biolinguistics

volume 6
issue 3-4

summer-fall
2012

ISSN 1450-3417

★ EDITORIAL ★

Introducing Embodiment of Language

Kleanthes K. Grohmann

The old saying that writing a book is like giving birth to a baby (or an elephant) certainly captures the process of putting this special issue together. But in the end we made it and are happy to present a comprehensive collection that reviews and advances the emerging field of "embodiment of language".

To say a few words about the genesis of this special issue, we were first approached in September 2011 by a *Biolinguistics* editorial board member with the suggestion to put together a volume based on the 'Embodied Language' conference which took place at New College, Oxford University (26–28 September 2011, http://www.newcollegeembodiedlanguage.com). We were intrigued by the possibility, least of all because it does offer new ways of probing language in the species, but also because of the top-notch research on mirror neurons coming from the field. We are very excited about presenting some of that research and — through some review articles — its history.

The organizer of that conference, Dr. Robin M. Allott, was kind enough to get things going as guest editor. He invited the participants, and some colleagues beyond, to submit their full-fledged papers to *Biolinguistics* for a full peer-review process. It is then that things went somewhat awry. Not being specialists on the topic ourselves (and the same could be said for most of the reviewers we have on file), we relied on Dr. Allott's expertise to help us find suitable reviewers. To cut a long story short, we invited a total of 67 colleagues to serve as peer-reviewers for the collection, intending two different reviewers for each submission, of who 42 turned down our request or failed to respond. The entire process took considerably longer than planned, and we had to find a number of last-minute peer-reviewers ourselves as well. If that weren't enough, Dr. Allott pulled out of his envisioned guest editorship in the finishing stages of the project as well.

But all good things end well. And without further ado, we invite the reader on an enlightening journey through embodied language.

Kleanthes K. Grohmann
University of Cyprus
Department of English Studies
75 Kallipoleos
P.O. Box 20537
1678 Nicosia
Cyprus
kleanthi@ucy.ac.cy

ISSN 1450–341

Language: From Sensory Mapping to Cognitive Construct

Bernard H. Bichakjian

This paper places embodiment in an evolutionary perspective and endeavors to show that as incipient speakers began forging a linguistic system, they molded their grammatical distinctions and syntactic functions on their perception of the outside world, but that in the course of evolution, these perceptually-tinkered features were gradually replaced with mental constructs, specifically conceived to serve linguistic purposes and serve them with increased potentiality and greater efficiency. The shift from perceptual to conceptual implements is perhaps most conspicuously visible in writing, where open-ended figurative hieroglyphs were replaced with a small set of abstract letters, but the process is pervasive. In syntax, the phenomenal notion of agency, so deeply anchored in our activities, and the entire grammatical system built thereupon were replaced with a model where agency is irrelevant and syntax is structured on the purely mental constructs of subject and object. The paper continues with further cases of disembodiment.

Keywords: argument alignment; cognitive constructs; embodied language; language evolution; perceptual mapping

1. The Canonical View Questioned

For some fifty years the prevailing theory in mainstream linguistics was Chomsky's nativist hypothesis. It stated that in the course of their evolution humans became endowed with an innate linguistic model that enabled them initially to build grammars and thereafter learn in their early years the language spoken in their linguistic environments. Since this innate model was postulated to have genetic correlates much like "an organ such as the eye or heart" (Chomsky 1980: 37), and since, barring a major mutation, these genetic correlates would permanently remain the same, all languages — extent or extinct — were by way of corollary ruled to be gratuitous variants of one another. This theoretical framework meant in turn that while languages do undergo changes, those changes are gratuitous: "There is no more reason" stated Postal emphatically, "for languages to change than there is for automobiles to add fins one year and remove them the next" (1968: 283). Languages, therefore, do not evolve; they remain with neutral changes the external manifestation of a

ISSN 1450–341

permanent innate model.

Initially, the nativist theory and its expression in formal language exerted an undeniable fascination, but more and more it became apparent that the necessary empirical support was lacking (cf. Vargha-Khadem *et al*. 1995: 930; see also Vargha-Khadem *et al*. 1998: 12699). Today, with counter evidence mounting, the nativist theory is openly contested. In a seminal paper published in a journal that once hosted Pinker and Bloom's comprehensive presentation of the nativist tenets, Evans and Levinson argued that

> [t]he claims of Universal Grammar … are empirically false, unfalsifiable, or misleading in that they refer to tendencies rather than strict universals. Structural differences should instead be accepted for what they are, and integrated into a new approach to language and cognition that places diversity centre stage (2009:429).

Speaking of a follow-up paper (Dunn *et al*. 2011) published by a partially different team, but from the same research center, the lead author told BBC online (Apr. 15, 2011):

> We show that each of these [four] language families evolves according to its own set of rules, not according to a universal set of rules.
>
> That is inconsistent with the dominant 'universality theories' of grammar; it suggests rather that language is part of not a specialised module distinct from the rest of cognition, but more part of broad human cognitive skills.

The cumulative message of the two papers, both based on a vast survey of languages, is clear and unmistakable: There is no linguistic evidence for the existence of a universal grammar coded in our genes, and languages pursue their own individual evolutionary courses.

The Dunn *et al*. paper is by no means flawless, and suggestions have been made to improve the approach (cf. e.g., Longobardi & Roberts 2011), but it does display a rigorous methodological approach applied, admittedly, to only one feature, namely the shift from head-last to head-first word order, but conducted across no less than one third of the world's languages. Their conclusion is therefore well grounded: Languages evolve and they set their own evolutionary courses. Dunn and his colleagues' innovative paper marks an important step in the study of linguistic change, but it begs the next one: If changes are discussed in an evolutionary framework, the discussion must then invariably include an assessment of the selective advantages of the output over the input. Such comparative assessments are indispensable if we are to understand why such sweeping changes have taken place and/or are ongoing, why they normally are irreversible, and why some applications of a broad shift may take exception (cf. e.g., the situation in English, where modifying adjectives are head last in a predominantly head first language).

The application of evolutionary criteria to the study of language has also been my pioneering activity for decades. Without a computational apparatus, but on the basis of diachronic data carefully placed in their historical context and properly extrapolated material from language typology, I have been arguing that, under normal circumstances, languages proceed in the direction of ever-higher efficiency, upgrading the power of expression of their implements while reducing their neuromuscular cost. Each language pursues such a course on its own — at its own rate and along its own pathway. Since no new alternative has only advantages, and no new alternative is the only one to present advantages, it is understandable that each language makes its own choices and takes its own pathway towards greater efficiency (cf. Bichakjian 2002 for an elaborate presentation).

2. The Present Objective

In this paper, I will argue that one of the important ways of achieving higher efficiency has been the shift from linguistic features initially molded on the sensory mapping of the external world to cognitive alternatives especially conceived to serve linguistic purposes.

This observable trend in the history of languages ties in with the adaptive nature of language, whose phylogenetic acquisition provided humans with a cognitive dimension that enables us to elaborate knowledge "not only from sensory mappings that we share with other anthropoids as well as most mammals, but by important inputs to the mapping that comes from our language 'sense' as it has evolved in *Homo sapiens*" (Jerison 2001: 384).

The shift from perceptual to conceptual implements also ties in with the embodiment issue, and the observed linguistic process can be seen as a case of disembodiment. The units of measurement provide a clear illustration. The ancestral ones were generally based on the dimension of body parts — they were literally embodied: The inch was the standard width of the thumb, the foot the standard length of the eponymous organ; the Egyptian cubit represented the length of the forearm, and the yard that of the extended arm. With notable exceptions, these anthropomorphic units of measurement, molded on the perception of the outside world, have been replaced with the conceptually devised metric system, which has considerable selective advantages. The evolution of the units of measurement clearly illustrates the shift to and thence the evolutionary trend toward disembodiment, while the cases of resistance to the modern system reveal the clash and competition between visceral feelings and mental deliberation. In linguistics, the older quantitative and qualitative vowel alternations (cf. Lat. *ĕdimus* vs. *ēdimus* 'we eat *vs.* we ate' and Engl. *we break* vs. *we broke* anchored in our potential for rhythmicity would fall in the category of embodied features whereas the modern opposition based on mentally created auxiliaries would constitute a case of disembodiment.

3. From a Perceptual Beginning

Perhaps Athena burst forth from Jupiter's forehead fully armed, but that certainly was not the case of language, and the idea of treating language as an all-or-none entity is counterfactual. Language developed in the course of time through the speakers' unconscious, yet intuitively guided efforts. Linguists are unfortunately unable to reconstruct the utterances of incipient speakers, but they have access to sufficient data provided by internal reconstruction and typological surveys to trace the developmental trend. Incipient speakers started from scratch, but *ex nihilo, nihil fit*, from nothing, nothing comes. We all know that animals do not speak, but it is part of their survival strategy to observe and categorize the elements of the outside world and the activities taking place around them. So, when incipient speakers began cobbling a system of verbal communication, they brought to the task the knowledge and experience that were already theirs. Since that knowledge was essentially perceptual, the linguistic system that they first built was based on distinctions and functions molded on those observed in the outside world. These were gradually either abandoned when they proved unnecessary or replaced with mentally-constructed alternatives that provided greater efficiency.

The evolution of languages has therefore been a steady shift from perception- to conception-based grammatical distinctions and syntactic functions. This general trend can be observed in several important parts of language, but it is perhaps most conspicuous in the evolution of writing and the resulting development of the alphabet. It is true that the graphic representation of speech only plays an ancillary role, but its special illustrative value justifies its being included in the discussion before the focus is laid on the evolution of nouns, the development of adjectives, the realignment of arguments, the rise of temporal distinctions and the coining of marking devices.

3.1. The Evolution of Writing: From Pictograms to Letters

The evolution of writing is well known, and its course from perceptual pictograms to conceptual alphabetic letters is no secret. When they wanted to commit a word to a slab of stone, a clay tablet, or a papyrus scroll, scribes sketched the image of the referent, provided, of course, the referent was concrete. So, the outline of a snake, for example, represented the word *snake*, but also the words for items and attributes associated with snakes, such as *venom* and *perfidy*. As such, the pictograms functioned as ideograms — they were meant to be read as words, the word for the depicted item or those associated therewith.

But pictograms could also have a phonetic function, one with far reaching consequences for the history of writing. In the absence of the diacritic mark indicating that the image must be read semantically, pictograms could be read phonetically. An imaginary English example can illustrate the point. The word *tail* can easily be represented with the image of a tail, and when accompanied with the proper diacritic mark the pictogram will refer to the organ and its figurative and associative meanings. But without the diacritic mark, called determinative, the image simply refers to the sound of the word *tail* and as such could also be used to represent the less "photogenic" but like-sounding word *tale*. The pictogram for the word *tail* could also be combined in a rebus with the image of an *oar* to form the hieroglyph of the word *tailor*.

It is this phonetic use of pictograms that led to the development of alphabetic letters. It occurred over time through a two-track process, one mental, the other graphic. Mentally, the initial consonant was extracted out of the continuous flow of speech sounds associated with the word and recognized as a specific entity, an entity without a semantic backing, but an entity, nevertheless. Graphically, the stylized and simplified form of the image of the full word became the sign of that abstract entity, the sign of a speech sound. It took more than two millennia for the Greek alphabet to evolve out of the Egyptian hieroglyphs. The process was admittedly slow, but writing systems did come up with conceptual alternatives for the initially created perceptual implements, and it was well worth the time and effort because an open-ended array of 700 hieroglyphs was replaced with a set of less than fifty characters that can code for an infinite number of words and that can, barring accidental cases of homophony, do so unambiguously. Writing systems became fully accurate and considerably more efficient. It does require a formal learning process, but considering the yield, it is well worth the effort (for a less cursory presentation of the evolution of writing systems, cf. Bichakjian 2002: 221–258; see also the classic works of Diringer 1948 and Gaur 1984).

3.2. Noun Classes

Since we cannot reconstruct the utterances of incipient speakers, it is impossible to tell with certainty whether the first nouns were subcategorized in classes (human, animal, vegetal, solid, liquid, long, compact, etc.), but the existence of such grammatical distinctions in aboriginal languages, the survival of active/stative doublets such as the Germ. *Wasser* 'water', neuter, and the Fr. *eau* (< Lat. *aqua*) 'water', feminine and the fact that in the Indo-European languages neuter nouns are unmarked in the nominative suggest that the prototypical vernaculars probably subcategorized nouns according to the physical characteristics they attributed to their referents. The fact that the ancestral language had two words for 'water' — one neuter, i.e. inanimate, another feminine, i.e. animate — suggests that the ancestral speakers had a dual representation of 'water'. In one perspective, they perceived 'water' as a material item and assigned it to the class of inanimate nouns, in the other as an entity endowed with cleansing virtues or calamitous powers and assigned it to the class of animate nouns. The absence of nominative markers of neuter nouns suggests that their etyma in an earlier language belonged to the class of nouns that could never appear in the agentive case — the forerunner of the nominative case — because their referents, like that of Germ. *Wasser* could never be the agent of an action (cf. *inter alia* Ashton *et al.* 1954 for an example of a language with noun classes, Meillet 1965: 219–220 on doublets, Diakonoff 1965: 55–56 for class evidence in Afro-Asiatic, and Schmidt 1979: 337 *et seq.* on neuter nouns).

The class distinction has not been completely eliminated everywhere; it often survives in the form of grammatical gender. The lineage is not complete and continuous, but one may reasonably surmise that incipient speakers built their grammar with distinctions observed and experienced in the outside world. But with speech making successive generations of speakers capable of greater abstraction, these perceptual markers proved to be redundant and thereby more taxing than informative. Classes gave way to grammatical genders, which in turn were gradually reduced or eliminated altogether as in the case of English, but also Armenian, Bengali, Chinese, and many other languages.

3.3. *The Adjectival Gap*

While the subcategorization of nouns in classes such as human, animal, vegetal, solid, liquid, long, compact, etc. was a likely feature of incipient speech, the distinction between active and inert or stative nouns was fundamental. Nouns, like their referents could be active or stative. In Latin the words for 'hand', 'foot', and 'tongue' were masculine or feminine, i.e. originally active since these are active organs, while the words 'head', 'heart', and 'liver' were neuter, i.e. originally stative since they were considered to be the seats, respectively, of intelligence, memory, and emotions. The active/stative distinction also applied to verbs: In Latin, 'to kill' was an active verb since it implies an activity on the part of the agent; 'to die' was a deponent since the experiencer is the seat of the action, not its author. So, originally, or at least in very ancient times, verbs were subcategorized into verbs of action and verbs of state, and since being white or red was a state, the characteristics that are expressed with adjectives in modern languages were expressed then with verbs of state (cf. Klimov 1979: 328 and fossilized tokens such as Lat. *albeo* 'to be white', *rubeo* 'to be red'). It stands to reason that the incipient speakers' first task was to label objects and coin words for actions and states. Conceptualizing quality and developing adjectives to express it came about later — the new part of speech needed greater mental application and higher abstraction.

3.4. *Argument Alignment from Agent/Patient to Subject/Object*

The subcategorization of nouns and verbs in active and stative classes is directly linked to how arguments were aligned in ancient times. The system was based on the incipient speakers' observation of events in the outside world. When narrating the event of a hunter killing an antelope, *hunter* would be in the agentive case and so marked, while *antelope* would be in the "patientive" case and left unmarked. Since the "patientive" form is the unmarked or bare form of the noun, the "patientive" case has been called the "absolutive" case, while the agentive case has been dubbed "ergative" (< Gk. *ergon* 'deed, action'). However, if an old man was there, his witnessing the scene of the hunting would be expressed with a stative verb, and *old man* would appear in the absolutive case, not in the ergative one, because a witness is not an agent. Both the old man and the antelope are considered to be part of the scene, not one of the movers — hence their being treated as patients and put in the absolutive case.

No exceptionally deep insight is needed to see that the incipient argument alignment was based on the perception and interpretation of events in the outside world. Since actions have agents and optionally one or more patients, their roles were carried over into grammar and made into syntactic functions. It should also be observed how important agency was considered to be. Not only *dying, witnessing,* and the like are not activities and the one involved does not qualify as an agent, but the fact that it is the ergative form that carries a morphological marker indicates how important agency was — in real life and in grammar.

Gradually this perceptual model morphed into the conceptual one of the nominative languages, where the active/stative and agent/patient dichotomies borrowed from the outside world no longer play a role. The key players of the new model are subject and object, mentally constructed functions that make it possible for any noun to be the subject of any verb and for speakers to describe actions in all perspectives — the agent's, the patient's and even the beneficiary's. Cf.

(1) The hunter showed the antelope to the old man.

(2) The antelope was shown by the hunter to the old man.

(3) The old man was shown the antelope by the hunter.

The shift from perceptual to conceptual grammatical functions has thus made grammar more flexible and more powerful.

3.5. Breaking the Bonds of the Present

While the thoughts of incipient speakers no doubt wandered about the experiences they had had in the past or those they foresaw in the future, speaking always took place in the present, and, in the present, actions were either ongoing or completed or possibly resulting from previous actions (cf. Germ. *wissen* 'to know', akin to Lat. *videō,* meaning originally 'to know for having seen').Those were the ancient verbal distinctions — all of them in the present. They were not temporal, but aspectual, because initial speakers expressed what they beheld in the material world at the time they spoke. The three ancestral aspects were called *present, aorist,* and *perfect* in the traditional terminology; *imperfective, perfective,* and *stative* are the preferred labels today.

As languages evolved, aspectual distinctions did not disappear altogether. English, for instance, makes an **aspectual** distinction — aorist *vs*. perfect, or perfective *vs*. stative — between *I ate* and *I have eaten*. Likewise, but somewhat differently, French and the romance languages in general make a distinction between the perfective and imperfective aspects of the past tense as in *j'ai mangé* vs. *je mangeais* 'I ate *vs*. I was eating'. Aspect is no doubt a useful distinction and that's why it has been partially preserved, but one should not lose sight of the fact that as languages evolved aspectual systems as a whole have morphed into temporal systems (on the ancestry of the aspectual system and its shift to a temporal one, cf. *inter alia* Meillet 1928: xii and Kuryłowicz 1964: 130). Just as motion picture cameras provide a kaleidoscopic view of events, temporal systems enable speakers to travel through time, narrate events of a foregone past, and structure those of a yet-to-occur future. Developing such a verbal system required a mental effort, a far greater effort than witnessing that an action is ongoing or completed.

3.6. Grammatical Marking: From Modulation to Free Morphemes

Let us imagine the following predicament of incipient speakers. They have coined nouns for objects and verbs for actions and states, but how are they going to make a distinction between an on-going action and one that is completed, one that is performed in one go and one that is performed repeatedly and in reduced form, or one performed normally and the same one with especial intensity. The original impulse seems to have been some form of stem modulation: Either the quality or the quantity of the root vowel was changed or the initial syllable was repeated — along with the change of quantity was at times a concomitant change of quality while stem reduplication often triggered a vowel reduction in the added syllable and occurred at times along with a vowel change in the stem (cf. e.g., Gk. *d**é**rkomai/déd**o**rka* I see/I have seen', Lat. *v**ĕ**nimus/v**ē**nimus* 'we come/we came', and *f**a**cimus/f**ē**cimus*' we make/we made', *pendimus/**pe**pendimus* 'we ponder/we pondered', *currimus/**ce**currimus* later by assimilation *cucurrimus* 'we run/we ran', and *c**a**nimus/**ce**c**i**nimus* 'we sing/we sang'. Reduplication was also used elsewhere as in *quisquis* 'whoever' lit. 'who-who', *quōquō* 'wherever' lit. 'where-where' and *alter alterum* 'each other' lit. 'other of two-other of two'. Outside the Indo-European family, reduplication can also serve to express the intensive or iterative forms of verbs (cf. Arabic *ka**s**ara* 'he or she broke' vs. *ka**ss**ara* 'he or she smashed to bits' and Swahili *piga* 'to strike' vs. *piga**piga*** 'to strike repeatedly') or even the plural as in Malay *rumah* 'house' ***rumah**-rumah* 'houses'.

While some cases of vowel alternation have survived, especially in the Germanic languages (cf. Engl. *sing/sang/sung*) and reduplication can occur today in baby talk and pet names (cf. Engl. *itsy bitsy*, Fr. Riri < Henri), stem modulation has generally been sidelined in the course of evolution and replaced with suffixes and infixes, which in turn have been partially replaced with particles and full-fledged words serving grammatical functions.

Stem modulation is not a feature molded on occurrences taking place in the outside world, but there is something physical or even visceral and echolalic about alternations and reduplications. These are indeed embodied linguistic features. The first task of incipient speakers was understandably to label items and actions and states; they then bent and remodeled these contents words to form their paradigmatic variants. But with the use of language stimulating greater abstraction and increased use of mental power, over the years, speakers developed a broad range of specific morphological segments with grammatical functions as referents. The following step was having full-fledged grammatical words next to the first coined contents words. That was a major step in the conceptualization of grammatical implements. Function words are sometimes called "empty" words. The adjective *empty* contrasts with *contents*, but it also reveals the level of abstraction and the mental effort that is required to coin words that have "no" contents. The three steps can be observed in the following sequence from Early Latin to Modern French, where stem modulation is successively replaced first with suffixation, later with the use of an auxiliary.

(4) cano/**ce**c**i**ni > canto/cant**avi** > je chante/j'**ai** chanté
'I sing/I sung'

These evolutionary steps, which constitute a case of disembodiment, took place because each new alternative had selective advantages that the ancestral one did not have. Stem modulation has a certain charm, echolalic or cadential, but such processes can only provide a limited number of distinctions, while suffixation offers unlimited possibilities. Suffixes are indeed open ended and as such more advantageous, but they have their own downside: They can trigger morphological irregularities and thereby create language acquisition problems and delays (cf. Slobin, 1971: 347 on the difficulty of acquiring flectional systems). Function words have no quantitative restrictions; they are easy to acquire and powerful to operate. While the shift to mentally-generated grammatical markers and the corollary disembodiment were driven by the pursuit of greater efficiency, it should be born in mind that no feature is exclusively advantageous, nor exclusively deleterious. Mentally generated linguistic implements have indeed the selective advantage of being more efficient, but embodied ones can also have theirs, such as the charm of reduplicatives in hypocoristics and nursery rhymes (e.g. *The Incy Wincy Spider*) and the more subtle pleasure of alliteration in adult language.

