an explanation of the reason for this gap, see Angrave & Glenberg, 2007; Gallese & Glenberg, 2012; Wolpert & Kawato, 1998).

The idea is therefore that the sensorimotor system is involved in the perception of action, an activity that would be at the basis of verb acquisition (Pulverman et al., 2006). On the other hand, some studies show how, in adults, the sensorimotor processing of verbs can be a result of associative learning (Cooper et al., 2013; Heyes, 2010), for example as a consequence of training through the use of action pseudoverbs: what is taking place here is the mapping of new verbs onto unfamiliar actions (Fargier et al., 2012). In general, associative learning is involved in the sensorimotor processing of action representations already at an early stage of development, as evidenced by studies showing the role of associative learning in the sensorimotor processing of sounds linked to actions in children between 7 and 9 months (Gerson et al., 2015; Paulus et al., 2013, 2012). Finally, a recent study by Antognini and Daum (2019) showed that toddlers' (18 and 24 months old) sensorimotor system is active during the processing of action-related verbs, concluding that the sensorimotor system plays a role in the processing of action verbs during initial phase of linguistic acquisition. In fact, as pointed out by the authors, the first verbs learned by toddlers are „to a great extent verbs that describe observable actions of people” (p. 82), while the more abstract ones are learned later.

6. Motor simulation: phono-articulatory level, semantic level and ostensive communication. Recognizing intentions in language

So let us now recap some of the statements we have encountered so far:

1) What made ostensive communication possible (i.e. the expression and recognition of communicative and informative intentions) is also what made language possible.

2) Mirror neurons are involved in the recognition of intentions, through the mechanism of motor simulation.

3) Motor simulation is involved in language understanding and in language acquisition.

My conclusion is that the recognition of communicative and informative intentions, through motor simulation, may play an important role in language acquisition, and that it might have played a role in language evolution (see Figure). In the first case (communicative intention) we would have a motor simulation at the phono-articulatory level, occurring with activation of motor cortex areas involved in speech production, which could have a role in the recognition of communicative intentions in verbal communication, a role that I assume to be important for language acquisition during development. What is the evidence for a motor simulation at the phono-articulatory level? In a TMS experiment (Fadiga et al., 2002) it was highlighted how listening to the phonemes induces an increase in motor evoked potentials (MEPs) amplitude registered from muscles of the tongue normally involved in their production: the result was interpreted as a resonance mechanism acoustically connected to the phonological level, a phenomenon confirmed by several other studies (Gallese, 2007). Furthermore, McGuigan and Dollins (1989) showed by electromyography that the muscles of the tongue and lips are activated in the same way both during normal speech production and in covert speech.

What I would like to underline here is that, similarly to what happens in cases of activation of the mirror system and motor areas when listening to sounds linked to actions, phono-articulatory resonance is involved in the processing of the communicative act in itself (Fischer & Zwaan, 2008). „If a listener's speech motor system responds to hearing the word "kick", then this would be an example of communicative motor resonance; the motor system is simulating the production of the utterance" (p. 837). The hypothesis is therefore that this mechanism plays a role in the recognition of communicative intentions. However, this does not mean that motor simulation necessarily *always* (i.e. in any case) has a role in the recognition of communicative intentions (or even informative intentions, as we will see soon) during linguistic communication, but that it plays an important role (for the recognition of communicative and informative intentions) for the purpose of language acquisition. When a child listens to a word or utterance, as a consequence there is a resonance at the level of the phono-articulatory system, although the recognition of a communicative intention may occur in different ways—for example through the perception of facial expressions or gestures, even in linguistic communication (Wilson &

Sperber, 2002)—my hypothesis is that the communicative resonance mechanism plays an important role in the development, to ensure that attention is directed to language and not to other systems of communication. More specifically, this is the recognition of a *linguistic* communicative intention: it is possible that this mechanism is the basis for the recognition of a linguistic informative intention.