4. The Evolving Instrument of a cerebral species

It is a trivial observation that we humans do not have the tigers' fangs, the antelopes' speed, the eagles' wings, or even the turtles' shield or the elephants' mass. We have none of the weapons and none of the defenses that other animals have, but we have a major trump card: We have, relative to our size, an exceptionally large and highly-developed brain (cf. *mutatis mutandis* Gould 1977: 402). Brain power is our most valuable asset. We do not have fangs, but we have invented fire arms to hunt with; we cannot run as fast as antelopes, but we have engineered automobiles that transport us even faster; we cannot fly, but we have built aircrafts that make air travel not only possible, but fast and effortless.

Likewise, humans have started with rudimentary linguistic implements and developed ever-more efficient alternatives from agent-patient to subject-object argument alignment, from verbs of state to adjectives, from stem modulation to suffixation and thence to an array of free grammatical morphemes such as pronouns, prepositions, auxiliaries, adverbs, and articles, and also from glottal and glottalized consonants, which involve an "intricate coordination of the actions of the larynx with the actions of the articulators in the mouth," to simple oral consonants (Maddieson 2011).

Unlike all other species, humans are the only ones that have endeavored to find mentally generated and, as such, ever more efficient alternatives to the physical resources that are part of our endowments or immediately available in the outside world. The shift that was observed in language and also elsewhere from prototypes molded on the perception of objects and actions around us to mentally designed alternatives is therefore part of a truly human strategy.

It is our cerebral nature that explains the developments that were discussed in the foregoing, and they in turn support and confirm the view that language is not an instinct or a steady-state attribute coded in our genes, an organ as it was once claimed (Chomsky 1980: 37), but an instrument that keeps evolving — becoming ever more cerebral and, by so doing, ever more efficient.

Glossary of some of the technical terms used in this paper

Argument alignment. The patterning of elements in a sentence. The **nominative** (also called **accusative**) **alignment** is that of the nearly universal model composed, independently of their order, of a verb, its subject, and optionally one or more objects. The much less common **ergative alignment** is a model based on the distinction between agent and patient, and active and stative verbs. Depending upon whether the action has an actual author or not, the verb will be active and combine with an agent in the ergative case and eventually a patient in the absolutive case, or stative and will combine simply with a patient in the absolutive case.

Aorist. The traditional word for the perfective aspect, which denotes an action apprehended in its completion. Also used as a temporal distinction to denote a point action in the past without resultative connotations. Cf. e.g., *I ate* as opposed to *I have eaten,* which has resultative connotations meaning 'I am full', 'I don't have to eat', etc.

Aspect. Whereas temporal distinctions are about the relative time of the action, namely past, present, and future, as in *I see, I saw, I shall see,* aspectual distinctions apply to the flow of the action. The main aspects are imperfective, perfective, and perfect, as in *I was eating, I ate, I have eaten.* Pedagogical grammars all too often conflate tense and aspect, but seen properly, the distinction between *I see* and *I saw* is temporal, between *I saw* and *I have seen* is aspectual.

Echolalia. Feature of two-syllable words where the second is much like the echo of the first. Cf. e.g., *bye-bye*, *oink-oink*, *tic-tac*.

Embodied language. Items of language possibly shaped by aspects of the human body.

Marker. A morphological device — often a suffix — used to indicate a paradigmatic variant. In English, *–s* is the regular plural marker of nouns, while *–ed* is the past tense marker of regular English verbs.

Marked/unmarked. These have two meanings:

They may simply refer to the presence or absence of a marker. In English, plural nouns are marked, but plural adjectives are unmarked.

In the **theory of markedness, marked** means a step or steps away from the basic or most natural variant. The vowel [ɛ] as in French *mère* 'mother' is unmarked — it is the most natural of all vowels — but the French vowel [œ], as in *sœur* 'sister', is marked for roundness because the spontaneous pronunciation of a front vowel is without lip rounding. The lip rounding is achieved through an extra effort— hence the marking. The vowel [œ̃], as in French *brun* 'brown', is doubly marked since it requires two extra efforts: One for lip rounding, one for nasalization.

Universal Grammar is the set of the structural properties common to all natural languages claimed to be hard-wired into the human brain.

References

Ashton, Ethel O. et al. 1954. *A Luganda Grammar.* London: Longmans, Green & Co.

Bichakjian, Bernard H. 1999. Language evolution and the Complexity Criterion. Target article. *Psycoloquy* [http://www.cogsci.ecs.soton.ac.uk/cgi/psyc/newpsy?10.33]

Bichakjian, Bernard H. 2002. *Language in a Darwinian Perspective*. Frankfurt: Peter Lang.

Chomsky, Noam. 1980. On Cognitive Structures and their Development: A Reply to Piaget. In Massimo Piattelli-Palmarini (ed.), *Language and Language Learning. The Debate between Jean Piaget and Noam Chomsky*, 35–52. London: Routledge and Kegan Paul.

Diakonoff, Igor M. 1965. *Semito-Hamitic Languages. An Essay in Classification*. Moscow: Nauka.

Diringer, David. 1948. *The Alphabet: A Key to the History of Mankind*. London: Hutchinson's Scientific and Technical Publication.

Dunn, Michael, Simon Greenhill, Stephen Levinson & Russell Gray. 2011. Evolved structure of language shows lineage-specific trends in word-order universals. *Nature* 473, 79–82.

Evans, Nicholas & Stephen C. Levinson. 2009. The Myth of Language Universals: Language Diversity and its Importance for Cognitive Science. *Behavioral and Brain Sciences* 32, 429–492.

Gaur, Albertine. 1984. *A History of Writing*. London: The British Library.

Gould, Stephen J. 1977. *Ontogeny and Phylogeny*. Cambridge, MA: Harvard University Press.

Jerison, Harry J. 2001. Adaptation and preadaptation in hominid evolution. In Phillip V. Tobias *et al.* (eds.), *Humanity from African Naissance to Coming Millennia*, 381-86. Florence: Florence University Press & Johannesburg, South Africa: Witwatersrand University Press.

Klimov, Georgij A. 1979. On the position of the ergative type in typological classification. In Frans Planck (ed.), *Ergativity. Towards a Theory of Grammatical Relations*, 327–332. London: Academic Press.

Kuryłowicz, Jerzy. 1964. *The Inflectional Categories of Indo-European*. Heidelberg: Winter.

Longobardi, Giuseppe & Ian Roberts. 2011. Non-arguments about non-Universals. *Linguistic Typology* 15, 483–495.

Maddieson Ian. 2011. Glottalized consonants. *The World Atlas of Language Structures Online*. [Http://wals.info/chapter/7]

Meillet, Antoine. 1928. *Esquisse d'une Histoire de la Langue Latine*, 6th edn. Paris: Hachette.

Meillet, Antoine. 1965. La catégorie du genre et les conceptions indo-européennes. In *Linguistique historique et linguistique générale*, 2nd edn, 211–229. Paris: Champion.

Pinker, Steven & Paul Bloom. 1990. Natural language and natural selection. *Behavioral and Brain Sciences* 13, 707–784.

Postal, Paul. 1968. *Aspects of Phonological Theory*. New York: Harper & Row.

Schmidt, Karl H. 1979. Reconstructing active and ergative stages of pre-proto-Indo-European. In Frans Plank (ed.), *Ergativity: Towards a Theory of Grammatical Relations*, 333–345. London: Academic Press.

Slobin, Dan I. 1971. Grammatical development in Russian-speaking children. In Aaron Bar-Adon & Werner F. Leopold (eds.), *Child Language: A Book of Readings*, 344–348. Englewood Cliffs, NJ: Prentice-Hall.

Vargha-Khadem, Faraneh, Kate Watkins, Katie Alcock, Paul Fletcher & Richard Passingham. 1995. Praxic and nonverbal cognitive deficits in a large family with a genetically transmitted speech and language disorder. *PNAS USA* 92, 930–933.

Vargha-Khadem, Faraneh, K. E. Watkins, C. J. Price, J. Ashburner, K. J. Alcock, A. Connelly, R. S. J. Frackowiak, K. J. Friston, M. E. Pembrey, M. Mishkin, D. G. Gadian & R. E. Passingham. 1998. Neural basis of an inherited speech and language disorder. *PNAS USA* 95, 12695–12700.

Bernard H. Bichakjian
Prof. van der Grintenstraat 3
6524 RG Nijmegen
The Netherlands

BHB@Post.Harvard.edu

Is Embodiment All That We Need?
Insights from the Acquisition of Negation

Valentina Cuccio

Simulation of propositional content does not sufficiently explain real-life linguistic activity, even for action-related language. In addition, how we get from propositional content to implicit and inferential meaning needs to be explained. Indeed, simulative understanding is immediate, automatic and reflex-like while an explicit interpretative act, even if not always needed, is still a part of many linguistic activities. The aim of this paper is to present the hypothesis that speaking is a complex ability realized by means of at least two different mechanisms that are likely developed at different and consecutive steps of cognitive and linguistic development. The first mechanism has a neural explanation grounded in the notion of embodied simulation. The second implies socio-cognitive skills such as Theory of Mind. In order to fully develop the second mechanism, a symbolic communication and interaction with a cultural community are needed. This hypothesis will be tested by looking at the acquisition of linguistic negation.

Keywords: embodied language; inferential meaning; mindreading; negation

1. Introduction

> There is an assumption, popular among philosophers, that the brain processes that make for cognition are one sort of thing and that the brain processes that contribute to motor control belong to an entirely different category.
>
> (Churchland 1986: 451)

It is interesting to note that philosophers such as Epicuro, Campanella, Vico, and Condillac all held the hypothesis that cognition is deeply grounded in our corporal and motor experience even though they did not have the knowledge of neurobiology that we have today.

However, since Patricia Churchland's book came out in the late eighties, experimental data (especially the discovery of mirror neurons in primates; Gallese *et al*. 1996; Rizzolatti *et al*. 2001, 1996) has confirmed the hypothesis that cognition and language are embodied. Consequentially, philosophers who still considered brain processes for cognition as radically different from processes for motor control have moved away from that bias. However, the embodiment of language and cognition might not fully explain human cognitive and linguistic

biolinguistics

ISSN 1450–341
Biolinguistics 6.3–4: 259–275, 2012
http://www.biolinguistics.eu

abilities. The aim of this paper is to present the hypothesis that speaking is a complex ability realized by means of at least two different mechanisms that are likely developed at different and consecutive steps of linguistic and cognitive development. The first mechanism has a neural explanation grounded in embodied semantics and mental simulation (Gallese & Lakoff 2005). The second implies socio-cognitive skills such as Theory of Mind. In order to fully develop the second mechanism, a symbolic communication and interaction with a cultural community are needed. Hence, the origin of man's complex species-specific language and cognition is in both the brain and in culture (Deacon 1997; Tomasello 1999; Arbib 2009). This hypothesis will be tested by looking at the acquisition of linguistic negation.

2. The Acquisition of Negation in Normal Development

In first-language learning three broad categories of negation consecutively arise (see Dimroth 2010 for a review): (1) rejection/refusal, (2) disappearance/non-existence/unfulfilled expectation, (3) denial. According to many studies (Volterra & Antinucci 1979; Pea 1980; Choi 1988), rejection is the first category of negation to be acquired. Children use "no" to express refusal of something existing in their present context. However, we can find examples of rejection in human pre-linguistic gestures and even in animal behaviour. In fact, before the time children start to produce the single word "no" to express rejection, they have already expressed rejection non-linguistically. Rejection, according to Pea (1980) does not require abstract mental representations, while non-existence and denial do require them.

The second category of linguistic negation to arise is non-existence/unfulfilled expectation. At this point, children are able to signal the absence or disappearance of an expected referent in the context of speech or to indicate something that violates their expectations, based on previous experience (for instance, malfunctioning toys).

Lastly, the third category to be acquired is denial. Denial implies negation of a predication. The referent is usually symbolically expressed. As Bloom (1970) argues, to deny children must have the ability to discern between their own knowledge of the world and the knowledge of their listener. In order to deny a sentence, children have to deal with two propositions, one affirming and one negating the same predication; and they have to ascribe one of these to the person they are speaking to. "To deny the truth of another person's statement entails the understanding that the other person may hold different beliefs, or that language is itself a representation of reality, not reality itself" (Tager-Flusberg 1999: 328). Denial is usually acquired by the age of two and a half years.

According to Antinucci & Volterra (1979) categories of negation are acquired according to the complexity of the inferences that they entail. At the beginning of this process, children are only able to make inferences about their present perceptual situation. Thus, at first, children can only negate (rejecting, prohibiting, or expressing non-existence) something currently present in the perceptual context of speech or something that was just present in the speech context. Later on, as children start to express denial, they become able to read their listeners presuppositions. At this time, children rely both on perceptual and pragmatic context.

Choi (1988), in her longitudinal study on English-, French-, and Korean-speaking children aged between 1;7 and 3;4 (years;months), went into more depth in the description of semantic categories of negation. She identified nine functions of negation that usually arise in three different phases.

(1) a. Phase 1: non-existence, prohibition, rejection, failure.
  b. Phase 2: denial, inability, epistemic negation.
  c. Phase 3: normative negation, inferential negation.
(Choi 1988: 525)

In Phase 3, Choi introduced the category of inferential negation that indicates "the child's inferences about the listener" (Choi 1988: 524). She refers to an interesting example of inferential negation:

(2) (Kyle has broken a few crayons. The experimenter has been scolding Kyle for breaking crayons. Kyle picks up a broken crayon which he did not break and looks at the experimenter.)
K: I not broken this.
(Choi 1988: 525)

This example clearly requires false belief understanding. In fact, when Kyle says "I not broken this", he is reasoning with the experimenter's belief. Specifically, Kyle supposes that the experimenter believes that he has broken that crayon while this was not the case (false belief). This example of inferential negation was recorded when Kyle was 2;8.

After the one-word utterance period, when children start to utter their first sentences, according to Bloom's (1970) study, non-existence is the first category of negative sentences to arise, not rejection. Denial is still the last category to be acquired, despite the fact that the syntactic structure of denial is less complex than non-existence. According to Bloom, denial requires more cognitive effort from children.

Hence, according to the data we have seen thus far, we can say, following Pea (1980: 165) that the expression of negation, from simplex to complex forms, requires underlying cognitive representations of increasing complexity. The first expression of negation does not require internal abstract representation because the rejected object is present in the perceptual scene; later on, with the expression of a disappearance, abstract mental representation is required because the negated object or person is no longer present in the speech event context; finally, when truth-functional negation is used to deny a predication, a second-order abstract representation is required.

Nevertheless, Pea (1980) does not agree with the opinion of Antinucci & Volterra (1979) that, in order to deny, children must attribute a presupposition to their listener. Many times, he argues, children express negation without addressing a person. Moreover, he claims, there is not enough independent experimental evidence supporting Antinucci & Volterra's assertion that two-year-olds are able to infer other people's mental states.

However, we know that the development of Theory of Mind begins early. During the second year of life, for example, language acquisition seems to heavily rely on the ability of reading other people's intentions (De Villiers 2000). Furthermore, children start to use mental verbs like *think*, *know*, etc. in their third year of life (i.e. before they are able to pass the false belief task). Currently, we also have evidence showing that even 15-month-old children can understand false beliefs (Onishi & Baillargeon 2005).

Moreover, it is worth noting that linguistic negation, at least denial, is a metalinguistic operator. Negation cannot be referentially used. Even in a negative descriptive sentence (e.g., *It is not raining*.), negation does not have its own referent. Negation never concerns a fact in a real or fictional world, nor an abstract concept like "elegance" or "rationality", nor an action. Negation is metalinguistic because it implies an operation on a proposition. That is to say, negation is the operation of setting a false value for the proposition it is referring to (and this is the same operation we also make in the more complex linguistic action of lying). Thus, negation, or at least denial, seems to imply a second order mental representation. In fact, by expressing a denial toward a listener, the child is representing a content, negating that content (by setting a false value) and attributing the negated content to the listener.

Thus, following the considerations of Antinucci & Volterra (1988), we can argue that the complex forms of negation, which are metalinguistic in their nature, regardless of what different names they may be called, entail the ability to attribute mental states to others and even to understand false beliefs.

Thus, psycholinguistic studies on linguistic negation show that negation can express very different functions that arise at different steps of language acquisition and that have different cognitive requirements. In what follows, it will be argued that we need two different mechanisms in order to account for the acquisition of these different functions of linguistic negation.

3. What Model of Linguistic Development can Account for the Acquisition of Negation?

So far, we have seen that there are at least three steps of increasing complexity in the acquisition of linguistic negation. The question that will be addressed now is: What model of linguistic development can account for the acquisition of negation?

A two-level model for language acquisition will be presented. In particular, the hypothesis that we can account for the first categories of negation, rejection and non-existence (the first step of language competence in this model) and even for the comprehension of the negated content of a denied sentence in a simulative paradigm will be proposed. Embodied simulation, however, does not sufficiently explain denial (the second step of language competence). While in a simulative account, based on motor simulation, our understanding of others is "immediate, automatic and almost reflex-like" (Gallese 2007), and our comprehension of action-related language is realized by means of a neural simulation of the pertinent action, as neuro-imaging and Transcranic Magnetic Stimulation studies have widely shown (Buccino *et al*. 2005; Tettamanti *et al*. 2005), an explicit interpretative act is needed in order to explicitly attribute negated/simulated information to someone else. Hence, simulation of the propositional content is not enough to explain denial. We need to introduce more complex inferential abilities that are not only the product of our brain, but also of our interaction with a cultural community.

The latter claim about the explicitness of denial could be challenged by saying that there is not any need to envisage an explicit act in order to attribute the negated information to someone else. Indeed, recently Gallese & Sinigaglia (2011) have argued that by means of embodied simulation a given content can be implicitly attributed to others because embodied simulation entails a functional, non-representational form of mental state attribution. However, what characterizes linguistic denial is exactly the explicit negation of someone else's mental state. Thus, in this case by accepting the claim that an explicit attribution is not necessary to denial because an implicit functional attribution of mental states can be realized by means of embodied simulation, it would follow that when denying we are explicitly negating a content that is implicitly attributed. And this means that, at some point, that content should become explicitly accessible. Then again, even if we buy into Gallese & Sinigaglia’s proposal about the functional and non-representational nature of mental state attribution realized by means of embodied simulation, we still need to introduce an inferential process that acts on explicit accessible content. This, obviously, does not happen for every inferential activity in which we are involved. But it is necessary for some human activities, and denial seems to be among them.

3.1. Simulating Negation

According to many authors (Barsalou 1999; Pulvermüller 1999, 2002; Gallese & Lakoff 2005; Gallese 2008), linguistic meaning is embodied. This means that the comprehension of an action-related word or sentence activates the same neural structures that enable the execution of that action. Gallese (2008) presented this hypothesis as the “neural exploitation hypothesis”. Language exploits the same brain circuits as action. According to this hypothesis, our linguistic and social abilities are grounded in our sensory-motor system. The Mirror Neurons System (MNS) is the neural structure that supports both our motor abilities and our social skills, language included. Thus, in this account, actions and language comprehension are mediated by motor simulation. This even holds true for the understanding of abstract linguistic meanings. Indeed, in that case, *mental imagery* and *metaphorical thought* allow us to map from a sensory-motor domain to an abstract domain. This mechanism, according to Gallese & Lakoff (2005), is the basis for the construction and comprehension of abstract linguistic meaning.

Moreover, it has been pointed out that language comprehension also relies on other kinds of embodied simulation, not only on motor simulation. For example, Engelen *et al*. (2011) elaborated and tested a model for the development of language comprehension in which perceptual simulation is central. According to these authors, a perceptual simulation of an event is created every time we listen to or read a description of that event. In this view, children exploit the process of perceptual simulation even if they have a limited knowledge of the objects involved. The role of sensory simulation in language comprehension has also been highlighted by Wojciehowski & Gallese (2011). Their approach focuses on the idea of the "Feeling of the Body" that is the product of sensory-motor simulation processes that enable us to immediately understand not only basic motor intensions but also the feelings and emotions of others. The experience of the "Feeling of the Body", in the authors' view, is fundamental when we approach literary texts and explains the sense of identification with and connectedness to narrated characters that we feel while reading stories. Recently, also further research conducted on the role of embodied emotion and introspection in language comprehension (Barsalou *et al.* 2003, Kousta *et al.* 2011) has contributed to giving us a richer and more powerful picture of the embodied language studies.

According to this view, many features of language seem to be bodily grounded. Lately many studies have also been devoted to the understanding of simulation processes in the comprehension of negated sentences (for example, Kaup et *al.* 2006; Tettamanti *et al.* 2008). To date, the neural mechanisms that underlie the processing of negation are still unclear. Tettamanti and colleagues carried out an experiment that aimed to identify the brain mechanisms that underlie the comprehension of negative sentences.

As Tettamanti *et al.* (2008) noted, psycholinguistic data on the comprehension of negative utterances seems to lead to different conclusions. On one hand, data from sentence comprehension (for example, Carpenter & Just 1975) suggests that a negative sentence is cognitively more demanding than the corresponding affirmative sentence. And this should indicate a stronger neural activation for a negative sentence than for its affirmative counterpart. On the other hand, data from studies on the accessibility of information (for example, Kaup & Zwaan 2003) suggest that negated information is cognitively less accessible.

The main aim of the Tettamanti *et al.* study was to test whether the impact of negation on neural activation is dependent on the semantic field of the negated content. Considering previous results, two possible outcomes were predicted by the authors. Negation could determine a reduced accessibility of the content of the negated sentence with, consequently, a lower neural activation for simulation in the left fronto-parieto-temporal and in the posterior cingulate cortex (content dependent hypothesis). Alternatively, negation could determine higher cognitive loads, due to a greater syntactic complexity. In this case, the processing of negation will lead to a higher, content independent, neural activation in the left perisylvian areas.

Results of Tettamanti *et al.*'s (2008) study confirmed previous findings on the accessibility of information (Kaup & Zwaan 2003). The authors say that "Negation is encoded by our brain in terms of a reduced activation of the areas representing the negated information".

Thus, the Tettamanti *et al.* study seems to be congruent with the hypothesis proposed by Kaup & Zwaan (2003, 2007) that the processing of negation enables a lower neural activation of the negated content in the brain, both of action-related and abstract content.

For the sake of clarity it should be noted that Kaup & Zwaan (2007) proposed a *two-step simulation hypothesis of negation*. According to their model, the comprehension of a negative sentence is realized firstly by means of a simulation of the negated content that is soon followed by a brief simulation of the actual state of affairs. At this point, information about the negated content is less accessible, this finding is congruent with Tettamanti *et al.*'s (2008) results. Moreover, this hypothesis can even account for data from sentence comprehension tasks (for example Carpenter & Just 1975). Indeed, in the first step of the comprehension of a negative sentence, the comprehender has to manage with two simulations, i.e. the simulation of the negated content and the simulation of the actual state of affairs. Hence, it is plausible that in this first step, a negative sentence is cognitively more demanding resulting in a longer response time.

The two-step simulation hypothesis was tested by means of a set of experiments that assessed time responses in recognition tasks. Participants read a sentence and then were shown a picture. They had to decide whether objects in the picture matched objects mentioned in the sentence. Responses were faster when depicted objects matched those mentioned in the sentences, even for negated contents. Thus, the same response-time pattern was found both for affirmative and negative sentences. In a different condition, pictures were shown with a longer delay. In this case, different response-time patterns were observed for negative and affirmative sentences, with negative sentences showing significantly slower responses. Thus, the results seemed to confirm the two-step simulation hypothesis, with less accessible information about the negated content in the second step.

It should be noted that in the studies on the simulative processes of negative sentences, negation is considered to be a monolithic syntactic function. The authors did not consider that different functions, with different cognitive and pragmatic demands, can be carried out in language by the same morpheme. Thus, to simulate the content, both the actual or the negated content of a sentence, does not supply a complete explanation of the comprehension of negation. "No" can express a rejection or a denial and the cognitive distance from the former to the latter is considerable. Moreover, it should be noted that Pragmatics usually classifies two different kinds of negation: descriptive and metalinguistic negation (Horn 1985). Descriptive negation refers to a state of affairs in the world while metalinguistic negation acts on presuppositions, implicatures and formal aspects of language like morphology or phonology. According to Kaup and colleagues, "the experiential view conceptualizes language comprehension as the performance of a sensory-motor simulation of the described sequence of events" (Kaup *et al.* 2007: 265). Hence, these studies seem to only be focused on descriptive negation. Language, however, is not always just the description of an abstract or action related state of events. Even if a simulation of the content occurs, and this seems to be the case, this does not sufficiently explain the complex inferential process that occurs in linguistic denial. In real life, in many cases, by denying a sentence, a speaker is denying presuppositions or implicatures of his co-speaker.

From the point of view of the speaker, this entails the ability of attributing mental states to the others and of following their inferences. The speaker is explicitly attributing a mental content to the interlocutor and he is then negating that content. Thus, the speaker is holding the actual and the negated content of the inference in his mind and he is explicitly attributing one of them to the interlocutor. In addition, in order to realize this attribution the speaker needs to rely on background and shared knowledge.

From the point of view of the interlocutor, the comprehension of negative sentences is not always just a matter of simulation of the propositional content. In fact, in order to correctly understand a cancelling implicature negation, the interlocutor needs to understand the inferences that can be drawn from the negated content of the sentence. This entails explicit mindreading for the interlocutor as well. Intentional and highly inferential communication is often explicit in the minds of speakers.

A simulation process of the propositional content can also be taking place with cancelling implicature negation. This is not questioned here. However the simulation of the propositional content is only a part of the full process of language comprehension.

It would be interesting to test the simulative processes at work during the production of inferential negative sentences. Is the content of inferential communication simulated as well? Do we simulate both the literal and the inferential meaning of those sentences? Which are the brain circuitries that implement our more complex inferential abilities? These issues are still waiting for an answer.

Interesting studies have been carried out on the ability to draw inferences during language comprehension and to integrate semantics and world knowledge. For instance, Chow *et al*. (2008) carried out an fMRI study in which participants were asked to read short passages and predict the development of the situations by drawing correct inferences. To trigger the participants to generate the correct inference, each short passage was followed by a lexical-decision task. The authors manipulated the predictability of the target words in the lexical decision-task. Predictable words were those not explicitly mentioned in the short passages but inferable from them. This study, thus, investigated the neural mechanisms sub-serving inference processes. However, it does not seem to address the problem of so-called pragmatic intrusion, namely to what extent is semantics pragmatically determined? In other words, this study is focused on the problem of predictive inferences in language comprehension while it does not address the different problem of the "intended meaning" of an utterance. It seems likely that we need inferential abilities to understand the meaning of an utterance, what the speaker is talking about, and not only to predict the event that can follow from the uttering of that sentence.

Another issue, then, is worthy of remark here. Studies on the simulative processes of negative sentences are carried out in the field of embodied semantics. The definition of meaning in this field of research seems to be problematic. Action-related meanings are represented in the brain through a mental simulation realized in the same sensory-motor circuitries that enable actions. Thus, in this view, a part of our vocabulary is built on our sensory-motor possibilities. Metaphorical thought and mental imagery allow us to have abstract meanings (Gallese & Lakoff 2005).

The problem with this account is that it considers semantics, specifically action-related words, as a stable field (dictionary model of language). For example, Meteyard *et al*. (2012) reviewed the work that has been done in the last decade on the topic of embodied language. The authors classified language theories, putting them on a continuum between those that are fully embodied and those that are fully disembodied. Semantics, in all of these works, seems to be conceived of as the link between a word and its representation, that can be more or less embodied. These representations also seem to be pretty stable, no room is left for pragmatics in the definition of semantic representations. However, language does not work as a rigid code. We can play with words, we can use irony, we can lie or we can simply be misunderstood. This is possible because meaning, in real-life linguistic activity, is mostly constructed in the context of speech. Indeed, many authors have been suggesting that we should abandon the distinction between semantics and pragmatics. Meaning, in this view, is constructed in every context of speech on the basis of the speakers, their background knowledge, their level of shared knowledge, their goals (Carapezza & Biancini 2012) and the physical context of the speech act.

Imagine a boy that returns home. His father sees him and asks: "So?" — and the boy answers with a smile: "It was fine". This conversation could only be understood by one who shares their same background knowledge. For example, the boy could have returned from an exam or a job interview. Or a date with a girl that he really likes. The father is asking about the outcome. Thus, it is likely that in this case both the father and the son are performing a mental simulation. But is the mental simulation pertinent to the words *So* and *It was fine* or, more likely, to the implicit meanings that can be inferred from those words? Moreover, these very same words uttered in a different context by different people would have a very different meaning.

Thus, studies on embodied language should discuss the notions of meaning and semantics further. Linguistic activity seems to not only be realized by means of a fixed and conventional repertoire of meanings, even in the case of action-related language. Up to now, the role of pragmatics and inferential processes are still missing in this research paradigm. Hence, inferential processes involved in language comprehension seem to be the next step that should be addressed in embodied language research.

In particular, studies of simulative processes of negation show an overt bias, the consideration of negation as mainly as descriptive while ignoring its metalinguistic functions (Horn 1985). Examples of negation that cancels presuppositions and implicatures, i.e. metalinguistic negation, can be easily found in everyday conversations. Below, are a few example dialogues:

(3) (Andy meets Barry.)
A: I saw you at the restaurant yesterday.
B: I did not move from my apartment.
A: Sorry, I was pretty sure that you were the man I saw with your wife!
After this conversation Barry goes home and says to his wife Carol:
B: You were at the restaurant with another man yesterday!
(Carol replies with a very classical answer.)
C: It isn't what you think!

(4) In a big company there was a cash deficit. The boss addresses one of his employees with these words:
Boss: You have recently bought a very expensive car, John.
John: I did not do anything that was beyond my possibilities.

In all of these examples, negative sentences negate a presupposition or implicature of the interlocutor. In the first case, Barry is negating the presupposition of Andy that Barry was the man at the restaurant. Carol is negating the implicature of Barry that she was cheating on him. Finally, the employee is negating the implicature of his boss that he stole money.

It is likely that in understanding these negative sentences, we start with a simulation. Following Kaup & Zwaan (2007), it is likely that we start with a simulation of the negated content and then we simulate the actual state of affairs. However, a simulation of the actual or the negated content is not sufficient in order to understand these dialogues. The boss is not explicitly saying that the employee stole the money. He is implying this sentence. And the employee is not explicitly negating this implicature. His negative sentence is an implicit negation of the boss's implicature.

Thus, the simulation of the propositional content is not sufficient in order to explain real-life linguistic activity. In addition, we need to explain how we get from the propositional content to those inferential meanings that are not explicitly stated but are explicitly accessed by speakers.

Simulative understanding is the first step in language acquisition that we reach during ontogeny and very likely it was the first step in phylogeny as well. However, at some point we got more complex socio-cognitive abilities that made more sophisticated linguistic activities possible. In competent speakers, these two mechanisms interact, making it hard to isolate one from the other.

Concerning the acquisition of linguistic negation, we can understand rejection and non-existence in the simulative paradigm. However, it is very unlikely that denial can be accounted for in this same paradigm, in particular for cases of presupposition cancelling negation or implicature cancelling negation. In order to account for denial we need to introduce explicit inferential abilities. The inferential level of linguistic negation will be the topic discussed in the next section.

3.2. *The Inferential Level of Negation*

The examples discussed in the previous paragraph seem to suggest that simulating the negated or actual content of a proposition is not always sufficient in order to understand negative sentences even if these are action-related. If our understanding of propositions such as those presented in the previous examples were based on a simulation of the propositional content, then much of the meaning of those dialogues would be lost. A simulation is most likely taking place. However, many questions arise about the content and temporal dynamic of this simulation. Only further empirical investigation can help to answer these questions.

Up to now, we have argued that complex socio-cognitive abilities are needed in order to produce and comprehend denial, the most complex form of negation. The next question we are going to address will be: Is there any empirical evidence supporting this claim?

We try to answer this question by looking at the acquisition of linguistic negation in autistic children. Autism is a neurodevelopmental disorder with three characteristic features: social impairments, communicative-linguistic impairments, and repetitive and stereotyped behaviors (Tager-Flusberg 1999). The first two aspects of autism are usually referred to as a condition of "mind-blindness", that is, the inability to read our and other people's mental states (mindreading). Hence, according to this account, autistic subjects have a specific deficiency in understanding other people's mental states and, as a consequence of this lack of comprehension, they have communicative and social deficiencies. Indeed, the ability of attributing mental states such as intentions, beliefs, desires, etc., to other people is the ground for social behaviors and linguistic communication.

Because they lack the mindreading ability, autistic individuals have difficulties in interpreting other people's communication and behavior. Different hypotheses have been proposed to explain this disorder. Baron-Cohen (2002), following a hypothesis originally proposed by Asperger (1947), gave empirical support to the idea that autism can be considered as an extreme form of the male brain. In other words, autistic individuals demonstrate hyper-developed systemizing abilities, i.e. they perform over the average in tasks that require identifying the variables of a system and predicting its rule-governed behaviors while they perform poorly in all the tasks that require empathizing with others. It is worth noting that both men and women have the systemizing and empathizing skills but usually women have more developed empathizing skills then men and vice versa, men have more developed systemizing abilities. The extreme male brain that seems to characterize autism, thus, is a particular combination of these abilities that consists in hyper-developed systemizing skills and hypo-developed empathizing skills. This is, of course, only one of the hypotheses that have been put on the table about this enigmatic pathology. Lately, the alternative explanation that subjects with autism spectrum disorder have a dysfunction of the mirror neurons system consisting in its disordered activity has also been considered (Iacoboni & Dapretto 2006). According to this hypothesis, this dysfunction of the mirror system is responsible for the lack of that immediate and automatic attunement with others that makes us able to understand their basic motor intentions.

In any case, whatever can be considered to be the origin of this pathology, one of the defining characteristics of autism is mind-blindness. However, this mindreading deficiency can be experienced to different degrees and this means that differences among individuals in the severity of symptoms are likely to be observed.

So far, we have made two points. Firstly, we have examined the cognitive requirements underlying the acquisition of linguistic negation, arguing that in its complex forms, linguistic negation requires second-order mental representations and complex mindreading abilities. Secondly, we have identified a neuro-developmental disorder, autism, where subjects are expected to show a specific deficiency in mindreading. Now, if these two assumptions are correct, autistic individuals should find complex forms of linguistic negations difficult to interpret correctly. And this seems, in fact, to be the case.

Shapiro & Kapit (1978) looked at the use of linguistic negation in autistic children and in typically developing 3- and 5-years-old controls. Subjects had to follow an experimenter's instructions eliciting comprehension, production or imitation of negative sentences. Findings suggested that autistic children used a different strategy than their controls in accomplishing the tasks. In fact, they performed better than controls in the imitation task but had significantly lower performances in production. All groups performed better in comprehension than in imitation. Still the autistic children's performances in comprehension were lower than the 5-years-olds and even lower than one of the two groups of 3-year-old normally developing controls. Shapiro & Kapit (1978: 349) find that "the autistic subjects produce fewer and more rigid negations as well as imitating well, suggesting adequate registration and reply but poor integrative processing of linguistic form for social and communicative use".

Moreover, Tager-Flusberg *et al.* (1990) looked at language acquisition in autistic children and children with Down syndrome. Children were visited in their homes and video-taped while playing with their mothers. Conversations were subsequently transcribed by the experimenters. Results showed that autistic and Down syndrome children acquired syntactic structure to express negation in the same order as typically developing children. However, autistic children only used syntactic structures of negation to express rejection and non-existence while at later stages, children with Down syndrome also express the function of denial. Significantly, the expression of denial was absent in the linguistic production of autistic children. The authors interpreted these findings as a result of a lack in Theory of mind, under the assumption that denial requires attributing mental states to the listener. "This paucity of denial reflects impairments in Theory of mind. […] These aspects of mental state understanding are specifically impaired in autism and it is therefore not surprising that this function of language, denial, is almost never used by young children with autism" (Tager-Flusberg 1999: 332).

Delayed negation processing was found by Schindele *et al.* (2008) even in adults diagnosed with high-functioning autism and Asperger's syndrome. Subjects were required to read short stories ending with either a negative or affirmative sentence. Depending on the context, the last sentence could be pragmatically felicitous or infelicitous. Normal controls only read negative final sentences in the pragmatically infelicitous context slowly while the two clinical groups had no context-related effect, showing longer reading times for negative sentences in both conditions.

The data about the production and comprehension of negative sentences in autistic individuals seems to support the conclusion that complex socio-cognitive abilities are needed in order to produce and comprehend denial. Indeed, when the ability to understand other people's mental states is impaired, production and comprehension of denial is impaired as well.

The next step will be to put all of these elements in a unitary model for language acquisition.

3.3. *Putting Together Simulative and Inferential Abilities into a Two Step-Model for Language Acquisition*

So far, we have seen that different kinds of embodied simulations occur in language comprehension. We have also seen that in many occasions the simulation of the propositional content does not sufficiently explain the pragmatic level of linguistic communication. Most linguistic meaning is implicitly communicated through inferential processes. What can all of this data tell us about the acquisition of language?

Children have a language-ready brain, but they need to be in a speaking community in order to start speaking themselves. Even Noam Chomsky made this kind of assertion (Chomsky 1988). Unlike what Chomsky believed, the claim here is that having a language-ready brain is not the same as having an innate, domain-specific, device for language.

Arbib (2009: 265) hypothesizes that the language-ready brain resulted from the evolution of a progression of mirror systems and linked brain regions ''beyond the mirror'' that made the full expression of their functionality possible.

Following the mirror system hypothesis advanced by Arbib and Rizzolatti for phylogeny (Arbib & Rizzolatti 1997, Rizzolatti & Arbib 1998), the proposal made here for ontogeny is that children have a language-ready brain that, based on mental simulation implemented by the mirror neuron system, makes the start of the acquisition of language possible. At this first step, the mirror neuron system provides us with the ability to approximately comprehend intentions, mainly motor intentions, and to start the process of language acquisition. In fact, the mirror neuron system can explain the so-called construction grammar model for language acquisition (Tomasello 1999) in terms of intention understanding, imitative learning and simulative understanding.

At this point, when the child enters the linguistic game, it is language itself that affects his or her cognitive functions and even brain. For example, the acquisition of literacy affects brain lateralization. Illiterate subjects, according to Arbib (2009), are more right-lateralized than literate controls.

Language makes our socio-cognitive abilities more complex. In particular, a full language shared by a cultural community makes our mindreading ability far more complex. Once the child acquires an explicit mindreading ability he or she is now ready to enter the second step of language acquisition. In fact, an explicit mindreading ability allows children to produce and comprehend inferential and implicit linguistic communication. The simulation of the propositional content by itself is not enough, even if it is very likely that simulations are taking place as well (Gallese & Sinigaglia 2011). Thus, at least two different neural mechanisms seem to be involved in language comprehension. On the one hand, language is deeply grounded in our sensory-motor abilities which are exploited via the mechanism of embodied simulation. On the other hand, we probably make use of a different neural mechanism that sub-serves inferential reasoning and that is an integral part of our linguistic skills (see for, example, Chow *et al*. 2008). Indeed, as pragmatics seems to suggest, linguistic symbols are a highly inferential communication system. And the more implicitly we are communicating, the more we need to be able to explicitly reason about other people mental states.

4. Conclusions

Summarizing the data discussed so far, negation can be considered a good example supporting the hypothesis that speaking is a complex ability realized by means of at least two different mechanisms that are likely developed at different and consecutive steps of cognitive and linguistic development. In fact, categories of linguistic negation have different levels of complexity with different cognitive requirements. Rejection and non-existence can be explained in a simulative account. Denial, instead, needs an inferential explanation.

It is worth noting again, that these two steps of language competence, despite the fact that they develop consecutively, interact in adult language making it really difficult to isolate one from the other.

The final remark is about human nature, its phylogeny and ontogeny. Embodiment does not sufficiently explain human language and cognition. We also must look at the co-evolutionary relationship of culture/language and the brain.

References

Arbib Michael A. 2009. Evolving the language-ready brain and the social mechanisms that support language. *Journal of Communication Disorders* 42, 263–271.

Arbib, Michael A. & Giacomo Rizzolatti. 1997. Neural expectations: A possible evolutionary path from manual skills to language. *Communication and Cognition* 29, 393–424.

Asperger, Hans. 1944. Die autistischen psychopathen im Kindesalter. *Archiv für Psychiatrie und Nervenkrankheiten* 117, 76–163.

Baron-Cohen, Simon. 2002. The extreme male brain theory of autism. *Trends in Cognitive Science* 6, 248–254.

Barsalou, Lawrence W. 1999. Perceptual symbol systems. *Behavioral Brain Science* 22, 577–609. Cambridge, MA: MIT Press.

Barsalou, Lawrence W., Paula M. Niedenthal, Aran K. Barbey & Jennifer A. Ruppert. 2003. Social embodiment. *Psychology of Learning and Motivation* 43, 43–92.

Bloom, Lois. 1970. Language development: Form and function in emerging grammars. New York: Columbia University dissertation.

Buccino, Giovanni, Lucia Riggio, Giorgia Melli, Ferdinand Binkofski, Vittorio Gallese & Giacomo Rizzolatti. 2005. Listening to action-related sentences modulates the activity of the motor system: A combined TMS and behavioral study. *Cognitive Brain Research* 24, 355–363.

Carapezza, Marco & Pierluigi Biancini. 2012. On Wittgenstein's language game. In A. Capone, F. Lo Piparo & M. Carapezza (eds.), *Perspectives on Pragmatics and Philosophy*. Berlin & New York: Verlag-Springer.

Carpenter, Patricia A. & Marcel A. Just. 1975. Sentence comprehension: A psycholinguistic processing model of verification. *Psychological Review* 82, 45–73.

Choi, Soonja. 1988. The semantic development of negation: A cross-linguistic longitudinal study. *Journal of Child Language* 15, 517–531.

Chomsky, Noam. 1988. *Language and Problems of Knowledge*. Cambridge, MA: MIT Press.

Chow, Ho Ming, Barbara Kaup, Markus Raabe & Mark W. Greenlee. 2008. Evidence of fronto-temporal interactions for strategic inference processes during language comprehension. *NeuroImage* 940–954.

Churchland, Patricia. 1986. *Neurophilosophy*. Cambridge, MA: MIT Press.

Deacon, Terrence. 1997. *The Symbolic Species*. New York: W. W. Norton & Co.

De Villiers, Jill. G. 2000. Language and Theory of Mind: What are the developmental relationships? In S. Baron-Cohen, H. Tager-Flusberg & D. Cohen (eds.), *Understanding Other Minds: Perspectives from Cognitive Neuroscience*, 83–123. Oxford: Oxford University Press.

Dimroth, Christine. 2010. The acquisition of negation. In L. R. Horn (ed.), *The Expression of Negation*, 39–72. Berlin & New York: Mouton de Gruyter.

Engelen, Jan A. A., Samantha Bouwmeester, Anique B. H. de Bruin & Rolf A. Zwaan. 2011. Perceptual simulation in developing language comprehension. *Journal of Experimental Child Psychology* 110, 659–675.

Gallese, Vittorio. 2007. Before and below theory of mind: Embodied simulation and the neural correlates of social cognition. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 362, 659–669.

Gallese, Vittorio. 2008. Mirror neurons and the social nature of language: The neural exploitation hypothesis. *Social Neuroscience* 3, 317–333.

Gallese, Vittorio, Luciano Fadiga, Leonardo Fogassi & Giacomo Rizzolatti. 1996. Action recognition in the premotor cortex. *Brain* 119, 593–609.

Gallese, Vittorio & George Lakoff. 2005. The brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology* 21, 455–479.

Gallese, Vittorio & Corrado Sinigaglia. 2011. What is so special about embodied simulation? *Trends in Cognitive Sciences* 15, 1–8.

Iacoboni, Marco & Mirella Dapretto. 2006. The mirror neuron system and the consequences of its dysfunction. *Nature Reviews Neuroscience* 7, 942–951.

Horn, Lawrence R. 1985. Metalinguistic negation and pragmatic ambiguity. *Language* 61, 121–174.

Kaup, Barbara, Jana Ludtke & Rolf A. Zwaan. 2006. Processing negated sentences with contradictory predicates: Is a door that is not open mentally closed? *Journal of Pragmatics* 38, 1033–1050.

Kaup, Barbara & Rolf Zwaan. 2003. Effects of negation and situational presence on the accessibility of text information. *Journal of Experimental Psychology: Learning, Memory and Cognition* 29, 439–446.