In the second case (informative intention) we would have a semantic simulation, occurring with activation of motor cortex areas involved in processing action content words, which would have a role in the recognition of informative intentions in verbal communication, a role that I assume—also in this case—to be important for language acquisition during development, insofar as words map onto actions. This hypothesis is compatible with the idea that the environment in which our ancestors lived triggered selection pressures in favour of expression of vocal information with action content: communication and language evolved for the purpose of action (Borghi & Caruana, 2016). So, semantic level motor simulation made it possible to think that language understanding has an embodied basis: as regards words and phrases that express action contents, neural structures that preside over the execution of the action could play a role also in understanding the semantic content of the same actions when they are verbally described. As we saw, this emerges from the study by Glenberg and Kaschak (2002), the so-called Action-sentence Compatibility Effect. Studies with TMS showed that pronouncing names of tools, as opposed to those of animals, differentially activates the left middle temporal gyrus, which is also activated with action tasks, as well as the left premotor cortex, which is activated generally when participants imagine themselves grasping objects with their dominant hand (Martin et al., 1996). Other studies instead show that exposure to words that indicate actions or tools produces a motor resonance, which manifests itself with an activation of the motor areas. Exposure to action verbs and words referring to tools semantically related to actions produces a stronger activation of the frontal-central cortical area than exposure to words referring to objects (Martin et al., 1996; Preissl et al., 1995; Pulvermuller et al., 1999). Specifically, action words related to movements of the face, arms or legs (Hauk et al., 2004), activate the fronto-central cortex in a somatotopic way, coherently with the affirmation that the sensorimotor cortex processes aspects of the meaning of words related to action (Pulvermuller, 2005). Further evidence of the automatic activation of motor representations following exposure to action verbs comes from a study conducted with high-density magnetoencephalography (Pulvermuller et al., 2005): subjects were engaged in a task with a distractor as they listened to words that denoted actions involving the leg or face. Different patterns of cortical activation were identified for words referring to leg or face in premotor areas: stimuli for the face-words activated lower front-central areas much more than for the leg-words, while for the opposite an activation of the upper central areas was highlighted. In addition, activations occurred 170 ms after the start of words. Pulvermuller and colleagues

(2005) interpreted the results as a reflex of the cortical somatotopic arrangement of motor actions signified by the words. This shows that access to meaning in the recognition of action words is an early automatic process, evidenced by the space-time indications of the activity evoked by the words (Fischer & Zwaan, 2008).

It is important here to highlight that semantic simulation is involved in the recognition of informative intentions since, through associative learning, it is possible to create a correspondence between the recognition of the goal of the action and the word intended to express the content of that action. This is also possible because the ability to perceive actions emerges quite early in development, during a prelinguistic stage (Antognini, Daum, 2019). Furthermore, infants already at the age of 6 months „perceive actions as being directed towards goals". What is important here, therefore, is not to demonstrate the role of motor simulation in every single aspect of linguistic processing (a controversial hypothesis that in recent times has been replaced by dual theories of understanding. See Paternoster & Calzavarini, 2020), but rather its role in the recognition of informative intentions—certainly linked to the semantic processing—of words with action content. In fact, this is important because, as already mentioned, the first verbs learned by infants are largely words that refer to observable actions (Antognini & Daum, 2019).

I would like to underline that my hypothesis is that in both cases, phono-articulatory simulation and semantic simulation, what are simulated are, respectively, the communicative intentions and the informative intentions of verbal communication. Simulation at the phono-articulatory level, as a resonance of the human communicative system linked to the production of speech, could trigger the recognition of communicative intentions, while semantic simulation, being sensitive to the content of the words, could trigger the recognition of informative intentions. It is difficult to say to what extent these mechanisms are involved in language processing, but if expression and recognition of communicative and informative intentions are at the basis of the production/understanding between sender and recipient in language (Scott-Phillips, 2015), then motor simulation must have had a role in the evolution of language, in particular in the transition from manual gestures to vocal gestures. This may have been the initial infrastructure that led to the use of recursive mindreading in ostensive communication (Scott-Phillips, 2015).

7. Conclusions

As I tried to demonstrate in this paper, there are two subsets of motor simulation involved in recognizing informative and communicative intentions: phono-articulatory simulation and semantic simulation. The first, which occurs with activation of motor cortex areas involved in speech production, is involved in the recognition of communicative intentions; the second, which occurs with activation of motor cortex areas involved in processing action content words, has a role in the recognition of informative intentions. The hypothesis is that both have a role in the acquisition of language during development, that is, by means of the recognition of intentions through motor simulation. As already seen, some experiments have tested the hypothesis according to which embodied theories of language comprehension predict that when individuals listen to words or phrases that imply actions, a modulation of the mirror system should correspond (Buccino et al., 2005; Hauk et al., 2004; Tettamanti et al., 2005): in turn this would affect the activation of the primary motor cortex. Overall, several studies support the finding that motor resonance occurs automatically during exposure to words with action content (nouns, verbs, adjectives) (Fischer & Zwaan, 2008).