Kaup, Barbara & Rolf A. Zwaan. 2007. The experiential view of language comprehension: How is negation represented? In F. A. Schmalhofer & C. A. Perfetti (eds.), *Higher Level Language Processes in the Brain: Inference and Comprehension Processes*. Mahwah, NJ: Erlbaum.

Kousta, Stavroula-Thaleia, Gabriella Vigliocco, David P. Vinson, Mark Andrews, Elena Del Campo. 2010. The representation of abstract words: Why emotion matters. *Journal of Experimental Psychology: General*, doi:10.1037/a0021446.

Meteyard, Lotte, Sara R. Cuadrado, Bahador Baharami & Gabriella Vigliocco. 2012. Coming of age: A review of embodiment and the neuroscience of semantics. *Cortex* 48, 788–804.

Onishi, Kristine H. & Renée Baillargeon. 2005. Do 15-month-old infants understand false beliefs? *Science* 308, 255–258.

Pea, Roy. 1980. The development of negation in early child language. In D. R. Olson (ed.), *The Social Foundations Of Language & Thought: Essays in Honor of Jerome S. Bruner*, 156–186. New York: W. W. Norton.

Pulvermüller, Friedmann 1999. Words in the brain's language. *Behavioral and Brain Sciences* 22, 253–336.

Pulvermüller, Friedmann. 2002. A brain perspective on language mechanisms: From discrete neuronal ensembles to serial order. *Progress in Neurobiology* 67, 85–111.

Rizzolatti, Giacomo & Michael A. Arbib. 1998. Language within our grasp. *Trends in Neuroscience* 21, 188–194.

Rizzolatti, Giacomo, Luciano Fadiga, Vittorio Gallese & Leonardo Fogassi. 1996. Premotor cortex and the recognition of motor actions. *Cognitive Brain Research* 3, 131–141.

Rizzolatti, Giacomo, Luciano Fogassi & Vittorio Gallese. 2001. Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience* 2, 661–670.

Schindele, Rebecca, Jana Lüdtke & Barbara Kaup. 2008. Comprehending negation: A study with adults diagnosed with high functioning autism or Asperger's syndrome. *Intercultural Pragmatics* 5, 421–444.

Shapiro Theodore & Richard Kapip. 1978. Linguistic negation in autistic and normal children. *Journal of Psycholinguistic Research* 7, 337–351.

Tager-Flusberg, Helen. 1999. A psychological approach to understanding the social and language impairments in autism. *International Review of Psychiatry* 11, 325–334.

Tager-Flusberg Helen, Susan Calkins, Tina Nolin, Therese Baumberger, Macia Anderson & Ann Chadwick-Dias. 1990. A longitudinal study of language acquisition in autistic and Downs syndrome children. *Journal of Autism and Developmental Disorders*, 20, 1–21.

Tettamanti, Marco, Giovanni Buccino, Maria Cristina Saccuman, Vittorio Gallese, Massimo Danna, Paola Scifo, Ferruccio Fazio, Giacomo Rizzolatti, Stefano F. Cappa & Daniela Perani. 2005. Listening to action related sentences activates fronto-parietal motor circuits. *Journal of Cognitive Neuroscience* 17, 273–281.

Tettamanti Marco, Rosa Manenti, Pasquale A. Della Rosa, Andrea Falini, Daniela Perani, Stefano Cappa & Andrea Moro. 2008. Negation in the brain: Modulating action representations. *NeuroImage* 43, 358–367.

Tomasello, Michael. 1999. *The Cultural Origin of Human Cognition*. Harvard, MA: Harvard University Press.

Volterra, Virginia & Francesco Antinucci. 1979. Negation in child language: A pragmatic study. In E. Ochs & B. Schiefflin (eds.), *Developmental Pragmatics*, 281–303. New York: Academic Press.

Wojciehowski, Hanna & Vittorio Gallese. 2011. How stories make us feel: Toward an embodied narratology. *California Italian Studies*, 2(1), http://escholarship.ucop.edu/uc/item/3jg726c2.

Valentina Cuccio
Università degli Studi di Palermo
Department of Philosophy, Philology, History, Arts, and Criticism of Knowledge
Edificio 12 Viale delle Scienze
90128 Palermo
Italy
valentina.cuccio@unipa.it

Embodied Social Cognition and Embedded Theory of Mind

Marco Fenici

Embodiment and embeddedness define an attractive framework to the study of cognition. I discuss whether theory of mind, i.e. the ability to attribute mental states to others to predict and explain their behaviour, fits these two principles. In agreement with available evidence, embodied cognitive processes may underlie the earliest manifestations of social cognitive abilities such as infants' selective behaviour in spontaneous-response false belief tasks. Instead, late theory-of-mind abilities, such as the capacity to pass the (elicited-response) false belief test at age four, depend on children's ability to explain people's reasons to act in conversation with adults. Accordingly, rather than embodied, late theory-of-mind abilities are embedded in an external linguistic practice.

Keywords: embodied and embedded cognition; false belief test; social cognition; social understanding; theory of mind

1. Introduction

Recent years have seen the birth of a new conception of the mind, namely, embodied cognition (Varela *et al*. 1991; Steels & Brooks 1995; Clark 1997, 2008; Lakoff & Johnson 1999; Shapiro 2011). Briefly, embodied cognition asserts that our physical constitution, that is, the body, matters to the definition of our mental life. In opposition to traditional cognitive psychology, according to which cognitive activity depends on the manipulation of amodal representations that control motor responses, embodied cognition states that perception and action are constitutive of mental representations. Accordingly, motor as well as sensory processes have a central role in the definition of cognition.

Embodied cognition supports a principle of economy in the definition of cognitive processes: It suggests substituting, until possible, reference to amodal representations with workable hypotheses about the functioning of sensory and motor systems.[1] This is also consistent with evolutionary explanations:

I wish to thank Jay Garfield, Zuzanna Rucinska, Silvano Zipoli Caiani, and two anonymous reviewers for useful discussion and comments on previous versions of this article. Preparation of this article was supported by a Short Research Grant for Doctoral Candidates and Young Academics and Scientists from the German Academic Exchange Service (DAAD).

[1] Consider, for instance, how Barsalou (1999) replaces Paivio's (1986) reference to a symbolic

ISSN 1450–341

Biolinguistics 6.3–4: 276–307, 2012
http://www.biolinguistics.eu

"[E]volution capitalized on existing brain mechanisms to implement conceptual systems rather than creating new ones" (Yeh & Barsalou 2006: 374). Theoretical economy and evolutionary plausibility thus make the framework of embodied cognition appealing and desirable for the study of cognition. However, economy and plausibility are not compelling reasons to accept embodied cognition. Whether it defines a valuable framework to explain cognitive activity, in general, and more specific cognitive competences, in particular, is an empirical issue, which is worth of consideration.

In this article, I investigate whether embodied cognition is compatible with social cognitive development and, in particular, with the capacity to attribute mental states (such as beliefs, desires, and intentions) to others in order to predict and explain their behaviour.[2] Although investigation on children's acquisition of this capacity dates back almost thirty years ago (Wimmer & Perner 1983; Baron-Cohen *et al*. 1985), traditional accounts of social cognitive development have never considered the possibility that it stands for an embodied capacity. This came to a reason. On the one hand, modularist accounts of theory-of-mind acquisition (Perner 1991; Baron-Cohen 1994, 1995; Leslie 1994, 1995; Scholl & Leslie 1999) usually referred to sentence-like representations to describe the processing of the cognitive mechanism implementing theory-of-mind abilities; thereby, they assumed an amodal medium of representation that is incompatible with embodied cognition. On the other hand, child-as-scientist accounts (Gopnik 1990, 1996; Carey & Spelke 1996; Wellman & Gelman 1997; Wellman 2002) preferentially focused children's theoretical understanding of folk psychology and left aside the analysis of cognitive processes.[3]

Recent research however allows rethinking the embodiment of theory-of-mind abilities. Based on evidence from the false belief test (FBT) paradigm, traditional explanations of theory-of-mind acquisition assumed that children acquire the ability to attribute false beliefs to others in their fourth year of life (Wellman *et al*. 2001; Wellman & Liu 2004; Liu *et al*. 2008). Recent results however demonstrated that even infants in their second year of life seem to attribute false beliefs when simplified versions of FBT and behavioural responses are considered (see Baillargeon *et al*. 2010 and Sodian 2011 for updated reviews and discussion about alternative interpretations). Because these recent results demonstrate that theory-of-mind abilities are acquired earlier than previously reported, they raise the possibility that theory of mind is embodied in early

---

code for mental representations with reference to a network of multi-modal associations.

[2] The acquisition of this capacity has traditionally been interpreted as equivalent to the possession of a 'theory of mind' (Premack & Woodruff 1978). Further evidence however importantly questioned the idea that such a capacity is acquired all in once due to the maturation of one cognitive mechanism. In what follows, I will hence distinguish *early* and *late* social cognitive abilities — I will consciously employ the expression 'social cognitive abilities' as a synonym for the more specific capacities manifested in false belief tasks. Whenever the term 'social cognition' will be used, it will preferentially refer to early social cognitive abilities, such as those manifested in spontaneous-response tasks. Instead, I will talk about 'theory of mind' to refer to late social cognitive abilities as manifested by the capacity to pass the traditional false belief test (see further).

[3] Simulation theory (Gordon 1986, 2007; Harris 1989; Heal 1986, 1998; Gallese & Goldman 1998; Goldman 2006) represents a case apart. I will discuss embodied accounts of theory-of-mind acquisition related to it in the long of this article.

sensory–motor skills. They also suggest that the capacity to pass FBT is not as central as previously thought to the acquisition of a theory of mind — as some had already claimed (e.g., Fodor 1992; Bloom & German 2000). But these results are merely suggestive. They leave open the question of *which* competences underlie the acquisition of this ability and whether *those competences* fit the framework of embodied cognition.

This article is devoted to the exploration of these two perspectives. I will claim that *early* social cognitive abilities are probably embodied inasmuch as available evidence is consistent with their implementation by cognitive processes integrating sensory–motor information. On the other hand, I will argue that *late* social cognitive abilities are embedded in social and dialogical practices — and, in particular, that the ability to pass FBT at age four denotes the acquisition of a minimal capacity to explain people's reasons to act. Embodiment and embeddedness are two logically distinct hypotheses about the nature of cognition, each appropriate to some cognitive skills, and not to others. Late social cognitive abilities thus fall beyond the borders of embodiment. I will conclude that theory of mind is a composed competence that stands in a complex relationship with the principle of embodiment: It is likely partially embodied, and partially not, but the part that is not is likely embedded.

Section 1 clarifies which conception of embodiment is at stake when discussing whether theory-of-mind abilities are embodied. It also distinguishes embodiment and embeddedness as two logically different principles about the nature of cognition. In section 2, I discuss how the empirical plausibility of an embodied approach to early social cognition is challenged by mentalist interpretations. I argue that embodiment accounts advance a coherent and plausible interpretation that is not dismissed by mentalist pre-theoretical intuitions.

In section 3, I claim that the crucial argument in favor of mentalist interpretations presupposes that early social cognitive abilities develop in continuity with later theory-of-mind capacities. However, I show that empirical evidence disconfirms continuity in social cognitive development. It follows that mentalist interpretations are not in a better position than embodied approaches in describing the earliest forms of social cognition. Empirical investigation should take very seriously the task of deciding to what extend infants' social cognitive abilities can be accounted for by relatively simple embodied processes and mechanisms.

Section 4 turns on four-year-olds' acquired capacity to pass FBT and rejects three different explanations of its developmental pattern, one of them being based on the role of the executive function two others on different aspects of language acquisition. In section 5, I propose as an alternative that the capacity to pass FBT depends on a minimal ability to explain people's reasons to act. I review empirical evidence supporting my proposal, and conclude that late theory-of-mind abilities fit the principles of embedded rather than embodied cognition.

2. Embodied, Situated, and Embedded Cognition

Discussing whether theory-of-mind abilities are embodied first requires clarifying which conception of embodiment is at stake. Generally speaking, embodied cognition asserts that our physical constitution, that is, the body, matters to the definition of our mental life. Although this general principle can be refined or expanded in several ways (Wilson 2002; Anderson 2003; Kiverstein & Clark 2009), it minimally requires only that the processes implementing cognitive abilities importantly rely on sensory (e.g., somatosensitive, interoceptive, proprioceptive) and motor representations (Goldman & de Vignemont 2009; Gallese & Sinigaglia 2011).

This formulation may appear inadequate for at least two reasons. First, even traditional cognitivists acknowledge that sensory and motor processes trivially have a role in cognition. By requiring that they must play an 'important' role, embodied cognition stresses that sensory and motor processes must be central even to the definition of high-level cognitive abilities. Second, anti-representationalist embodiment theorists (e.g., Varela *et al*. 1991; Thelen & Smith 1994; Steels & Brooks 1995; Chemero 2009) would say that this formulation is too weak because sensory–motor processes sufficiently define cognitive activity without the need of positing inner representations. Still, embodiment as a principle neither requires nor denies the existence of mental representations. I believe that this formulation would be accepted by most of its non-radical supporters. In what follows, I will hence assume that this formulation correctly states a viable minimal definition of embodiment: In order to be embodied, social cognitive abilities need to be implemented by cognitive processes that importantly rely on sensory and motor information.[4]

Before going further, it is important to disentangle embodiment from two close principles. Making this distinction will come at help later when discussing how it relates to different components of theory-of-mind abilities. *Situated* cognition asserts that we cannot artificially separate the body, thereby cognitive activity, from the environment in which it is placed.[5] Situated cognition differs from embodiment in that it stresses the role of background information to the processing of any stimulus (Yeh & Barsalou 2006; Barsalou 2009), whereas the latter focuses on the role of the body in actively engaging the organism with the environment.

Embodied cognition has also to be distinguished from *embeddedness*, which highlights the role of external structures in supporting and scaffolding cognitive activity.[6] Embeddedness supports a principle of conceptual economy for cogni-

---

[4] This is substantially a re-proposal of Goldman & de Vignemont's (2009) definition of embodied cognition. It is less exposed to anti-representational concerns because 'representations in a bodily format' have been replaced by reference to their vehicles, i.e. cognitive processes integrating sensory and motor information.

[5] Situatedness is closely related to ecologism (Gibson 1979), which states that behavior cannot be studied independently of the environment in which it occurs. It is also presupposed by enactivist approaches — in both their representationalist (Grush 2004; Noë 2005) and anti-representational versions (Chemero 2009; e.g., Varela *et al*. 1991; Thelen & Smith 1994; Steels & Brooks 1995) — according to which cognition is the outcome of the interaction between the body and the environment so that action, not only perception, is constitutive of cognitive activity. Situatedness is however a weaker principle than enactivism, which also assumes the truth of embodiment.

[6] We continuously modify and construct the space around us disseminating information in it to our next benefit. Think about how we re-locate objects in our house not to stub our toe on

tion in a direction opposite to embodiment. Where embodiment points to conceptual economy in the inwards direction of the modal nature of mental representations, which are the inner vehicles of cognition, embeddedness points instead to the outwards direction of the environment, which simplifies cognitive processes by scaffolding cognitive activity.

Moreover, not every form of embedded cognition is situated (and vice versa but I will not pursue this here). Language, for instance, is a powerful tool to discharge the computational complexity of a task (e.g., remembering a long sequence of actions) on an external support (e.g., a piece of paper, or a sentence one can rehearse by the help of auditory memory) (Vygotsky 1934, 1978). However, although language is learned in interaction with the (social) environment, it is later internalised (Berk 1991; Winsler *et al*. 2003), and can be used as a symbolic tool in isolation from the environment (Clark 1998). Language acquisition thus deeply impacts on cognition disregarding the situatedness of the cognizer (Karmiloff-Smith 1992; Clark & Karmiloff-Smith 1993).[7]

2. Embodied Cognition and Early Social Cognitive Abilities

I have made explicit a minimal conception of embodiment. I have also distinguished it from the two closely related principles of situated and embedded cognition. We can now start investigating whether the capacity to attribute mental states can be defined in a cognitive system so characterized.

Recent research employing violation-of-expectancy and first looking paradigms recently showed that even infants seem to attribute false beliefs to others in their second year of life. For instance, Onishi & Baillargeon (2005) found that 15-month-olds look significantly longer when they see an experimenter acting incoherently with respect to her false beliefs. Their result, obtained in a violation-of-expectancy paradigm, was replicated considering 25-month-olds' anticipatory looking, which is a clearer index of infants' expectations (Southgate *et al*. 2007). In addition, Surian *et al*. (2007) found that 13-month-olds are already sensitive to one agent's knowledge or ignorance of a situation.

The studies above strictly focused on visual stimulation and responses. Further research investigated infants' processing of others' beliefs obtained through sensory modalities other than vision and showed that infants' social cognitive abilities are not restricted to the exclusive elaboration of visual input. Infants have been found sensitive to one agent's false beliefs induced through

---

them; or how we leave post-its on the fridge and knotted handkerchiefs in our pockets to remind of next duties; or how we fill our environment with road and shop signs. These activities allow us to discharge the computational complexity of cognitive processes in the environment. They relieve the cognitive load of memory and simplify both perception and action planning (Kirsh 1995; Clark 1997).

[7] The scaffolding role of language is not limited to cognitive agents in isolation. Thanks to linguistic communication, cognitive processes can be distributed across the members of a group — as the crew of a ship (Hutchins 1995a), or the aircrew of a plane (Hutchins 1995b), or a surgery team — thereby supporting the execution of complex cognitive tasks. Linguistically mediated communication also allows the emergence of important forms of cultural transmission across generations (Dawkins 1976; Latour 1986).

proper communication (Song *et al*. 2008), through incorrect deductions from perceptual cues (Song & Baillargeon 2008), and through tactile perception (Träuble *et al*. 2010). In addition, not only they consider others' non-visually induced false beliefs, but they can also actively react to them (Buttelmann *et al*. 2009; Southgate *et al*. 2010; Knudsen 2011).

According to a first interpretation of these results, early forms of social cognition can be explained by cognitive processes that operate on perceptual input and mostly automatically trigger low-level motor responses (e.g., sustained attention and anticipatory looking). These processes integrate visual information that infants obtain by observing other agents, but they likely involve also motor representations. Extensive data indeed show that processing others' actions involves the activation of pre-motor areas in adults (Wilson & Knoblich 2005; van Overwalle 2009), and the same likely happens even in infants (Del Giudice *et al*. 2009). This first interpretation of early forms of social cognition is thus consistent with the definition of embodiment introduced in section 1 because, on this view, early social cognitive abilities are implemented by cognitive processes subserving both sensory and motor information. Call this the *embodied view* of early social cognitive abilities.

The embodied view is compatible with very different accounts of the capacity to attribute mental states advanced both in the philosophical and scientific literature.[8] For instance, Gallese (2005, 2007; Gallese & Sinigaglia 2011) argues that the same sensory–motor processes (i.e. the mirror mechanism, Rizzolatti & Craighero 2004) implementing one's own mental states — e.g., one's intention to act — are also used when functionally attributing the same mental state to another — e.g., when understanding another's intention to act. Similarly, Goldman (2006, 2009) claims that mirror neurons play an important role in 'low-level' mindreading and support the attribution of mental states to others. According to Gallagher (2008, 2011), interpreting others' mental states depends on perceptual, rather than inferential, capacities that are employed in situated social interaction and rely on low-level sensory–motor associations developed since early infancy. Finally, according to De Jaeger, "social understanding emerges from a dynamical process of interaction and coordination of two embodied subjects coupled to each other" (Fuchs & De Jaegher 2009: 470; see also De Jaegher 2009, McGann & De Jaegher 2009). Accordingly, we cannot disentangle infants' elaboration of a perceptual input from the motor processes driving infants' reaction to it.

These accounts differ from one another with respect to several issues: which kinds of mental states are attributed by the cognitive processes implementing early social cognitive abilities; how often these processes are at work in everyday social interaction; whether they can be interpreted in representational terms; and how they ground or implement theory-of-mind capacities. With respect to the last point, in particular, these accounts provide very different interpretations of the activity of the mirror neuron system when we observe others' actions. According to Gallese, for instance, mirror neurons

---

[8] The point is not that all the following accounts embraced the embodied view. Rather, they may agree with an embodied explanation of early social cognitive abilities.

underpin out understanding of motor intentionality. For Goldman, instead, they enable us to enter the same mental states that we observer in another person. Gallagher interprets the mirror neuron system as a neural mechanism supporting 'smart' perception. Finally, De Jaegher is very critical of neurological explanations of social cognitive abilities.[9] Nevertheless, she considers that "this is not to say that the link between action and perception found in mirror neuron research does not play an important role for social understanding" (Fuchs & De Jaegher 2009: 469).

Despite these differences, these accounts are nonetheless unanimous with regard to the following theses: (i) the attribution of mental states to others exploits cognitive mechanisms that are developing since early infancy; (ii) sensory–motor processes such as the mirror neuron system constitute the core of these mechanisms. Call these embodied social cognitive processes. The embodied view explicitly adds that (iii) embodied social cognitive processes ground infants' performance in spontaneous-response false belief tasks.

The embodied view presupposes that manifest behavior encoded through visual processes and processed by the mirroring system constitutes the fundamental source of data that infants process in spontaneous-response false belief tasks. In this sense, the embodied view is sympathetic with those proposals explaining infants' sensitivity to others' false beliefs in the terms of different capacities to track more superficial, observational features.[10] For example, it has been argued that infants' performance on spontaneous-response false belief tasks depends on behavior-reading capacities (Penn & Povinelli 2007; Perner 2010; Butterfill & Apperly 2013), on the capacity to remember others' encounter with objects (Apperly & Butterfill 2009; Butterfill & Apperly 2013), or to create triadic associations (Perner & Ruffman 2005; de Bruin & Newen 2012), or even on sensitivity to affordances (de Bruin *et al*. 2011).

Notably, these (more or less strictly) behavioural accounts and the embodied view may disagree about the *interpretation* of the cognitive processes underlying infants' performance in spontaneous-response false belief tasks. However, they are much more in agreement about the empirical *nature* of these processes. Behavioural accounts indeed argue that (i) infants' performance in spontaneous-response false belief tasks does not demonstrate the capacity to attribute (false) beliefs, and that (ii) the cognitive processes underlying infants' looking behavior primarily process others' motor intentions and goal-directed behavior. Analogously, it is the empirical significance of a minimal interpretation of the embodied view that the capacities to process goal-directed behavior and motor intentions are sufficient to ground the earliest forms of social cognition. The two views are thus minimally consistent: They both stress the importance of processing overt

---

[9] These explanations indeed "single out one section only of the whole circle of organism–environment interaction. They fail to address social interaction as a structured and structuring process which in turn influences brain functions" (Fuchs & De Jaegher 2009: 469).

[10] Therefore, Gallagher writes: "What the enactive position adds to the behavioral abstraction position concerns the nature of the meaning that I see in the other's actions. The other's actions have meaning for me in terms of how I may be able to interact with her. […] I think this is consistent with your [the behavioural abstraction] view, but offers a specification about the meaning" (Gallagher & Povinelli 2012: 154).

behavior to display the kind of expectations manifested in spontaneous-response false belief tasks, and they avoid commitment to strong mentalistic interpretations of early social cognitive abilities. Indeed, it is possible that the attribution of mental states merely globally supervenes on the sensory–motor processes underpinning infants' basic abilities to process others' behavior, and is not an explicit independent representational activity.

The alternative to the embodied view, the *mentalist* view, instead claims that infants' early social cognitive abilities already involve the capacity to attribute mental states such as beliefs. For instance, Leslie (1994, 1995) advocates for the existence, at 18 months, of a Theory of Mind Mechanism (ToMM) that allows the use representations as meta-representations, thus constitutes the basic computational mechanism beyond both pretend and belief representations. Similarly, Baillargeon (Scott & Baillargeon 2009; Baillargeon *et al*. 2010) advances that early social cognitive abilities are provided by the maturation of a new modular component in the infant's mind in the second year of life, Subsystem-2, which allows infants to hold in mind a separate representation of a scene.

It is a hallmark of the mentalist view that infants' early social cognitive abilities do not exploit any behavior-reading heuristic.[11] This view rejects both behavioural interpretations of infants' performance in spontaneous-response false belief tasks and the embodied view, which is minimally consistent with them. And, in fact, mentalist accounts of early social cognitive abilities are often associated with criticisms to the fundamental importance of sensory–motor processes to the ability of attributing mental states (e.g., Csibra 2007; Grafton 2009).

Despite the arguments advanced by mentalist theorists, nonetheless, the opposition between their rich explanation of spontaneous-response false belief tasks and the minimal interpretation defended by both embodied and (more or less strictly) behavioural accounts is far from being settled. Of course, this is an empirical debate, and empirical evidence may provide some reason to assess the contrast. If it were found, for instance, that early social cognitive abilities are not flexible enough to properly match mental state attributions — because, for example, they do not retain attributed beliefs beyond short time threshold, or because they are insensitive to some perceptual modality in the process of belief formation —, this would constitute evidence against the mentalist view. On the contrary, the embodied view is challenged by any result showing the complexity of early social cognitive abilities. In front of a very flexible behavior manifested by infants in a variegated set of false belief tasks, it would be more difficult to explain their performance in the terms of the mere capacity to process sensory–motor information. The choice to treat their capacity as theory-of-mind abilities would be theoretically more parsimonious, thereby also preferable.

Both the embodied and the mentalist view nevertheless have general strategies to explain their empirical flaws. In particular, the mentalist view can always maintain that non-flexible manifestations of early social cognitive abilities

---

[11] Adduced motivations are variegated. Leslie claims that ToMM is the essential core of theory-of-mind reasoning because it permits and promotes children's attention to early intentional insight into the behaviors of others, thereby it allows them to learn about these states. Instead, according to Baillargeon, Subsystem-2 implements genuine theory-of-mind capacities because of reasons of parsimony (Onishi & Baillargeon 2005: 257).

are explained by limitation of the computational resources available to the working of the theory-of-mind mechanism (Fodor 1992; Leslie *et al*. 2005; Scott & Baillargeon 2009). On the other hand, the embodied view can always reduce the complexity of mentalist interpretations of infants' behavior by elaborating behavior-reading strategies of some sort (Perner 2010; Butterfill & Apperly 2013).

I take those principled objections as demonstrating that the opposition between the embodied and the mentalist view is also partially a matter of theoretical preference about how to describe very simple capacities manifested in infancy. Although I acknowledge that solving the dispute is lastly a matter of empirical discussion, I want herein to consider further assumptions not clearly spelled out in the current debate. In defense of the embodied view, I will claim that it advances a coherent and plausible interpretation, which is not dismissed by mentalist pre-theoretical intuitions. It thus defines a concrete proposal, and it is should be in the agenda of future empirical investigation assessing to what extend infants' social cognitive abilities can be accounted for by relatively simple embodied processes and mechanisms.

The issue whether (amodal) mental states can in principle be computed by cognitive processes that principally integrate sensory–motor information is particularly relevant to the assessment of the assumptions in favor and against embodied interpretations. With this respect, the embodied view favors that amodality can effectively be reduced to interwoven cross-modal connections (Barsalou 2005; Goldman & de Vignemont 2009; Gallese & Sinigaglia 2011). Accordingly, also the attribution of mental states to others can be implemented by cognitive mechanisms processing sensory–motor information and directly triggering automatic motor responses. Instead, the mentalist view holds that processing sensory–motor information cannot account for the attribution of mental states for the very nature of the modally-non-neutral information that is processed. Nothing less than theory-of-mind processes can account for social cognitive abilities even in infancy.

It is important to note that the embodied view advances a specific claim about the modal nature of information, which can be empirically investigated. The mentalist alternative, on the other hand, merely relies on a principled and, as I see it, unsuccessful objection. Moreover, as discussed above, the embodied view also suggests a viable alternative explanation to data concerning social cognitive abilities in infancy in agreement with (more or less strictly) behavioural accounts. Therefore, if we only consider social cognitive capacities apparent in the second year, the available evidence does not decide between embodied and mentalist interpretations of social cognitive abilities. But if we look more broadly, the evidence supports the embodied alternative. For the sake of parsimony, indeed, there is no need to assume that infants can attribute (false) beliefs if the same cognitive abilities can be explained by more basic capacities to process manifest behavior and motor intentions.

There is, however, a second argument advanced in favor of the mentalist view. Rather than focusing on the second year of life, it hinges on the gradual development of social cognitive abilities from infancy to early childhood. I will assess it in the next section.

3. Social Cognitive Development from Infancy to Early Childhood

The argument from the continuity of social cognitive development states that (i) infants' selective behavior in spontaneous-response false belief tasks appears before their capacity to pass FBT at age four, and that (ii) this capacity is usually interpreted as the explicit manifestation of the possession of the concept of belief, and so that (iii) infants' performance in spontaneous-response false belief tasks is the implicit manifestation of the concept of belief.

The argument underlies many mentalist interpretations of early social cognitive abilities. For instance, Poulin-Dubois *et al*. (2009) report data from longitudinal studies finding that children's performance on traditional false belief tasks is predicted by earlier ability to understand goal-directed actions with computer-animated geometric forms (Yamaguchi *et al*. 2009) or to identify behavioural cues of intentional action in an imitation task (Colonnesi *et al*. 2008). Hence they conclude that "the current data suggest continuity in social cognitive development that provides support for the hypothesis that the sophisticated social cognitive abilities have their roots in infancy" (p. 91).

Unfortunately for the mentalist view, however, that early social cognitive abilities develop before the capacity to pass FBT does not demonstrate that they are the precursors of this capacity. This conclusion follows only if this capacity is demonstrated to develop in strict continuity with them. Therefore, continuity in social cognitive development is the test bed to decide whether mentalist interpretations of early social cognitive abilities are to be preferred to the embodied view. It is on this issue that I will now turn my attention.

Some empirical evidence attests gradual development in social cognition. Southgate *et al*. (2007) found that 25-month-olds gaze in anticipation towards a location where a person would be expected to search if she had a false belief. This extends Onishi & Baillargeon's (2005) result by relating early social cognitive abilities to a more active behavior (i.e. anticipatory gaze). Still two-year-olds are limited in the kind of stimulation that can enhance their anticipatory-looking response. In Southgate and colleagues' study, infants anticipatory looking was prompted by a visual stimulation, but Clements & Perner (1994) and Garnham & Ruffman (2001) found that it cannot be triggered by verbal prompting until age three.

These studies suggest the following developmental pattern for social cognitive abilities: (i) after 15 months, the cognitive processes responsible for social cognitive abilities can already direct infants' attention at the incongruent behaviour of an agent; (ii) after age two, they also start driving anticipatory looking reactions; (iii) at age three, they start bring prompted by verbal stimulation; (iv) finally around age four, they fully integrate with linguistic abilities, thereby also allow children to correctly answer FBT.

Considering this evidence, Baillargeon (Scott & Baillargeon 2009; Baillargeon *et al*. 2010; cf. also Leslie 2005) claimed that young children fail elicited-response FBT because it involves the functioning of at least three different processes. In particular: (i) a process to represent others' false-beliefs, (ii) a process to select the proper response when asked about others' behavior, and (iii) a process to inhibit the tendency to answer the test question based on one's own knowledge. Since spontaneous-response tasks only tap psychological-reasoning, they are passed earlier than traditional elicited-response false belief tasks. As soon as response-selection mechanisms develop (or interface themselves with psychological-reasoning processes) children's anticipatory-looking starts responding to verbal prompts. Finally, when response-inhibition processes properly develop, children also become able to pass elicited-response tasks.

Important considerations nevertheless reject continuity in the development from early social cognition to late theory-of-mind abilities. A first hint comes when considering a possible double dissociation between early and more mature social cognitive abilities.[12] Senju and collaborators (Senju *et al*. 2009, 2010; see also Senju 2011 for a discussion) found that autistic people are impaired on spontaneous-response false belief tasks while at the same time they pass elicited-response tasks (Happé 1995) — their performance being strongly related to their linguistic abilities (Tager-Flusberg & Joseph 2005). This pattern is opposed to the one of three-year-olds, who are impaired on elicited-response tasks while at the same time they pass spontaneous-response tasks. This suggests that the two tasks map different capacities.

Secondly, if the cognitive processes implementing early social cognitive abilities progressively develop in continuity with more advanced social competence, one would expect cognitive biases affecting late social cognitive abilities to be present even at earlier developmental stages. However, a central bias to the capacity to pass FBT such as the 'curse of knowledge' (Birch & Bloom 2003, 2004, 2007) genuinely affects only four-year-olds' performance on elicited-response tasks, while it spares infants' early social cognitive abilities.[13] This challenges the hypothesis that passing FBT at age four depends on the same processes already in place around age two (Samson & Apperly 2010).

Finally and critically, increasing evidence supports a multi-process theory of social cognitive abilities. On the one hand, empirical findings suggest that beliefs are not automatically attributed in FBT. Apperly *et al*. (2006) reasoned that if this were the case, we should consider others' beliefs even when not requested to do so. They thus probed experimental subjects with unpredictable questions about what was happening in a video; the questions concerned either the location of an object, which participant were requested to track, or a false belief of the main character, which were irrelevant to the task goal. They found that longer response times and higher errors where connected to answers about the character's false belief, suggesting that subjects normally did not track it. Also, explicitly requiring subjects to track the character's belief eliminated the asymmetry between belief- and reality-answers, suggesting that such asymmetry

---

[12] Though see Scerif & Karmiloff-Smith (2005) for a warning about the misuse of double dissociations in cognitive neuroscience.

[13] The curse of knowledge refers to the fact that children as well as adults find it difficult to stop considering their own knowledge when asked to assess others' perspectives. That the curse of knowledge spares early social cognitive abilities provides no surprise in the experiment by Southgate *et al*. (2007), where the object that is the content of the false belief is taken out of the scene before infants' response is prompted. No actual knowledge of the object's location thus misleads infants' reaction. This is however not the case in the experiments by Clements & Perner (1994), Garnham & Ruffman (2001), and Onishi & Baillargeon (2005).

depended on the cost of retrieving the character's belief.

Contrary to the case of belief attribution, other findings instead suggest that adults automatically compute others' visual experience even when they themselves have a different view (Samson *et al*. 2010). This result has been demonstrated in six-year-olds (Surtees & Apperly 2012) and, surprisingly, even in seven-month-olds (Kovács *et al*. 2010).[14] However, this capacity is importantly limited in many respects: It does not consider level-2 visual perspective taking (Surtees *et al*. 2011), and it is impaired when the other's perspective includes complex scenarios (Keysar *et al*. 2000, 2003).

In the light of these results, Apperly & Butterfill (2009; cf. also Frith & Frith 2006; Apperly 2010) suggested that adults compute others' mental states by two kinds of cognitive process. *High-level* social cognitive processes develop in early childhood and allow children to pass complex tasks such as elicited-response FBTs. They are highly flexible but cognitively demanding, therefore they do not get automatically employed. In contrast, *low-level* social cognitive processes develop in infancy and have likely been naturally selected. They are cognitively efficient, because they rely on the elaboration of simple features of the perceived input, and explain infants' performance in spontaneous-response false belief tasks. However, the same reason why they are cognitively efficient also makes them inflexible. Indeed, they are very limited both in the kind of information they can process and in how their outcome can influence other cognitive processes. That is, they are *encapsulated* and *impenetrable*: They are activated only by some specific available input, and are of no help to solve general domain problems (Fodor 1983; Coltheart 1999).

Importantly to the present discussion, empirical investigation indicates that early and late social cognitive abilities are provided by completely different cognitive processes. Accordingly, cognitive development does not progress continuously from infancy to early childhood. This rejects the mentalist theorist's argument that the cognitive processes underlying early forms of social cognition must be interpreted in strong mentalistic terms because they represent the early roots of mature theory-of-mind abilities.

In light of the empirical inadequacy of the argument from the continuity in social cognitive development, and considering that mentalist interpretations about early social cognitive abilities inconclusively oppose embodied accounts, we must thus leave it open how to interpret infants' performance in spontaneous-response false belief tasks. Given the severe limitations of infants' early social cognitive abilities, adopting the full vocabulary of folk psychology to describe them may be incorrect, whenever misleading (cf. Kagan 2008 for the same argument against very young infants' possession of the concepts of number and object).

Concluding, it is up to future empirical investigation deciding to what extend infants' social cognitive abilities can be accounted for by relatively simple embodied processes and mechanisms. However, theoretical reflection

---

[14] Interestingly, Kovács and colleagues interpret their result in the terms of the capacity to consider others' *beliefs*, although what they really assessed is the subject's capacity to recall what other agents *saw*. This is a good example of over-interpretation of experimental evidence.

demonstrates that we are not committed to interpret them in a strong mentalistic vocabulary. Consistently with the embodied view, the cognitive processes underlying early social cognitive abilities may be the outcome of a minimal capacity to attribute motor intentions and goal-directed behavior. Coherently with (more or less strictly) behavioural accounts, they may even depend on less sophisticated embodied competences that do not have a direct translation in the vocabulary of folk psychology. Inasmuch as these two interpretations do not mutually exclude but agree about the empirical nature of the cognitive processes underlying early social cognitive abilities, rejecting strong mentalist accounts paves the way to an alternative interpretation coherent with the embodied view.

4. Explaining Theory-of-Mind Acquisition in Early Childhood

If early social cognitive abilities already reflected the capacity to attribute beliefs to others, learning to pass FBT at age four would not be a milestone in children's social cognitive development. This would constitute *a priori* reason not to investigate whether this ability is implemented in embodied cognitive processes. However, the discontinuity between early and late social cognitive abilities attests that elicited- and spontaneous-response false belief tasks are rather distinct. Passing elicited-response FBT thus identifies an autonomous competence in child development.[15] Accordingly, it is still worth investigating whether this ability fits the framework of embodied cognition. In this section, I discuss and reject three explanations of the acquisition of the capacity to pass FBT. This will clear the field to my alternative proposal.

A first attempt to explain children's acquisition of the capacity to pass FBT has appealed to the maturation, around age four, of several components of the executive function:[16] in particular, the capacity to inhibit stimulus-dependent answers (Carlson & Moses 2001; Jacques & Zelazo 2005; Sabbagh *et al*. 2006), cognitive flexibility (Carlson & Moses 2001; Müller *et al*. 2005; Guajardo *et al*. 2009), and visual perspective taking (Harris 1992; Gopnik *et al*. 1994; Farrant *et al*. 2006; Bigelow & Dugas 2008). This explanation of four-year-olds' acquired capacity to pass FBT is also provided by contemporary modularist accounts of the theory of mind — namely, those accounts supporting the mentalist view of early social cognitive abilities (sections 2 and 3).

---

[15] And indeed, the capacity to pass FBT has been demonstrated extremely robust and unlikely depending on minor changes in previous cognitive development. Allowing children to respond by sticking surprised or non-surprised facial expressions (de Villiers & de Villiers 2000), or proper thought (Wellman *et al*. 1996; Woolfe *et al*. 2002), as well as hide and retrieve tasks (Call & Tomasello 1999; Figueras-Costa & Harris 2001) did not improve four-year-olds' performance in any sensitive way, while only mild improvements were found when allowing children to respond by appropriate hand-gesture (Carlson *et al*. 2005), betting coins (Ruffman *et al*. 2001) and lying by deceiving pointing rather than explicit verbal communication (Perner *et al*. 2002).

[16] By 'executive function', cognitive psychologists refer to the suite of cognitive functions supporting goal directed behavior and cognitive control across conceptual domains, including inhibitory control (or response inhibition), working memory, error monitoring, rule representation and use, planning, behaviour organisation, cognitive flexibility, and attentional control (Zelazo *et al*. 2008).

Several findings nevertheless suggest that the executive function really provides only a marginal contribution to the development of theory-of-mind abilities. Firstly, although autistic children do not pass FBT, they normally perform on executive function tasks when tested by a computer rather than by a person (Ozonoff *et al*. 1991; Ozonoff 1995). Secondly, language delayed deaf children raised by hearing parents are not at all impaired in executive function such as non-verbal working memory, inhibitory control, and conditional rule following; still they fail FBT (P.A. de Villiers 2005). Finally, children in Asian countries manifest earlier competence than their Western peers at executive function tasks, the effect perhaps being due to their education more inclined toward self-control. Nevertheless, early improved executive function does not translate into superior performance in FBT (Sabbagh *et al*. 2006; Liu *et al*. 2008; Oh & Lewis 2008; Lewis *et al*. 2009).[17]

Language acquisition constitutes a better candidate than the maturation of the executive function to account for children's late acquisition of theory-of-mind abilities. Meta-analyses showed that the capacity to pass FBT relates to linguistic competence, the correlation from linguistic abilities to social understanding being stronger than the opposite (Astington & Baird 2005b; Milligan *et al*. 2007). Still, even when focusing on the contribution of language to FBT passing, many different aspects of language acquisition may be relevant (Astington & Baird 2005; de Villiers 2007: 1869–1871). Investigating the embodiment of late social cognitive abilities thus depends on assessing their different contribution.

One explanation that may account for the correlation between language acquisition and FBT passing is that younger children lack the representational capacity to store others' (false) beliefs (Leekam & Perner 1991; Perner 1995; Leekam *et al*. 2008). Accordingly, FBT would measure children's meta-representational abilities. Language acquisition may thus impact children's capacity to pass FBT because, by enabling new representational formats (Karmiloff-Smith 1992), it enables and/or improves the representation of the mental states.

This explanation is strongly supported by de Villiers and collaborators' finding that syntax acquisition, and, in particular, the mastery of sentential complements — i.e. the sentences introduced by a 'that' in mental propositional attitudes (e.g., "he thinks *that-p*") as well as reporting attitudes (e.g., "he says *that-p*") — is predictive of children's ability to pass FBT (de Villiers & Pyers 2002; de Villiers & de Villiers 2003; J.G. de Villiers 2005, 2009).[18] On de Villiers' original interpretation, this was considered evidence that the mastery of sentential complements reshapes children's cognition by providing a new representational format to store meta-representations, therefore also to attribute beliefs to others.[19]

---

[17] Cf. also Sabbagh *et al*. (2010) for an extended criticism of the role of the executive function in promoting late social cognitive abilities.

[18] The result has been confirmed by comparative studies on different populations of deaf children (Peterson & Siegal 2000; Garfield *et al*. 2001; P.A. de Villiers 2005; Pyers & Senghas 2009; Schick *et al*. 2007), and by training studies, where children were trained in FBT, a Test for Complements and other relevant tasks (Hale & Tager-Flusberg 2003; Lohmann & Tomasello 2003; Lohmann *et al*. 2005).

[19] This proposal shares with embodied cognition the focus on cognitive development to explain social cognitive development. Nevertheless, it does not agree with embodied cognition on the representational format encoding belief attributions. Indeed, it supposes

Several findings nevertheless reject de Villiers' analysis that representing others' beliefs is the main problem in FBT. Indeed, according to de Villiers, there must be one moment in which children learn how to represent sentential complements. However, several studies demonstrated that children start mastering complementation at different ages depending on the context in which it occurs. This is revealed by considering sentential complements selected by desire verbs in German (Perner *et al*. 2003), by pretence verbs (Garfield *et al*. 2009), as well as relative clauses (Smith *et al*. 2003).[20] Moreover, de Villiers' proposal advances that children's difficulty with FBT depends on the general understanding that verbs of thought select either true or false sentential complements. However, the mastery of complementation likely predicts FBT passing only because it requires children to understand that verbs of thought can specifically select *false* complements (Cheung *et al*. 2004, 2009). And indeed, although the mastery of sentential complements is sufficient to pass FBT, children's difficulty also partially depends just on the comprehension of the deceiving character of the false beliefs (Lohmann & Tomasello 2003; Lohmann *et al*. 2005). Therefore, providing the right representational format to represent others' beliefs is unlikely the exclusive reason why language acquisition supports late social cognitive development.

A second attempted explanation of the correlation between language acquisition and FBT passing is that FBT requires children to master not just the representational format of attributed beliefs and desires, but also belief–desire reasoning, that is, the capacity to inferentially combine attributed mental states to make predictions about others' future actions. This proposal is largely shared among supporters of both the modularist and the child-as-scientist view of the theory (cf. the introduction), who advanced that passing FBT require children to develop (either implicit or explicit) inferential abilities. Accordingly, language acquisition would improve the capacity to pass FBT by bolstering children's belief–desire reasoning capacities.