A goal of future research may be, in addition to testing the role of intention recognition through motor simulation in toddlers' language acquisition, to check the involvement of motor simulation when we infer communicative intentions and informative intentions during verbal communication. One hypothesis is that phono-articulatory simulation is a mechanism, if not sufficient, at least necessary for communicative intentions recognition. As far as semantic simulation is concerned, is it also a necessary mechanism? Both of these aspects of embodied ostensive communication must be tested in the laboratory.

The second hypothesis to emerge from this paper is that, from the point of view of biological evolution, the environment in which our ancestors lived triggered selection pressures in favour of expression of vocal information with action content: communication and language evolved for the purpose of action. This could partly explain the experimental evidence showing the link between motor simulation and words / phrases with action content. Another aspect to deepen may be to understand the role played by motor simulation in the evolution of language, within the framework of embodied ostensive communication.

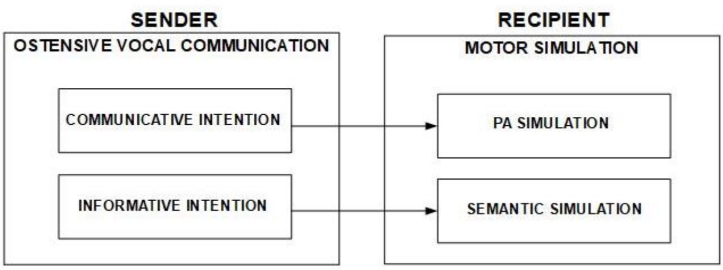

Figure: A diagram that shows how motor simulation may have guided the understanding and evolution of vocal communication.

References

Angrave L.C., Glenberg A.M. (2007). Infant Gestures Predict Verb Production One Year Later. *Paper presented at the annual meeting of the American Psychological Association.*

Antognini K., Daum M. M. (2019) . Toddlers show sensorimotor activity during auditory verb processing. *Neuropsychologia*, *126*, 82-91.

Apperly I.A. (2011). *Mindreaders: the cognitive basis of "theory of mind"*. Psychology Press, New York.

Arbib M.A. (2012). *How the brain got language: The mirror system hypothesis*, New York (NY), Oxford University Press.

Arbib M.A. (2013). Complex imitation and the language-ready brain. *Language and Cognition, 5*, 273–312.

Arbib M.A., Rizzolatti G. (1997). Neural expectations: A possible evolutionary path from manual skills to language. *Communication and Cognition, 29*, 393–424.

Armstrong, D. F., Stokoe,W. C., & Wilcox, S. E. (1995). Gesture and the nature of language. Cambridge, UK: Cambridge University Press.

Aziz-Zadeh L., Wilson S.M., Rizzolatti G., Iacoboni M. (2006). Embodied semantics and the premotor cortex: Congruent representations for visually presented actions and linguistic phrases describing actions. *Current Biology, 16*, 1818-1823.

Bak T.H., Hodges J.R. (2003). The effects of motor neurone disease on language: Further evidence. *Brain and Language, 89*, 354-361.

Bambini V. (2017). *Il cervello pragmatico,* Roma, Carocci.

Bernardis P., Gentilucci M. (2006). Speech and gesture share the same communication system. *Neuropsychologia, 44*, 178–190.

Binkofski F., Buccino G. (2006). The role of ventral premotor cortex in action execution and action understanding. *J. Physiol. Paris 99*, 396–405.

Borghi A., Caruana F. (2016). *Il cervello in azione*, Bologna, Il Mulino.

Buccino G., Binkofski F., Fink G.R., Fadiga L., Fogassi L., Freund H.J., et al. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: An fMRI study. *European Journal of Neuroscience, 13*(2), 400-404.

Buccino G. Binkofski F., Riggio L. (2004b). The mirror neuron system and action recognition. *Brain and Language, 89*(2), 370-376.

Buccino G., Riggio L., Melli G., Binkofski F., Gallese V., Rizzolatti G. (2005). Listening to action-related sentences modulates the activity of the motor system: a combined TMS and behavioral study. *Cogn. Brain Res. 24*, 355–363.

Buccino G., Solodkin A., Small S. (2006). Functions of the mirror neuron system: Implications for neurorehabilitation. *Cognitive Behavioral Neurology, 19*, 55-63.

Byom L. J. & Mutlu B. (2013). Theory of mind: Mechanisms, methods, and new directions. *Frontiers in human neuroscience*, *7*, 413.

Cavallo, A., Lungu, O., Becchio, C., Ansuini, C., Rustichini, A., & Fadiga, L. (2015). When gaze opens the channel for communication: Integrative role of IFG and MPFC. *NeuroImage*, *119*, 63–69.

Chomsky N. (1957). *Syntactic structures*, Princeton, Mouton and Co.

Cole, E.J., Barraclough, N.E., Andrews, T.J. (2019). Reduced connectivity between mentalizing and mirror systems in autism spectrum condition. *Neuropsychologia 122*, 88–97.

Cook R., Bird G., Catmur C., Press C., Heyes C. (2014). Mirror neurons: from origin to function. Behav. *Brain Sci. 37*, 177–192.

Cooper, R. P., Cook, R., Dickinson, A., & Heyes, C. M. (2013). Associative (not Hebbian) learning and the mirror neuron system. *Neuroscience Letters*, *540*, 28-36.

Corballis, M. C. (2002). From hand to mouth: The origins of language. Princeton, NJ: Princeton University Press.

Csibra G. (2007). Action mirroring and action understanding: an alternative account. In: P. Haggard, Y. Rosetti, M. Kawato, *Sensorimotor Foundations of Higher Cognition. Attention and Performance XII*, Oxford, Oxford University Press 2007 (pp. 453–459).

De Vega M. (2008). Levels of embodied meaning. From pointing to counterfactuals. In: De Vega M., Glenberg A.M., Graesser A.C., *Symbols, Embodiment, and Meaning*. Oxford, Oxford University Press 2008 (pp. 285-308).

Fadiga L., Craighero L., Buccino G., Rizzolatti G. (2002). Speech listening specifically modulates the excitability of tongue muscles: A TMS study. *European Journal of Neuroscience, 15*(2), 399-402.

Fargier, R., Paulignan, Y., Boulenger, V., Monaghan, P., Reboul, A., & Nazir, T. A. (2012). Learning to associate novel words with motor actions: Language-induced motor activity following short training. *Cortex*, *48*(7), 888-899.

Fischer M.H., Zwaan R.A. (2008). Embodied language: a review of the role of the motor system in language comprehension. *Quarterly journal of experimental psychology, 61*(6), 825-850.

Fogassi L., Ferrari P. F., Gesierich B., Rozzi S., Chersi F., Rizzolatti G. (2005). Parietal lobe: from action organization to intention understanding. *Science, 302*, 662–667.

Gallese V. (2003). A neuroscientific grasp of concepts: from control to representation. *Phil. Trans. R. Soc. B, 358*, 1231–1240.

Gallese V. (2007). Before and below 'theory of mind': embodied simulation and the neural correlates of social cognition. Phil. *Trans. R. Soc. B., 362*(1480), 659–669.

Gallese V., Glenberg A.M. (2012). Action-based language: a theory of language acquisition, comprehension, and production. *Cortex, 48*(7), 905–922.

Gallese V., Lakoff G. (2005). The brain's concepts: the role of the sensory-motor system in reason and language. *Cogn. Neuropsychol. 22*, 455–479.

Gentilucci M., Bernardis P., Crisi G., Volta R. D. (2006). Repetitive transcranial magnetic stimulation of Broca's area affects verbal responses to gesture observation. *J. Cogn. Neurosci., 18*, 1059–1074.

Gentilucci M., Santunione P., Roy A. C., Stefanini S. (2004a). Execution and observation of bringing a fruit to the mouth affect syllable pronunciation. *Eur. J. Neurosci. 19*, 190–202.

Gerson, S. A., Bekkering, H., & Hunnius, S. (2015). Short-term motor training, but not observational training, alters neurocognitive mechanisms of action processing in infancy. *Journal of Cognitive Neuroscience*, *27*(6), 1207-1214.

Glenberg A.M., Kaschak M.P. (2002). Grounding language in action. *Psychonomic Bulletin and Review, 9*(3), 558-565.

Glenberg A.M., Robertson D.A. (1999) Indexical understanding of instructions. *Discourse Processes, 28*, 1-26.

Glenberg A.M., Sato M., Cattaneo L. (2008). Use-induced motor plasticity affects the processing of abstract and concrete language. *Current Biology, 18*(7), R290-R291.

Goldin-Meadow S. (1999). The role of gesture in communication and thinking. *Trends Cogn. Sci., 3*, 419–429.

Hauk O., Johnsrude I., Pulvermueller F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron 41*, 301–307.