Despite its popularity, we should be cautious to adopt this solution: In fact, several reasons suggest that belief–desire reasoning is really not needed to pass FBT. A first weak argument is that we do not consciously perform belief–desire reasoning very often (Gallagher 2007). Secondly, folk psychology apparently works differently across different cultures (e.g., Lillard 1998; Vinden 1999); this should lead to the conclusion that passing FBT is culture dependent — a result for which partial evidence has been provided (Wellman *et al*. 2006; Liu *et al*. 2008; Shahaeian *et al*. 2011). Third, passing FBT by belief–desire reasoning poses a typical inverse problem (Csibra & Gergely 2007) which requires solving an abductive inference. This makes unlikely that children rely on belief–desire reasoning to pass the test (Apperly 2010; cf. also Ratcliffe 2007; Perner & Roessler 2010; de Bruin *et al*. 2011 for related discussion).

---

that others' beliefs are encoded in sentence-like representations, that is, in an amodal representational medium that is very different from sensory and motor representations.

[20] J. G. de Villiers (2005) opposed that *only* the mastery of that-clauses selected by verbs of thought is predictive of children's ability to pass FBT, and proposed that such a competence is scaffolded by their experience with verbs of speech (e.g., *saying*, *telling*). However, there is no evidence in the literature for a developmental gap between the mastery of verbs of speech and the mastery of thinking verbs.

Of course, denying that the mastery of belief–desire reasoning is not necessary to pass FBT requires explaining why belief–desire reasoning is apparently so pervasive in everyday life (Spaulding 2010). However, note that folk psychology and the attribution of mental states are often employed to *explain* past actions rather than to *predict* future ones. Therefore, the pervasiveness of belief–desire reasoning may well depend on its relevance in rationalizing people's behavior by reporting their reasons to act.[21] If that is the case, we can abandon the idea that passing FBT requires the mastery of belief–desire reasoning.

I have rejected three explanations of four-year-olds' acquired ability to pass FBT, one of them being based on the role of the executive function two others on different aspects of language acquisition. In the next section, I will introduce my alternative proposal. We will hence be in position to judge whether late theory-of-mind abilities fit the framework of the embodied cognition.

5. Embedded Cognition and Theory-of-Mind Acquisition

My previous analysis rejected two explanations of the correlation between language acquisition and four-year-olds' acquired capacity to pass FBT. However, it has not refuted the main idea that passing FBT depends on language acquisition. My proposal carves out an alternative explanation for that.

I suggest that passing FBT demonstrates the acquisition of a minimal capacity to explain people's reasons to act. Since the very early infancy, children are continuously exposed to stories and narratives that clarify the reasons why people acted in the way they did. Although full reasons defined by belief–desire pairs are rarely provided, these stories identify relevant constituents of these reasons (e.g., beliefs, desires, intentions, behavioural traits, personality features) and acquaint children with the domain of folk psychology (Hutto 2008; Nelson 2009). I claim that children's acquaintance with these narratives, and in particular with those stories focusing false utterances, the deceiving aspect of things, and lying behaviors, promotes their understanding of the reasons beyond (unsuccessful) action and improve their capacity to pass FBT. I also propose that dialogical ex-changes where people's behavior is explained by the attribution of (false) beliefs trigger the acquisition of explanatory capacities in the domain of folk psychology. Accordingly, language acquisition affects children's capacity to pass FBT because linguistic interaction in the social environment, and, in particular, specific dialogical exchanges where false beliefs are the matter of discussion, provide the main evidence necessary to them to pass FBT.

Before further discussion, let me introduce empirical evidence supporting my proposals. First of all, although I opposed de Villiers' claim that children younger than four lack the representational capacity to store others' (false) beliefs, her finding that the mastery of sentential complements is predictive of children's ability to pass FBT is in itself significant and requires explanation. According to my proposal, children must understand that false beliefs sometimes are good reasons for action before they can pass FBT. Now, sentential complements are the syntactic structures normally employed to report false belief

---

[21] A similar point has been suggested by Slors (2012) and Van Cleave & Gauker (2010).

attributions. It comes to a reason that children need to master complementation before they can pass FBT.

A second piece of relevant evidence comes from studies assessing the frequency of mental state lexicon in parental conversation. According to my proposal, understanding people's reasons to act is developed in specific dialogical exchanges where people's behavior is explained by the attribution of mental states. Accordingly, children who have more chances to take part to those dialogical exchanges should be expected to pass FBT earlier. On the contrary, finding that the amount of dialogical exchanges involving psychological discourse does not correlate with the ability to pass FBT would oppose my analysis.

Considering the empirical literature, many studies extensively showed that the frequency of mental terms in parental conversation predicts children's ability to pass FBT (Dunn *et al*. 1991; Furrow *et al*. 1992; Moore *et al*. 1994; Sabbagh & Callanan 1998; Ruffman *et al*. 2002; Meins *et al*. 2003; Dunn & Brophy 2005; Taumoepeau & Ruffman 2006). In addition, some evidence also suggests that this does not depend on the mere presence of mental lexicon in parental conversation, but on the quantity of discourse related to people's mental states even when mental states are not mentioned (Turnbull *et al*. 2008).

The third evidence for my proposal comes from studies about the quality of the conversation between the child and the caregiver. My proposal states that children should advance in their understanding of the mental domain proportionally to the quality of the conversation about the psychological domain they have with their caregivers. Accordingly, children whose caregivers tend to entertain more prolonged exchanges of such a kind and to provide more feedback should be expected to pass FBT earlier. On the contrary, evidence opposing my model would be that the caregiver's availability to converse with the child did not correlate with children's ability to pass FBT.

In the empirical literature, several indices have been advanced to assess the quality of parental conversation. Ontai & Thompson (2008) shaped an *elaborative discourse* index, which assesses the parental disposition to elaborate children's utterances by filling the gaps, providing explanations, and in general enriching the child's utterances. Similarly, Ensor & Hughes (2008) developed an index that they call *connectedness*, which assesses how much parental answers continue the child conversational contribution or whether they just push conversation further. Both studies found that those indices of the quality of the parental conversation correlate with the child's ability to pass FBT.[22]

My proposal is very close to Hutto's hypothesis that *folk psychological narratives* have a fundamental role in fostering "an understanding of the *forms* and *norms* of folk psychology" (Hutto 2007: 53), that is, "our everyday practice of making sense of intentional actions (i.e. our own and those of others) in terms of reasons" (Hutto 2009: 10). In particular, we share the same idea that dialogical interaction with the caregiver is the most important factor for the acquisition of

---

[22] While evidence reported above about the quantity of the conversational input was obtained through both correlational and transitional studies, evidence in this case is only correlational, therefore less significant. However, it still suggests that the more that adults are prone to elaborate children's utterances, the earlier the capacity to pass FBT is acquired.

the concept of belief while cognitive development only plays a minor role.

Although I am very sympathetic with Hutto's approach, there are nevertheless also substantial differences between his and my view. In particular, Hutto claims that "children's nuanced folk psychological skills only develop securely after ages four and five" (Hutto 2008: 26) and denies that passing FBT marks an important step in children's mastery of folk psychology.[23] Against this, I advance that passing FBT denotes an important improvement in children's mastery of folk psychology, because it marks their acquisition of an ability to explain people's behavior in folk psychological terms.

The dispute is partially theoretical and partially empirical. As for the theoretical facet, I believe that children's acquired capacity to explain others' reasons to act, which is manifested when they pass FBT, only denotes a minimal understanding of folk psychology, which needs time to be turned into a mature social competence. Therefore, Hutto does not really oppose my view when he claims that folk psychological skills fully develop only after age five or six.

However, against Hutto, I also advance a more specific empirical claim and propose that children start passing FBT because they learn to explain others' behavior by reporting their reasons to act. This makes a definitive claim about the timeline of children's acquisition of the capacity to pass FBT: Explanatory abilities in the domain of folk psychology should be acquired earlier than the predictive ability necessary to pass FBT. Therefore, my proposal would be supported by findings showing that explanatory capacities come in place earlier than the time children pass FBT. On the contrary, if it were found that children can pass FBT without still being able to express people's reasons to act, that would constitute opposing evidence to my model.

Referring to empirical evidence, several studies have already tested the correlation between traditional *predictive* FBT and a modified *explanatory* version, where children are asked to explain the behavior of a main character who just acted on the basis of a false belief. Many studies found that the explanatory version is as hard as the traditional one (Moses & Flavell 1990; Wimmer & Weichbold 1994; Wellman *et al*. 1996; Wimmer & Mayringer 1998; Perner *et al*. 2002; Atance & O'Neill 2004). This does not explicitly contradict my proposal, although it neither supports the presumed role of explanatory abilities in promoting the predictive abilities assessed by FBT. Notably, it nevertheless shows, against Hutto, that four-year-olds start manifesting important folk psychological competences: Not only did they correctly predict others' behavior that depends on the attribution of a false belief, but they also justify their predictions by correct explanations.

Evidence supporting my proposal comes instead by a few studies specifically finding that theory-of-mind predictive capacities are anticipated by explan-

---

[23] "There is a fairly widespread tendency to conflate the latter sort of ability [to understand and attribute beliefs] with a capacity to understand and attribute reasons. This mistake stems from assuming, as is commonly done, that children are already in the possession of the bulk of their theory of mind at the point at which they begin to pass false-belief tests. Hence, success on these tests is taken to be the mark of their having acquired the final piece of the theory of mind puzzle. Having mastered the core concept of belief, it is supposed that they have mastered the full set of folk psychological principles" (Hutto 2008: 25).

atory capacities (Bartsch & Wellman 1989; Bartsch *et al*. 2007). Careful looking at the methodology of these studies shows that they are flawed in the way in which they assessed children's psychological explanatory abilities.[24] Nevertheless, we should notice that finding a transitional period for the acquisition of abilities is always difficult: You can fail because either you look at too old children, or because you do not employ fine enough tools. Furthermore, all studies reported above always looked at explanatory ability as a yes-or-no competence and did not consider that there can be many levels of certainty in reporting one's reasons to act. Serious investigation instead would require keeping those levels separated. Future research, more respecting of the ecological validity of prompting answer methods and of the gradual acquisition of explanatory abilities in the domain of folk psychology, may bring clearer results about children's earlier capacity to pass explanatory rather than predictive versions of FBT.

Summarizing, predictive and correlational relations between children's capacity to pass FBT and (i) children's mastery of sentential complements, (ii) the quantity of parental conversation involving mental concepts, (iii) the quality of parental conversation, and (iv) children's explanatory capacities in the domain of folk psychology all support the claim that four-year-olds' capacity to pass FBT depends on their acquisition of a minimal capacity to report others' reasons to act.

This constitutes a significant improvement in our knowledge about social cognitive development and its triggering factors. It also leaves us in position to judge whether late social cognitive abilities fit the framework of the embodied mind. It follows indeed from my analysis that theory-of-mind capacities manifested by the ability to pass FBT are acquired by being engaged in a proper conversational context. The linguistic competence necessary to pass FBT is thus not localized and depends on the whole activity of a brain immersed in its natural and social (dialogical) environment. Accordingly, sensory and motor processes play a very peripheral role in the capacity to pass FBT: Late theory-of-mind abilities do not particularly fit the framework of embodied cognition.

Even though the capacity to pass FBT does not respect the strictest principles of embodiment, it is nevertheless compatible with the closer principles of situatedness and embeddedness. Indeed, according to my proposal, children's capacity to predict others' behavior depends on the mastery of an explanatory practice that children refine in conversation with their caregivers. This is a clear example of how the embeddedness of cognition in the child's social environment supports high-level cognitive processes such as social understanding.

6. Conclusions

Although theory of mind has been interpreted for a long as a unified capacity in

---

[24] Indeed, Bartsch & Wellman's (1989) method to prompt explanations was all but ecologically valid. Bartsch *et al*. (2007) used a more ecological prompting strategy, but their result depends on considering passers children that passed just one out of four explanatory false belief tasks: if at least two out of four trials are requested, their result is no longer significant.

the empirical literature, careful consideration of its development really demonstrates that it stands for a composed competence, which stands in a complex relation with the principle of embodiment. Available evidence does not exclude that early social cognitive abilities, which are manifested in spontaneous-response false belief tasks, depend on the activity of embodied cognitive processes. This challenges the mentalist view, according to which they must be interpreted in the terms of a capacity to attribute false beliefs.

Instead, late social cognitive abilities, such as the capacity to pass FBT, are the outcome of a process of enculturation: Children learn how to use at their own benefit and for predictive purposes the dialogical competence they have developed in conversation with their caregivers about others' reasons to act. This makes late social cognitive abilities not depending on the principles of embodied cognition. They are nonetheless compatible with it by falling within the borders of socially embedded cognition.

The present analysis acknowledges that embodied cognition indicates a 'unifying perspective' for psychology (Glenberg 2010). However, it suggests that embodiment alone is not sufficient to account for all forms of cognitive competences. Whereas the investigation of earlier forms of cognitive activity (e.g., infants' performance in spontaneous-response false belief tasks) requires pursuing research on the underlying embodied neural circuitries, expanding our knowledge about more advanced forms of comprehension (e.g., social understanding) needs to consider how social practices scaffold cognition and genuinely expand our cognitive competences.

With respect to the case of social understanding, a comprehensive explanation of the capacity to attribute mental states needs an analysis of the dialogical and social interaction between the child and the caregiver, which allows the former entering the 'community of minds' (Nelson 2009). The present analysis thus raises skeptical doubts about the empirical investigation of the neural circuitries underlying late social cognitive abilities such as the capacity to pass FBT (e.g., Saxe *et al.* 2004). Rather, it points to the study of the mechanics beyond dialogical exchanges (e.g., Pickering & Garrod 2004; Ruiter *et al.* 2006; de Ruiter *et al.* 2010) as a more promising field to start clarifying children's development of social understanding. Some research has already chosen this direction (Fernyhough 2008): It is my hope to have contributed to address further investigation along this path.

References

Anderson, Michael L. 2003. Embodied cognition: A field guide. *Artificial Intelligence* 149(1), 91–130.

Apperly, Ian A. 2010. *Mindreaders: The Cognitive Basis of 'Theory of Mind'*. Hove: Psychology Press.

Apperly, Ian A. & Stephen A. Butterfill. 2009. Do humans have two systems to track beliefs and belief-like states? *Psychological review* 116(4), 953–970.

Apperly, Ian A., Kevin J. Riggs, Andrew Simpson, Claudia Chiavarino & Dana Samson. 2006. Is belief reasoning automatic? *Psychological Science* 17(10), 841–844.

Astington, Janet Wilde & Jodie Alison Baird. 2005a. Introduction: Why language matters. In Janet Wilde Astington & Jodie Alison Baird (eds.), *Why Language Matters for Theory of Mind*, 3–25. New York: Oxford University Press.

Astington, Janet Wilde & Jodie Alison Baird (eds.). 2005b. *Why Language Matters for Theory of Mind*. New York: Oxford University Press.

Atance, Cristina M. & Daniela K. O'Neill. 2004. Acting and planning on the basis of a false belief: Its effects on 3-year-old children's reasoning about their own false beliefs. *Developmental Psychology* 40(6), 953–964.

Baillargeon, Renée, Rose M. Scott & Zijing He. 2010. False-belief understanding in infants. *Trends in Cognitive Sciences* 14(3), 110–118.

Baron-Cohen, Simon. 1994. The mindreading system: New directions for research. *Current Psychology of Cognition* 13, 724–750.

Baron-Cohen, Simon. 1995. *Mindblindness: An Essay on Autism and Theory of Mind*. Cambridge, MA: MIT Press.

Baron-Cohen, Simon, Alan M. Leslie & Uta Frith. 1985. Does the autistic child have a 'Theory of Mind'? *Cognition* 21(1), 37–46.

Barsalou, Lawrence W. 1999. Perceptual symbol systems. *Behavioral and Brain Sciences* 22(4), 577–660.

Barsalou, Lawrence W. 2005. Abstraction as dynamic interpretation in perceptual symbol systems. In Lisa Gershkoff-Stowe & David H. Rakison (eds.), *Building Object Categories in Developmental Time*, 389–431. Mahwah, NJ: Lawrence Erlbaum Associates.

Barsalou, Lawrence W. 2009. Simulation, situated conceptualization, and prediction. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364(1521), 1281–1289.

Bartsch, Karen, Michelle D. Campbell & Georgene L. Troseth. 2007. Why else does Jenny run? Young children's extended psychological explanations. *Journal of Cognition and Development* 8(1), 33–61.

Bartsch, Karen & Henry M. Wellman. 1989. Young children's attribution of action to beliefs and desires. *Child Development* 60(4), 946–964.

Berk, Laura. 1991. Children's private speech: An overview of theory and the status of research. In Rafael M. Diaz & Laura E. Berk (eds.), *Private Speech: From Social Interaction to Self-regulation*, 17–53. Hillsdale: Lawrence Erlbaum Associates.

Bigelow, Ann E. & Kevin Dugas. 2008. Relations among preschool children's understanding of visual perspective taking, false belief, and lying. *Journal of Cognition and Development* 9(4), 411–433.

Birch, Susan A. J. & Paul Bloom. 2003. Children are cursed: An asymmetric bias in mental-state attribution. *Psychological Science* 14(3), 283–286.

Birch, Susan A. J. & Paul Bloom. 2004. Understanding children's and adults' limitations in mental state reasoning. *Trends in Cognitive Sciences* 8(6), 255–260.

Birch, Susan A. J. & Paul Bloom. 2007. The curse of knowledge in reasoning about false beliefs. *Psychological Science* 18(5), 382–386.

Bloom, Paul & Tim P. German. 2000. Two reasons to abandon the false belief task as a test of Theory of Mind. *Cognition* 77(1), 25–31.

de Bruin, Leon & Albert Newen. 2012. An association account of false belief understanding. *Cognition* 123(2), 240–259.

de Bruin, Leon, Derek Strijbos & Marc Slors. 2011. Early social cognition: Alternatives to implicit mindreading. *Review of Philosophy and Psychology* 2(3), 499–517.

Buttelmann, David, Malinda Carpenter & Michael Tomasello. 2009. Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition* 112(2), 337–342.

Butterfill, Stephen A. & Ian A. Apperly. 2013. How to construct a minimal theory of mind. *Mind & Language*.

Call, Josep & Michael Tomasello. 1999. A nonverbal false belief task: The performance of children and great apes. *Child Development* 70(2), 381–395.

Carey, Susan & Elizabeth S. Spelke. 1996. Science and core knowledge. *Philosophy of Science* 63(4), 515–533.

Carlson, Stephanie M. & Louis J. Moses. 2001. Individual differences in inhibitory control and children's theory of mind. *Child Development* 72(4), 1032–1053.

Carlson, Stephanie M., Antoinette Wong, Margaret Lemke & Caron Cosser. 2005. Gesture as a window on children's beginning understanding of false belief. *Child Development* 76(1), 73–86.

Chemero, Anthony. 2009. *Radical Embodied Cognitive Science*. Cambridge, MA: MIT Press.

Cheung, Him, Hsuan-Chih Chen & William Yeung. 2009. Relations between mental verb and false belief understanding in Cantonese-speaking children. *Journal of Experimental Child Psychology* 104(2), 141–155.

Cheung, Him, Chen Hsuan-Chih, Nikki Creed, Lisa Ng, Sui Ping Wang & Lei Mo. 2004. Relative roles of general and complementation language in theory-of-mind development: Evidence from Cantonese and English. *Child Development* 75(4), 1155–1170.

Clark, Andy. 1997. *Being There: Putting Brain, Body, and World Together Again*. Cambridge, MA: MIT Press.

Clark, Andy. 1998. Magic words: How language augments human computation. In Peter Carruthers & Jill Boucher (eds.), *Language and Thought: Interdisciplinary Themes*, 162–183. Cambridge: Cambridge University Press.

Clark, Andy. 2008. *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*, 1st edn. Oxford: Oxford University Press.

Clark, Andy & Annette Karmiloff-Smith. 1993. The cognizer's innards: A psychological and philosophical perspective on the development of thought. *Mind & Language* 8(4), 487–519.

Van Cleave, Matthew & Christopher Gauker. 2010. Linguistic practice and false-belief tasks. *Mind & Language* 25(3), 298–328.

Clements, Wendy A. & Josef Perner. 1994. Implicit understanding of belief. *Cognitive Development* 9(4), 377–395.

Colonnesi, Cristina, Carolien Rieffe, Willem Koops & Paola Perucchini. 2008. Precursors of a theory of mind: A longitudinal study. *British Journal of Developmental Psychology* 26(4), 561–577.

Coltheart, Max. 1999. Modularity and cognition. *Trends in Cognitive Sciences* 3(3), 115–120.

Csibra, Gergely. 2007. Action mirroring and action interpretation: An alternative account. In P. Haggard, Y. Rosetti & M. Kawato (eds.), *Sensorimotor Foundations of Higher Cognition. Attention and Performance XXII*, 435–459. Oxford: Oxford University Press.

Csibra, Gergely & György Gergely. 2007. 'Obsessed with goals': Functions and mechanisms of teleological interpretation of actions in humans. *Acta Psychologica* 124(1), 60–78.

Dawkins, Richard. 1976. *The Selfish Gene*. New York: Oxford University Press.

Dunn, Judy & M. Brophy. 2005. Communication, relationships, and individual differences in children's understanding of mind. In Janet Wilde Astington & Jodie Alison Baird (eds.), *Why Language Matters for Theory of Mind*, 50–69. New York: Oxford University Press.

Dunn, Judy, Jane R. Brown & Lynn Beardsall. 1991. Family talk about feeling states and children's later understanding of others' emotions. *Developmental Psychology* 27(3), 448–455.

Ensor, Rosie & Claire Hughes. 2008. Content or connectedness? Mother–child talk and early social understanding. *Child Development* 79(1), 201–216.

Farrant, Brad M., Janet Fletcher & Murray T. Maybery. 2006. Specific language impairment, theory of mind, and visual perspective taking: Evidence for simulation theory and the developmental role of language. *Child Development* 77(6), 1842–1853.

Fernyhough, Charles. 2008. Getting Vygotskian about theory of mind: Mediation, dialogue, and the development of social understanding. *Developmental Review* 28(2), 225–262.