Hafri A., Trueswell J.C., Epstein R.A. (2017). Neural representations of observed actions generalize across static and dynamic visual input. *J. Neur*osci. 37, 3056–3071.

Hewes, G.W. (1973). Primate communication and the gestural origins of language. *Current Anthropology 14*, 5–24.

Heyes, C. (2010). Where do mirror neurons come from?. *Neuroscience & Biobehavioral Reviews*, *34*(4), 575-583.

Hickok G. (2009). Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *J. Cogn. Neurosci. 21*, 1229–1243.

Huttenlocher J., Smiley P., Charney R. (1983). Emergence of action categories in the child: Evidence from verb meanings. *Psychological Review, 90*(1), 72-93.

Iacoboni M., Molnar-Szakacs I., Gallese V., Buccino G., Mazziotta J., Rizzolatti G. (2005). Grasping the intentions of others with one's owns mirror neuron system. *PloS Biol. 3*, 529–535.

Jacob P., Jeannerod M. (2005). The motor theory of social cognition: A critique. *Trends Cogn. Sci. 9*, 21–25.

Kendon, A. (2011). Some modern considerations for thinking about language evolution: a discussion of the evolution of language by Tecumseh Fitch. *Public J. Semiotics 3*(1), 79–108.

Libero, L.E., Maximo, J.O., Deshpande, H.D., Klinger, L.G., Klinger, M.R., Kana, R.K. (2014). The role of mirroring and mentalizing networks in mediating action intentions in autism. *Mol. Autism 5* (1), 50.

Martin A., Wiggs C.L., Ungerleider L.G., Haxby J.V. (1996). Neural correlates of category-specific knowledge. *Nature, 379*, 649–652.

Masataka N. (2001). Why early linguistic milestones are delayed in children with Williams syndrome: late onset of hand banging as a possible rate-limiting constraint on the emergence of canonical babbling. *Devel. Sci. 4*, 158–164.

McGuigan F. J., Dollins A. B. (1989). Patterns of covert speech behavior and phonetic coding. *Pavlovian J. Biol. Sci. 24*, 19–26.

McNeill, D. (2012). How Language Began: Gesture and Speech in Human Evolution. Cambridge University. Press, Cambridge.

Mikulan E.P., Reynaldo L., Ibanez A. (2014). Homuncular mirrors: misunderstanding causality in embodied cognition. *Front. Hum. Neurosci., 8*, 1-4.

Morey, R. D., Kaschak, M. P., Díez-Álamo, A. M., Glenberg, A. M., Zwaan, R. A., Lakens, D., ... & Ziv-Crispel, N. (2021). A pre-registered, multi-lab non-replication of the action-sentence compatibility effect (ACE). *Psychonomic bulletin & review*, 1-14.

Nicholson T., Roser M., Bach P. (2017). Understanding the goals of everyday instrumental actions is primarily linked to object, not motor-kinematic, information: Evidence from fMRI. *PLoS One, 12*, 1–21.

Paternoster A., Calzavarini F. (2020). *Comprendere il linguaggio*. Il mulino.

Paulus, M., Hunnius, S., Van Elk, M., & Bekkering, H. (2012). How learning to shake a rattle affects 8-month-old infants’ perception of the rattle's sound: electrophysiological evidence for action-effect binding in infancy. *Developmental Cognitive Neuroscience*, *2*(1), 90-96.

Paulus, M., Hunnius, S., & Bekkering, H. (2013). Neurocognitive mechanisms underlying social learning in infancy: infants’ neural processing of the effects of others’ actions. *Social Cognitive and Affective Neuroscience*, *8*(7), 774-779.

Poizner H., Klima E.S., Bellugi, U. (1987). *What the hands reveal about the brain,* Cambridge MA, MIT Press.

Preissl H., Pulvermuller F., Lutzenberger W., Birbaumer N. (1995). Evoked potentials distinguish between nouns and verbs. *Neuroscience Letters, 197*, 81–83.

Premack D., Woodruff G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, 1*(4), 515-526.

Pulverman, R., Hirsh-Pasek, K., Golinkoff, R. M., Pruden, S., & Salkind, S. (2006). Conceptual foundations for verb learning: Celebrating the event. *Action meets word: How children learn verbs*, *2010*, 134-159.

Pulvermuller F. (2005). Brain mechanisms linking language and action. *Nature Reviews Neuroscience, 6*, 576–582.