Figueras-Costa, Berta & Paul L. Harris. 2001. Theory of Mind development in deaf children: A nonverbal test of false-belief understanding. *Journal of Deaf Studies and Deaf Education* 6(2), 92–102.

Fodor, Jerry A. 1983. *The Modularity of Mind*. Cambridge, MA: MIT Press.

Fodor, Jerry A. 1992. A theory of the child's Theory of Mind. *Cognition* 44(3), 283–296.

Frith, Chris D. & Uta Frith. 2006. How we predict what other people are going to do. *Brain Research* 1079(1), 36–46.

Fuchs, Thomas & Hanne De Jaegher. 2009. Enactive intersubjectivity: Participatory sense-making and mutual incorporation. *Phenomenology and the Cognitive Sciences* 8(4), 465–486.

Furrow, David, Chris Moore, Jane Davidge & Lorraine Chiasson. 1992. Mental terms in mothers' and children's speech: Similarities and relationships. *Journal of Child Language* 19(3), 617–631.

Gallagher, Shaun. 2007. Simulation trouble. *Social Neuroscience* 2(3-4), 353–365.

Gallagher, Shaun. 2008. Direct perception in the intersubjective context. *Consciousness and Cognition* 17(2), 535–543.

Gallagher, Shaun. 2011. Strong interaction and self-agency. *Humana.Mente* 15, 55–76.

Gallagher, Shaun & Daniel Povinelli. 2012. Enactive and behavioral abstraction accounts of social understanding in chimpanzees, infants, and adults.

Review of Philosophy and Psychology 3(1), 145–169.

Gallese, Vittorio. 2005. Embodied simulation: From neurons to phenomenal experience. *Phenomenology and the Cognitive Sciences* 4(1), 23–48.

Gallese, Vittorio. 2007. Before and below 'theory of mind': Embodied simulation and the neural correlates of social cognition. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 362(1480), 659–669.

Gallese, Vittorio & Alvin I. Goldman. 1998. Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences* 2(12), 493–501.

Gallese, Vittorio & Corrado Sinigaglia. 2011. What is so special about embodied simulation? *Trends in Cognitive Sciences* 15(11), 512–519.

Garfield, Jay L., Candida C. Peterson, B. Garson, A. Nevin & Tricia Perry. 2009. Let's pretend! The role of pretense in the acquisition of theory of mind. Ms., Smith College and University of Tasmania.

Garfield, Jay L., Candida C. Peterson & Tricia Perry. 2001. Social cognition, language acquisition and the development of the Theory of Mind. *Mind & Language* 16(5), 494–541.

Garnham, Wendy A. & Ted Ruffman. 2001. Doesn't see, doesn't know: Is anticipatory looking really related to understanding or belief? *Developmental Science* 4(1), 94–100.

Gibson, James Jerome. 1979. *The Ecological Approach to Visual Perception*. Mahwah, NJ: Houghton Mifflin Harcourt.

Del Giudice, Marco, Valeria Manera & Christian Keysers. 2009. Programmed to learn? The ontogeny of mirror neurons. *Developmental science* 12(2), 350–363.

Glenberg, Arthur M. 2010. Embodiment as a unifying perspective for psychology. *Wiley Interdisciplinary Reviews: Cognitive Science* 1(4), 586–596.

Goldman, Alvin I. 2006. *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading*. New York: Oxford University Press.

Goldman, Alvin I. 2009. Mirroring, mindreading, and simulation. In Jaime A. Pineda (ed.), *Mirror Neuron Systems*, 1–20. Totowa, NJ: Humana Press.

Goldman, Alvin I. & Frederique de Vignemont. 2009. Is social cognition embodied? *Trends in Cognitive Sciences* 13(4), 154–159.

Gopnik, Alison. 1990. Developing the idea of intentionality: Children's theories of mind. *Canadian Journal of Philosophy* 20(1), 89–113.

Gopnik, Alison. 1996. The scientist as child. *Philosophy of Science* 63(4), 485–514.

Gopnik, Alison, V. Slaughter & Andrew N. Meltzoff. 1994. Changing your views: How understanding visual perception can lead to a new theory of the mind. In Charlie Lewis & Peter Mitchell (eds.), *Children's Early Understanding of Mind: Origins and Development*, 157–181. Hove: Erlbaum.

Gordon, Robert M. 1986. Folk psychology as simulation. *Mind & Language* 1(2), 158–171.

Gordon, Robert M. 2007. Ascent routines for propositional attitudes. *Synthese* 159(2), 151–165.

Grafton, Scott T. 2009. Embodied cognition and the simulation of action to understand others. *Annals of the New York Academy of Sciences* 1156(1), 97–117.

Grush, Rick. 2004. The emulation theory of representation: Motor control, imagery, and perception. *Behavioral and Brain Sciences* 27(3), 377–442.

Guajardo, Nicole R., Jessica Parker & Kandi Turley-Ames. 2009. Associations among false belief understanding, counterfactual reasoning, and executive function. *The British Journal of Developmental Psychology* 27(3), 681–702.

Hale, Courtney Melinda & Helen Tager-Flusberg. 2003. The influence of language on theory of mind: A training study. *Developmental Science* 6(3), 346–359.

Happé, Francesca G. E. 1995. The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child Development* 66(3), 843–855.

Harris, Paul L. 1989. *Children and Emotion: The Development of Psychological Understanding*. Oxford: Wiley-Blackwell.

Harris, Paul L. 1992. From simulation to folk psychology: The case for development. *Mind & Language* 7(1-2), 120–144.

Heal, Jane. 1986. Replication and functionalism. In Jeremy Butterfill (ed.), *Language, Mind and Logic*, 135–150. Cambridge: Cambridge University Press.

Heal, Jane. 1998. Co-cognition and off-line simulation: Two ways of understanding the simulation approach. *Mind & Language* 13, 477–498.

Hutchins, Edwin. 1995a. *Cognition in the Wild*. Cambridge, MA: MIT Press.

Hutchins, Edwin. 1995b. How a cockpit remembers its speeds. *Cognitive Science* 19(3), 265–288.

Hutto, Daniel D. 2007. The narrative practice hypothesis: origins and applications of folk psychology. *Royal Institute of Philosophy Supplement* 60, 43–68.

Hutto, Daniel D. 2008. *Folk Psychological Narratives*. Cambridge, MA: MIT Press.

Hutto, Daniel D. (ed.). 2009. Narrative and folk psychology. *Journal of Consciousness Studies* 16(6-8).

Jacques, Sophie & Philip David Zelazo. 2005. Language and the development of cognitive flexibility: Implications for theory of mind. In Janet Wilde Astington & Jodie Alison Baird (eds.), *Why Language Matters for Theory of Mind*, 573–591. New York: Oxford University Press.

De Jaegher, Hanne. 2009. Social understanding through direct perception? Yes, by interacting. *Consciousness and Cognition* 18(2), 535–542; discussion 543–550.

Kagan, Jerome. 2008. In defense of qualitative changes in development. *Child Development* 79(6). 1606–1624.

Karmiloff-Smith, Annette. 1992. *Beyond Modularity: A Developmental Perspective on Cognitive Science*. Cambridge, MA: MIT Press.

Keysar, Boaz, Dale J. Barr, Jennifer A. Balin & Jason S. Brauner. 2000. Taking perspective in conversation: The role of mutual knowledge in comprehension. *Psychological Science* 11(1), 32–38.

Keysar, Boaz, Shuhong Lin & Dale J. Barr. 2003. Limits on theory of mind use in adults. *Cognition* 89(1), 25–41.

Kirsh, David. 1995. The intelligent use of space. *Artificial Intelligence* 73(1-2), 31–68.

Kiverstein, Julian & Andy Clark. 2009. Introduction: Mind embodied, embedded, enacted: One church or many? *Topoi* 28(1), 1–7.

Knudsen, Birgit. 2011. Infants' appreciation of others' beliefs in prelinguistic communication: A second person approach to mindreading. Paper presented at the *Other Minds: Embodied Interaction and Higher-Order Reasoning*, Bochum.

Kovács, Ágnes Melinda, Ernő Téglás & Ansgar Denis Endress. 2010. The social sense: Susceptibility to others' beliefs in human infants and adults. *Science* 330(6012), 1830–1834.

Lakoff, George & Mark Johnson. 1999. *Philosophy In The Flesh: The Embodied Mind And Its Challenge To Western Thought*. New York: Basic Books.

Latour, Bruno. 1986. Visualisation and cognition: Thinking with eyes and hands. In H. Kuklick (ed.), *Knowledge and Society Studies in the Sociology of Culture Past and Present*, vol. 6, 1–40. Bingley: Jai Press.

Leekam, Susan R. & Josef Perner. 1991. Does the autistic child have a metarepresentational deficit? *Cognition* 40(3), 203–218.

Leekam, Susan R., Josef Perner, Laura Healey & Claire Sewell. 2008. False signs and the non-specificity of theory of mind: Evidence that preschoolers have general difficulties in understanding representations. *British Journal of Developmental Psychology* 26, 485–497.

Leslie, Alan M. 1994. ToMM, ToBy, and agency: Core architecture and domain specificity in cognition and culture. In Lawrence A. Hirschfeld & Susan A. Gelman (eds.), *Mapping the Mind: Domain Specificity in Cognition and Culture*, 119–148. New York: Cambridge University Press.

Leslie, Alan M. 1995. A theory of agency. In Dan Sperber, David Premack & Ann James Premack (eds.), *Causal Cognition: A Multidisciplinary Debate*, 121–149. Oxford: Clarendon Press.

Leslie, Alan M. 2005. Developmental parallels in understanding minds and bodies. *Trends in Cognitive Sciences* 9(10), 459–462.

Leslie, Alan M., Tim P. German & P. Polizzi. 2005. Belief–desire reasoning as a process of selection. *Cognitive Psychology* 50(1), 45–85.

Lewis, Charlie, Masuo Koyasu, Seungmi Oh, Ayako Ogawa, Benjamin Short & Zhao Huang. 2009. Culture, executive function, and social understanding. *New Directions for Child and Adolescent Development* 2009(123), 69–85.

Lillard, Angeline. 1998. Ethnopsychologies: Cultural variations in theories of mind. *Psychological Bulletin* 123, 3–32.

Liu, David, Henry M. Wellman, Twila Tardif & Mark A. Sabbagh. 2008. Theory of mind development in Chinese children: A meta-analysis of false-belief understanding across cultures and languages. *Developmental Psychology* 44(2), 523–531.

Lohmann, Heidemarie & Michael Tomasello. 2003. The role of language in the development of false belief understanding: A training study. *Child Development* 74(4), 1130–1144.

Lohmann, Heidemarie, Michael Tomasello & Sonja Meyer. 2005. Linguistic communication and social understanding. In Janet Wilde Astington & Jodie Alison Baird (eds.), *Why Language Matters for Theory of Mind*, 245–265. New York: Oxford University Press.

McGann, Marek & Hanne De Jaegher. 2009. Self–other contingencies: Enacting social perception. *Phenomenology and the Cognitive Sciences* 8(4), 417–437.

Meins, Elizabeth, Charles Fernyhough, Rachel Wainwright, David Clark-Carter, Mani Das Gupta, Emma Fradley & Michelle Tuckey. 2003. Pathways to understanding mind: Construct validity and predictive validity of maternal mind-mindedness. *Child Development* 74(4), 1194–1211.

Milligan, Karen, Janet Wilde Astington & Lisa Ain Dack. 2007. Language and theory of mind: Meta-analysis of the relation between language ability and false-belief understanding. *Child Development* 78(2), 622–646.

Moore, Chris, David Furrow, Lorraine Chiasson & Maria Patriquin. 1994. Developmental relationships between production and comprehension of mental terms. *First Language* 14(42-43), 1–17.

Moses, Louis J. & John H. Flavell. 1990. Inferring false beliefs from actions and reactions. *Child Development* 61(4), 929–945.

Müller, Ulrich, Philip David Zelazo & Sarah Imrisek. 2005. Executive function and children's understanding of false belief: How specific is the relation? *Cognitive Development* 20(2), 173–189.

Nelson, Katherine. 2009. Narrative and folk psychology. *Journal of Consciousness Studies* 16(6-8), 69–93.

Noë, Alva. 2005. *Action in Perception*. Cambridge, MA: MIT Press.

Oh, Seungmi & Charlie Lewis. 2008. Korean preschoolers' advanced inhibitory control and its relation to other executive skills and mental state understanding. *Child Development* 79(1), 80–99.

Onishi, Kristine H. & Renée Baillargeon. 2005. Do 15-month-old infants understand false beliefs? *Science* 308(5719), 255–258.

Ontai, Lenna L. & Ross A. Thompson. 2008. Attachment, parent–child discourse and theory-of-mind development. *Social Development* 17(1), 47–60.

Van Overwalle, Frank. 2009. Social cognition and the brain: A meta-analysis. *Human Brain Mapping* 30(3), 829–858.

Ozonoff, Sally. 1995. Reliability and validity of the Wisconsin card sorting test in studies of autism. *Neuropsychology* 9(4), 491–500.

Ozonoff, Sally, Bruce F. Pennington & Sally J. Rogers. 1991. Executive function deficits in high-functioning autistic children: Relationship to theory of mind. *Journal of Child Psychology and Psychiatry* 32(7), 1081–1105.

Paivio, Allan. 1986. *Mental Representations: A Dual Coding Approach*. New York: Oxford University Press.

Penn, Derek C. & Daniel J. Povinelli. 2007. On the lack of evidence that non-human animals possess anything remotely resembling a 'theory of mind'. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 362(1480), 731–744.

Perner, Josef. 1991. *Understanding the Representational Mind*. Cambridge, MA: MIT Press.

Perner, Josef. 1995. The many faces of belief: Reflections on Fodor's and the child's theory of mind. *Cognition* 57(3), 241–269.

Perner, Josef. 2010. Who took the cog out of cognitive science? Mentalism in an era of anti-cognitivism. In Peter A. Frensch & Ralf Schwarzer (eds.), *Cognition and Neuropsychology International Perspectives on Psychological Science*, vol. 1, 241–261. Hove: Psychology Press.

Perner, Josef, Birgit Lang & Daniela Kloo. 2002. Theory of mind and self-control: More than a common problem of inhibition. *Child Development* 73(3), 752–767.

Perner, Josef & Ted Ruffman. 2005. Infants' insight into the mind: How deep? *Science* 308(5719), 214–216.

Perner, Josef, Manuel Sprung, Petra Zauner & Hubert Haider. 2003. Want that is understood well before say that, think that, and false belief: A test of de Villiers's linguistic determinism on German-speaking children. *Child Development* 74(1), 179–188.

Peterson, Candida C. & Michael Siegal. 2000. Insights into theory of mind from deafness and autism. *Mind & Language* 15(1), 123–145.

Pickering, Martin J. & Simon Garrod. 2004. Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences* 27(02), 169–190.

Poulin-Dubois, Diane, Ivy Brooker & Virginia Chow. 2009. The developmental origins of naïve psychology in infancy. In P. J Bauer (ed.), *Advances in Child Development and Behavior*, vol. 37, 55–104. San Diego, CA: Elsevier.

Premack, David & Guy Woodruff. 1978. Does the Chimpanzee Have a Theory of Mind? *Behavioral and Brain Sciences* 1(04), 515–526.

Pyers, Jennie E. & Ann Senghas. 2009. Language promotes false-belief understanding: Evidence from learners of a new sign language. *Psychological Science* 20(7), 805–812.

Ratcliffe, Matthew. 2007. *Rethinking Commonsense Psychology: A Critique of Folk Psychology, Theory of Mind and Simulation*, 1st edn. Basingstoke: Palgrave Macmillan.

Rizzolatti, Giacomo & Laila Craighero. 2004. The mirror-neuron system. *Annual Review of Neuroscience* 27, 169–192.

Ruffman, Ted, Wendy A. Garnham, Arlina Import & Dan Connolly. 2001. Does eye gaze indicate implicit knowledge of false belief? Charting transitions in knowledge. *Journal of Experimental Child Psychology* 80(3), 201–224.

Ruffman, Ted, Lance Slade & Elena Crowe. 2002. The relation between children's and mothers' mental state language and theory-of-mind understanding. *Child Development* 73(3), 734–751.

de Ruiter, J. P., Matthijs L. Noordzij, Sarah Newman-Norlund, Roger Newman-Norlund, Peter Hagoort, Stephen C. Levinson & Ivan Toni. 2010. Exploring the cognitive infrastructure of communication. *Interaction Studies* 11(1), 51–77.

Ruiter, Jan-Peter, Holger Mitterer & N. J. Enfield. 2006. Projecting the end of a speaker's turn: A cognitive cornerstone of conversation. *Language* 82(3), 515–535.

Sabbagh, Mark A., Jeannette E. Benson & Valerie Kuhlmeier. 2010. False belief understanding in infants and preschoolers. In M. Bornstein & M. Legerstee (eds.), *The Developing Infant Mind: Integrating Biology and Experience*. Toronto: Guilford Press.

Sabbagh, Mark A. & Maureen A. Callanan. 1998. Metarepresentation in action: 3-, 4-, and 5-year-olds' developing theories of mind in parent-child conversations. *Developmental Psychology* 34(3), 491–502.

Sabbagh, Mark A., Louis J. Moses & Sean Shiverick. 2006. Executive functioning and preschoolers' understanding of false beliefs, false photographs, and false signs. *Child Development* 77(4), 1034–1049.

Samson, Dana & Ian A. Apperly. 2010. There is more to mind reading than having theory of mind concepts: New directions in theory of mind research. *Infant and Child Development* 19, 443–454.

Samson, Dana, Ian A. Apperly, Jason J. Braithwaite, Benjamin J. Andrews & Sarah E. Bodley Scott. 2010. Seeing it their way: Evidence for rapid and involuntary computation of what other people see. *Journal of Experimental Psychology. Human Perception and Performance* 36(5), 1255–1266.

Saxe, Rebecca, Susan Carey & N. Kanwisher. 2004. Understanding other minds: Linking developmental psychology and functional neuroimaging. *Annual Review of Psychology* 55(1), 87–124.

Scerif, Gaia & Annette Karmiloff-Smith. 2005. The dawn of cognitive genetics? Crucial developmental caveats. *Trends in Cognitive Sciences* 9(3), 126–135.

Schick, Brenda, Peter A. de Villiers, Jill G. de Villiers & Robert Hoffmeister. 2007. Language and theory of mind: A study of deaf children. *Child Development* 78(2), 376–396.

Scholl, Brian J. & Alan M. Leslie. 1999. Modularity, development and 'theory of mind'. *Mind & Language* 14(1), 131–153.

Scott, Rose M. & Renée Baillargeon. 2009. Which penguin is this? Attributing false beliefs about object identity at 18 Months. *Child development* 80(4), 1172–1196.

Senju, Atsushi. 2011. Spontaneous theory of mind and its absence in autism spectrum disorders. *Neuroscientist* 18, 108–113.

Senju, Atsushi, Victoria Southgate, Yui Miura, Tomoko Matsui, Toshikazu Hasegawa, Yoshikuni Tojo, Hiroo Osanai & Gergely Csibra. 2010. Absence of spontaneous action anticipation by false belief attribution in children with autism spectrum disorder. *Development and Psychopathology* 22(02), 353–360.

Senju, Atsushi, Victoria Southgate, Sarah White & Uta Frith. 2009. Mindblind eyes: An absence of spontaneous theory of mind in asperger syndrome. *Science* 325(5942). 883–885.

Shahaeian, Ameneh, Candida C. Peterson, Virginia Slaughter & Henry M. Wellman. 2011. Culture and the sequence of steps in theory of mind development. *Developmental Psychology* 47(5), 1239–1247.

Shapiro, Lawrence A. 2011. *Embodied Cognition*. London: Taylor & Francis.

Slors, Marc. 2012. The model-model of the theory-theory. *Inquiry* 55(5), 521–542.

Smith, Mark, Ian A. Apperly & Victoria White. 2003. False belief reasoning and the acquisition of relative clause sentences. *Child Development* 74(6), 1709–1719.

Sodian, Beate. 2011. Theory of mind in infancy. *Child Development Perspectives* 5(1), 39–43.

Song, Hyun-joo & Renée Baillargeon. 2008. Infants' reasoning about others' false perceptions. *Developmental Psychology* 44(6), 1789–1795.

Song, Hyun-joo, Kristine H. Onishi, Renée Baillargeon & Cynthia Fisher. 2008. Can an agent's false belief be corrected by an appropriate communication? Psychological reasoning in 18-month-old infants. *Cognition* 109(3), 295–315.

Southgate, Victoria, Coralie Chevallier & Gergely Csibra. 2010. Seventeen-month-olds appeal to false beliefs to interpret others' referential communication. *Developmental Science* 13, 907–912.

Southgate, Victoria, Atsushi Senju & Gergely Csibra. 2007. Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science* 18, 587–592.

Spaulding, Shannon. 2010. Embodied Cognition and Mindreading. *Mind & Language* 25(1), 119–140.

Steels, Luc & Rodney Allen Brooks. 1995. *The Artificial Life Route to Artificial Intelligence: Building Embodied, Situated Agents*. Mahwah, NJ: Routledge.

Surian, Luca, Stefania Caldi & Dan Sperber. 2007. Attribution of beliefs by 13-month-old infants. *Psychological Science* 18(7), 580–586.

Surtees, Andrew D. R. & Ian A. Apperly. 2012. Egocentrism and automatic perspective taking in children and adults. *Child Development* 83(2), 452–460.

Surtees, Andrew D. R., Stephen A. Butterfill & Ian A. Apperly. 2011. Direct and indirect measures of level-2 perspective-taking in children and adults. *British Journal of Developmental Psychology* 30, 75–86.

Tager-Flusberg, Helen & Robert M. Joseph. 2005. How language facilitates the acquisition of false belief in children with autism. In Janet Wilde Astington & Jodie Alison Baird (eds.), *Why Language Matters for Theory of Mind*, 298–318. New York: Oxford University Press.

Taumoepeau, Mele & Ted Ruffman. 2006. Mother and infant talk about mental states relates to desire language and emotion understanding. *Child Development* 77(2), 465–481.

Thelen, E. & Linda Smith. 1994. *A Dynamic Systems Approach to the Development of Cognition and Action*. Cambridge, MA: MIT Press.

Träuble, Birgit, Vesna Marinović & Sabina Pauen. 2010. Early theory of mind competencies: Do infants understand others' beliefs? *Infancy* 15(4), 434–444.