Pulvermuller F. (2008). Grounding language in the brain. In: De Vega M., Glenberg A.M., Graesser A.C., *Symbols, Embodiment, and Meaning*. Oxford, Oxford University Press 2008 (pp. 85-116).

Pulvermuller F., Lutzenberger W., Preissl H. (1999). Nouns and verbs in the intact brain:

Evidence from event-related potentials and high-frequency cortical responses. *Cerebral Cortex, 9*, 498–508.

Pulvermuller F., Shtyrov Y., Ilmoniemi R. (2005). Brain signatures of meaning access in action word recognition. *Journal of Cognitive Neuroscience, 17*, 884–892.

Rizzolatti G., Arbib M.A. (1998). Language within our grasp. *Trends in Neurosciences, 21*(5), 188–194.

Rizzolatti G., Craighero L. (2004). The mirror-neuron system. *Annual Review of Neuroscience, 27*, 169-192.

Rizzolatti G., Fogassi L., Gallese V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nat.Rev. Neurosci., 2*, 661–670.

Salo, V. C., Ferrari, P. F., & Fox, N. A. (2019). The role of the motor system in action understanding and communication: Evidence from human infants and non-human primates. *Developmental psychobiology*, *61*(3), 390-401.

Sandler, W. (2013). Vive la différence: sign language and spoken language in language evolution. *Lang. Cogn. 5*(2–3), 189–203.

Scott-Phillips T. (2015). *Speaking our minds*, Palgrave Macmillan (trad. it. *Di' quello che hai in mente*, Roma, Carocci 2017).

Shannon C.E., Weaver W. (1949). *The mathematical theory of communication*, Illinois, The University of Illinois Press.

Sperber D., Wilson D. (1986). *Relevance: Communication and Cognition,* Oxford, Blackwell.

Spunt R.P., Adolphs R. (2014). Validating the why/how contrast for functional MRI studies of theory of mind. *Neuroimage, 99*, 301–311.

Spunt R.P., Lieberman M.D. (2013). The busy social brain: evidence for automaticity and control in the neural systems supporting social cognition and action understanding. *Psychol. Sci., 24*, 80–86.

Stokoe, W. C. (2002). Language in hand: Why sign came before speech. Washington, DC: Gallaudet University Press.

Tettamanti M. et al. (2005). Listening to action-related sentences activates frontoparietal motor circuits. *J. Cogn. Neurosci. 17*, 273–281.

Thompson E.L., Bird G., Catmur C. (2019a). Mirror neurons, action understanding and social interaction: implications for educational neuroscience. *Conference Abstract: 4th International Conference on Educational Neuroscience.*

Thompson, E. L., Bird, G., & Catmur, C. (2019b). Conceptualizing and testing action understanding. *Neuroscience & Biobehavioral Reviews, 105*, 106-114.

Tomasello, M. (2008). The origins of human communication. Cambridge, MA: MIT Press.

Wilson D., Sperber D. (2002). Relevance theory. G. Ward, L. Horn. *Handbook of Pragmatics,*

Blackwell.

Wolpert D.M., Kawato M. (1998). Multiple paired forward and inverse models for motor control. *Neural Networks, 11*(7), 1317-1329.

Wurm M.F., Ariani G., Greenlee M.W., Lingnau A. (2016). Decoding concrete and abstract action representations during explicit and implicit conceptual processing. *Cereb. Cortex, 26*, 3390–3401.

Wurm M.F., Lingnau A., (2015). Decoding actions at different levels of abstraction. *J. Neurosci., 35*, 7727–7735.

Zywiczynski, P., Gontier, N., & Wacewicz, S. (2017). The evolution of (proto-) language: Focus on mechanisms.

Angelo D. Delliponti is a graduate in Cognitive Science of Communication and Action, the master's degree course at the Department of Philosophy, Communication, and Entertainment of the Roma Tre University. He is a PhD student in Linguistics at Academia Copernicana, University Centre of Excellence IMSErt in Toruń. The title of his thesis was "Ostensive-inferential communication and embodied cognition". He worked in CosmicLab—a laboratory headed by Prof. Francesco Ferretti in Roma Tre University—on a behavioural experimental design aimed at investigating the comprehension of visuo-spatial metaphoric expressions in adults with visual impairments. Specifically, the expe-riment intended to evaluate whether a narrative context can facilitate metaphoric comprehension. He partecipated in the THUS 2021—Torun Humanities and Social Sciences Summer Programme—of the Nicolaus Copernicus University on experimental semiotics.