Turnbull, William, Jeremy I. M. Carpendale & Timothy P. Racine. 2008. Relations between mother-child talk and 3- to 5 year-old children's understanding of belief: Beyond mental state terms to talk about the mind. *Merrill-Palmer Quarterly* 54(3), 367–385.

Varela, Francisco J., Evan Thompson & Eleanor Rosch. 1991. *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT Press.

de Villiers, Jill G. 2005. Can language acquisition give children a point of view? In Janet Wilde Astington & Jodie Alison Baird (eds.), *Why Language Matters for Theory of Mind*, 186–219. New York: Oxford University Press.

de Villiers, Jill G. 2007. The interface of language and theory of mind. *Lingua* 117(11), 1858–1878.

de Villiers, Jill G. 2009. Complements enable representation of the contents of false beliefs: The evolution of a theory of theory of mind. In Susan Foster-Cohen (ed.), *Language Acquisition*. Basingstoke: Palgrave Macmillan.

de Villiers, Jill G. & Jennie E. Pyers. 2002. Complements to cognition: A longitudinal study of the relationship between complex syntax and false-belief-understanding. *Cognitive Development* 17(1), 1037–1060.

de Villiers, Jill G. & Peter A. de Villiers. 2000. Linguistic determinism and the understanding of false beliefs. In Peter Mitchell & Kevin Riggs (eds.), *Children's Reasoning and the Mind*, 191–228. Hove: Psychology Press.

de Villiers, Jill G. & Peter A. de Villiers. 2003. Language for thought: Coming to understand false beliefs. In Dedre Gentner & Susan Goldin-Meadow (eds.), *Language in Mind: Advances in the Study of Language and Thought*, 335–384. Cambridge, MA: MIT Press.

de Villiers, Peter A. 2005. The role of language in theory of mind development: What deaf children tell us. In Janet Wilde Astington & Jodie Alison Baird (eds.), *Why Language Matters for Theory of Mind*, 266–297. New York: Oxford University Press.

Vinden, Penelope G. 1999. Children's understanding of mind and emotion: A multi-culture study. *Cognition & Emotion* 13(1), 19–48.

Vygotsky, Lev Semenovich. 1934. *Thought and Language*. Cambridge, MA: MIT Press.

Vygotsky, Lev Semenovich. 1978. *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, MA: Harvard University Press.

Wellman, Henry M. 2002. Understanding the psychological world: Developing a theory of mind. *Blackwell Handbook of Childhood Cognitive Development*, 167–187. Oxford: Wiley-Blackwell.

Wellman, Henry M., David Cross & Julanne Watson. 2001. Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development* 72(3), 655–684.

Wellman, Henry M. & Susan A. Gelman. 1997. Knowledge acquisition. In William Damon & Nancy Eisenberg (eds.), *Handbook of Child Psychology*. New York: Wiley.

Wellman, Henry M., Michelle Hollander & Carolyn A. Schult. 1996. Young children's understanding of thought bubbles and of thoughts. *Child Development* 67(3), 768–788.

Wellman, Henry M. & David Liu. 2004. Scaling of theory-of-mind tasks. *Child Development* 75, 523–541.

Wellman, Henry M., Fuxi Fang, David Liu, Liqi Zhu & Guoxiong Liu. 2006. Scaling of theory-of-mind understandings in Chinese children. *Psychological Science* 17(12), 1075–1081.

Wilson, Margaret. 2002. Six views of embodied cognition. *Psychonomic Bulletin & Review* 9(4), 625–636.

Wilson, Margaret & Günther Knoblich. 2005. The case for motor involvement in perceiving conspecifics. *Psychological Bulletin* 131(3), 460–473.

Wimmer, Heinz & Heinz Mayringer. 1998. False belief understanding in young children: Explanations do not develop before predictions. *International Journal of Behavioral Development* 22(2), 403–422.

Wimmer, Heinz & Josef Perner. 1983. Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition* 13(1), 103–128.

Wimmer, Heinz & Viktor Weichbold. 1994. Children's theory of mind: Fodor's heuristics examined. *Cognition* 53(1), 45–57.

Winsler, Adam, Jesus René De Leon, Beverly A. Wallace, Martha P. Carlton & Angela Willson-Quayle. 2003. Private speech in preschool children: Developmental stability and change, across-task consistency, and relations with classroom behaviour. *Journal of Child Language* 30(03), 583–608.

Woolfe, Tyron, Stephen C. Want & Michael Siegal. 2002. Signposts to development: Theory of mind in deaf children. *Child Development* 73(3), 768–778.

Yamaguchi, Mariko, Valerie A. Kuhlmeier, Karen Wynn & Kristy van Marle.

2009. Continuity in social cognition from infancy to childhood. *Developmental Science* 12(5), 746–752.

Yeh, Wenchi & Lawrence W. Barsalou. 2006. The situated nature of concepts. *The American Journal of Psychology* 119(3), 349–384.

Zelazo, Philip David, Stephanie M. Carlson & A. Kesek. 2008. The development of executive function in childhood. In C. Nelson & M. Luciana (eds.), *Handbook of Developmental Cognitive Neuroscience*, 553–574. Cambridge, MA: MIT Press.

Marco Fenici
Università degli Studi di Firenze
Department of Philosophy
Via Bolognese, 52
50139 Florence
Italy
mfenici@hotmail.com

Cortical Motor Organization, Mirror Neurons, and Embodied Language: An Evolutionary Perspective

Leonardo Fogassi & Pier Francesco Ferrari

The recent conceptual achievement that the cortical motor system plays a crucial role not only in motor control but also in higher cognitive functions has given a new perspective also on the involvement of motor cortex in language perception and production. In particular, there is evidence that the matching mechanism based on mirror neurons can be involved in both phonological recognition and retrieval of meaning, especially for action word categories, thus suggesting a contribution of an action–perception mechanism to the automatic comprehension of semantics. Furthermore, a comparison of the anatomo-functional properties of the frontal motor cortex among different primates and their communicative modalities indicates that the combination of the voluntary control of the gestural communication systems and of the vocal apparatus has been the critical factor in the transition from a gestural-based communication into a predominantly speech-based system. Finally, considering that the monkey and human premotor-parietal motor system, plus the prefrontal cortex, are involved in the sequential motor organization of actions and in the hierarchical combination of motor elements, we propose that elements of such motor organization have been exploited in other domains, including some aspects of the syntactic structure of language.

Keywords: action; Broca's area; gestures; matching mechanism; monkey

1. Introduction

The aim of this article is to show the strong link existing between the motor system and language, with a particular emphasis on the relationship between the mirror neuron mechanism, thought to be involved in action understanding, and the mechanism underlying language comprehension. In order to address these issues, we have organized the current review as follows. First, we will describe the organization of the cortical motor system and how cognitive functions, and

---

We want to thank D. Mallamo for his help in preparing the figures. This work has been supported by the Italian MIUR (Prot. 2008J7YFNR_003), the Italian Institute of Technology (RTM), the European grant NEST-2004-Path-HUM (HANDTOMOUTH), and the P01 HD 064653-01 of NICHD (Functions and Development of the Mirror Neuron System).

biolinguistics

ISSN 1450–341
Biolinguistics 6.3–4: 308–337, 2012
http://www.biolinguistics.eu

more specifically, action understanding, derive from this organization. Second, we will review the evidence suggesting that the mirror matching mechanism can apply to phonological matching. Third, we will briefly describe vocal and gestural communication systems in non-human primates, suggesting that their combination at voluntary level can have played an important role in language evolution. Fourth, we will present the anatomical and neurophysiological evidence that suggests a homology between higher order monkey motor areas and frontal regions involved in language processing. We will conclude suggesting the possible link between sequential motor behavior and syntactic structure. Although several issues of the debate concerning the role of the motor system in speech perception and language require a deeper elaboration of the different perspectives, we deliberately focus on the motor system and the action–perception mechanisms to emphasize their central contribution in different aspects of perception processes occurring in language and to challenge a theoretical position in linguistics and cognitive sciences that considers sensory and motor information processes as separate domains.

2. The Organization of the Motor System and the Emergence of Motor Cognitive Functions

In the traditional way of conceiving brain processes, perception was considered the result of a higher order elaboration of sensory information, occurring in the posterior half of the brain. In this view, when we need to act on the external world, the outcome of this elaboration is fed to the anterior part of the cortex, the prefrontal and motor cortex, in order to plan actions and execute them. Thus, the posterior part of the cortex would have a main role in functions such as space perception and object perception, plus other cognitive functions, also considered as higher order elaboration of sensory functions, such as language comprehension, music, reasoning, memory, while the anterior part would simply guide our Behaviour. Although philosophers, psychologists, and neurophysiologists (Piaget 1951; Merleau-Ponty 1962; Jeannerod 1988) pointed to the motor system as to a fundamental tool for our knowledge of the world, for many years, this system was mainly considered in its executive aspects, such as, for example, movement parameters (Evarts 1968; Georgopoulos *et al*. 1982), or, at most, motor preparation (Weinrich *et al*. 1984). Interestingly, however, some researchers demonstrated that the motor cortex activates during motor imagery, which can be considered a mental function (Roland *et al*. 1980; Jeannerod 1994). Despite this evidence, the serial flow of information, i.e. from the so called associative areas of the temporal and parietal cortex to frontal areas, was not disputed, and also many computational models were organized according to this view (Poggio & Edelman 1990; Giese & Poggio 2003). In the last two decades the conceptualization about action, perception and cognitive functions radically changed. The two most important new concepts that emerged are the assignment to the motor system of a crucial role in cognitive functions and, linked with it, the recognition of its strong importance in perception. These conceptual changes have been possible due to the neurophysiological, neuroanatomical and psychological findings achieved in the last two decades. Neurophysiological experiments in the monkey showed that the motor cortex code the goal of motor acts and contain a storage of high level motor representations. Neuroanatomical studies showed that parietal and frontal cortices are linked by reciprocal connections (Rizzolatti & Luppino 2001), thus indicating that these cortical regions have both motor and sensory properties and that motor and perceptual aspects are strictly integrated and influence each other. Finally, psychophysics studies in humans strengthen this suggestion, showing that not only higher order sensory elaboration influences the parameters of motor reactions, but, most importantly, motor representations influence perceptual processes (Craighero *et al*. 1999; Loula *et al*. 2005; Casile & Giese 2006). The next section will concentrate on neurophysiological findings.

Single neuron recording studies allow to correlate the neuronal activity with the presentation of sensory stimuli or with specific behavioural events. Using this technique, it has been demonstrated that neurons of premotor cortex (Brodmann's area 6) activate during goal-related motor acts, such as reaching, pushing, grasping, manipulating, breaking, etc., rather than during simple movements, such as, for example, arm extension or fingers flexion. In particular, single neurons of ventral premotor area F5 code various levels of abstraction of motor acts. Some activate when a monkey executes motor acts such as grasping, manipulating, holding, tearing objects (Rizzolatti *et al*. 1988). Within the category of grasping neurons, some discharge when the monkey grasps food with the hand or the mouth, or when the same goal is achieved not only with the biological effector but also with a tool, after a period of motor training to use it (Umiltà *et al*. 2008). All these studies strongly demonstrate that the main role of the motor cortex is that of coding goals. The neurons coding these goals form, together, a 'storage' of motor representations, a kind of 'internal motor knowledge' of the individual. Coming back to the motor imagery function (see above), whenever an individual imagines to perform a motor act, the activation of the motor cortex reflects the activation of specific motor representations. Notably, for the occurrence of this function a previous activity of posterior, sensory or associative cortices, is not required.

Usually, our behavior occurs in response to sensory stimuli. Therefore, based on the above described motor organization, in order to transform a sensory input into a motor output, the former must be associated to the appropriate motor representation. This is possible through the anatomical connections between parietal and motor cortex. Note that this link ensures two functions: (i) the transformation of an external input (e.g., an object) into a motor format (e.g., grasping) and (ii) a sensorimotor matching mechanism, providing an automatic attribution of motor meaning to the sensory input addressing a specific motor representation. Neurophysiological studies demonstrated that this double function is present in several circuits (Rizzolatti & Luppino 2001). For example, a circuit linking an area of intraparietal cortex (anterior intraparietal area, AIP) with an area of ventral premotor cortex (area F5) is involved in transforming the physical object properties in grasping motor acts. At the same time, the visual response of F5 neurons to object presentation appears to derive from a mechanism matching the object visual characteristics with the corresponding motor representations (the grip used to grasp that object). Thus, this neuronal visual response represents a pragmatic description of the object because the object is described in motor terms. Another example of a sensorimotor matching process that allows individuals to achieve a specific understanding of the external world is the mirror matching mechanism. The system based on this mechanism in monkeys and humans will be described in the next section.

In conclusion, the existence of a variety of sensorimotor matching mechanisms suggest that these systems have been selected in the phylogenesis because of their pragmatic role and their capacity to perform two functions: sensorimotor integration and achievement of a motor knowledge of the external world.

2.1. Mirror Neurons in the Monkey

Mirror neurons were first discovered in area F5 (see Fig.1A) of the monkey ventral premotor cortex (Di Pellegrino *et al*. 1992; Gallese *et al*. 1996; Ferrari *et al*. 2003) and then in area PFG in the rostral part of the inferior parietal lobule (Gallese *et al*. 2002; Fogassi *et al*. 2005; Rozzi *et al*. 2008). Since the properties of mirror neurons of F5 and PFG are similar, we will describe them together. These neurons discharge when the monkey *performs* a hand or mouth goal-directed motor act (e.g., grasping, biting, tearing or manipulating an object), and when it *observes* the same, or a similar, act performed by the experimenter or by another monkey. They do not respond to the simple object presentation (differently from the neurons of the AIP-F5 circuit) and to the vision of the hand mimicking the motor act without the target. Notably, the motor acts effective in eliciting the mirror neurons response are the same as those coded by purely motor neurons, that is they correspond to the different goal types stored in ventral premotor and inferior parietal cortex.

The finding that mirror neurons respond visually to the presentation of a hand (or mouth) — object interaction strongly suggest that these neurons code the goal of the observed motor acts. This suggestion is corroborated by the results of two investigations. In the first it has been shown that mirror neurons discharged both when the monkey could fully observe a grasping act and when it could see only part of it because the hand-target interaction was hidden behind a screen (Umiltà *et al*. 2001). The access to memory of the object, combined with the vision of the reaching component of the act, allows the retrieval of the motor representation of the partially observed motor act and of its consequences. Thus, it has been concluded that the neuronal discharge obtained in the partially occluded condition codes the goal of the act, even though its target is not visible.

In the second study, monkey had to both observe motor acts (breaking, manipulating, dropping) and listen to the sound of these noisy acts, while the response of mirror neurons was recorded (Kohler *et al*. 2002) The results showed that a sub-category of them discharged not only during motor act observation but also during pure listening to the sound produced by that act ('audio-visual mirror neurons'). This indicates that the content (the goal) of the motor act can be accessed, and therefore coded, through different sensory modalities. Although the response of most mirror neurons is independent of many details of the observed act, such as space, type of object or hand, recently two studies have demonstrated that part of mirror neurons can provide also information on the visual details.

In the first study (Caggiano *et al*. 2009) the visual response of mirror neurons have been analyzed during observation of an experimenter grasping a piece of food within the monkey reaching space (peripersonal space), or far from the monkey (extrapersonal space). The study showed that half of the tested mirror neurons responded better in one of the two conditions, the two subcategories (peri- and extrapersonal neurons) being equally represented. This finding suggests that mirror neurons can code others' actions within different spaces, possibly in relation to the possibility to socially interact with others by performing different types of behavioural responses.

In a second study (Caggiano *et al*. 2011), the responses of F5 mirror neurons were investigated during observation of movies showing grasping motor acts seen from different perspectives (frontal, lateral, egocentric). The first interesting result of this study was that mirror neurons respond also to acts presented in movies, although the same motor acts, presented naturalistically, evoked a higher visual response. The second result was that while one quarter of the recorded mirror neurons responded to the visual presentation of motor acts in movies, independent of the visual perspective from which they were presented, the other three-quarters of neurons were tuned to specific visual perspectives.

Altogether, the two studies indicate that there are mirror neurons whose function is only that of encoding the goal of a motor act, while others, beyond this property, can also contribute to provide the observer with the details of the observed act, probably through feedback connections between motor cortex and posterior, higher order, visual areas. This mechanism, which is supported by the presence of reciprocal connections between anterior and posterior cortical areas, would explain a way in which motor representations and the corresponding sensory representations mutually interact.

The encoding property of a neuron can be read only from its output, i.e. from its discharge. A neuron has only a single output, meaning that it produces only one code. Since, however, its output depends on the integration of several inputs, it is important to know which of these inputs can drive a certain output and which cannot. In the case of mirror neurons, this input-output comparison becomes an analysis of the congruence between the observed and the executed motor act. This analysis leads to conclude that ninety percent of mirror neurons are congruent in terms of goal (Gallese *et al*. 1996; Rozzi *et al*. 2008). However, this congruence may be strict or broad. In 'strictly congruent' mirror neurons the observed and executed motor acts correspond both in the goal and the details of the act. In 'broadly congruent' mirror neurons there is also congruence of the goal but, for example, the range of observed motor acts effective in eliciting the visual response can be broader with respect to that of motor acts effective during execution. The congruence property of mirror neurons is very important for several reasons. The first is that the congruence manifests the occurrence of an efficient matching mechanism, which is the basis for understanding others' actions. In other words, the observation of a specific motor act performed by another individual elicits the activation of the corresponding motor representation in the motor system of the observer, as if he was executing the same act. Note that, during observation, an inhibitory mechanism must come into play, blocking the automatic execution of the observed act (see Kraskov *et al*. 2009). Second, a strict congruence can be very important if these neurons must be used in imitative processes (see below). Third, a broad congruence can be very helpful for the generalization of the meaning of the observed motor act. This aspect was confirmed by two studies showing that some mirror neurons can begin to respond also during observation of motor acts performed with tools, provided a long visual exposure to these acts (Ferrari *et al*. 2005) or after training to use these tools (Umiltà *et al*. 2008). Fourth, observation/execution matching can be very useful during social learning processes such as those occurring during language development.

Which is the source of visual information that contributes to the formation of the mirror matching mechanism? It is known that in the anterior part of the monkey superior temporal sulcus (STSa) there are neurons that fire during observation of biological movements, among which also hand motor acts (Perrett *et al*. 1989). These high order visual neurons are considered the source of visual input for parietal mirror neurons and, as a consequence, for ventral premotor mirror neurons. Note that both STSa and PFG, on one side, and PFG and F5, on the other, are reciprocally connected (Rozzi *et al*. 2006; Bonini *et al*. 2010), while there is no direct connection between STSa and F5. Thus, it is not known where is the first place where matching may occur.

2.2. *The Mirror System in Humans*

After the discovery of mirror neurons in the monkey, the presence of a mirror system (MS) in humans has been demonstrated with electrophysiological and neuroimaging techniques. For example, the transcranial magnetic stimulation (TMS) technique, that allows either to excite or to inactivate limited cortical regions, enabled several researchers to demonstrate the existence of a mirror matching system at the level of the motor cortex. In fact, if a TMS pulse at threshold intensity is given to the motor cortex of subjects observing motor acts performed by another individual, it is possible to enhance the electromyographic activation of the same muscles that would be active if the subjects themselves would perform the observed motor act (Fadiga *et al*. 1995, Gangitano *et al*. 2004). This enhancement is justified only if the stimulated region is pre-activated by observation.

While TMS provided indication of the occurrence of a matching mechanism, positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) techniques allowed to identify the anatomical location of areas active during observation of motor acts. Since up to now several brain imaging studies have been carried out on this topic, we will limit to sum up all these results. The main areas activated during observation are: A region around the superior temporal sulcus (STS), one in the supramarginal gyrus (part of the inferior parietal lobule, IPL), and a third in the ventral premotor cortex plus the posterior sector of the inferior frontal gyrus (IFG), namely areas area 44 and 45 that, in the left hemisphere, correspond to Broca's area (the 'speech' area) (see for review Rizzolatti & Sinigaglia 2009; Caspers *et al*. 2010). In many of these studies it has been underlined that the areas activated in IFG and IPL correspond anatomically to the areas where mirror neurons have been found in monkeys (F5 and PFG, respectively). In addition, the STS region corresponds to the monkey area described by Perrett *et al*. (1989), which is active only during observation of a motor act, but not during its execution. Among activated areas can also be included the anterior intraparietal area (AIP) — corresponding to the monkey area having approximately the same location — that is mainly activated by observation of hand motor acts (Shmuelof & Zohari 2008), and, in some studies, regions of the dorsal premotor cortex and superior parietal lobule, that can become active during observation of reaching motor acts (Filimon *et al*. 2007).

Finally, electroencephalographic (EEG) and magnetoencephalographic (MEG) studies were very useful not only to confirm the activation of the frontal cortex during action observation (Cochin *et al*. 1999; Nishitani & Hari 2000), but also to show the time course of activation. In fact, it has been shown that, after activation of the occipital lobe, the activation of IFG precedes that of precentral cortex.

In monkeys, MNs respond exclusively or stronger to goal-directed acts than to intransitive movements, the only exception being communicative MNs (see below). Differently from monkey studies, in humans also the observation of meaningless movements can elicit an activation of areas belonging to the motor system. For example, TMS studies showed that observation of meaningless movements determine a resonance in the motor cortex (Fadiga *et al*. 1995). Some neuroimaging studies showed that observation of meaningless movements activate a dorsal premotor-parietal circuit (Grèzes *et al*. 1998), others show that observation of mimed motor acts activate the same premotor, IFG and parietal regions activated by the goal-directed ones (Grèzes *et al*. 1998; Buccino *et al*. 2001). Interestingly, observation of symbolic gestures appear to activate both ventral premotor and inferior parietal cortex, but this latter activation involves more posterior sectors than those activated by observation of goal-directed motor acts (Lui *et al*. 2008). Thus, observation of non goal-directed movements may activate areas belonging to the MS, and these areas may be different depending on whether the observed movements are meaningful or meaningless.

Interestingly, it has been claimed that intransitive gestures (pantomimes) play a very important role in language evolution, because they facilitated the transition from object-related actions to intentional movements to “protosign, a manual-based communication system that broke through the fixed repertoire of primate vocalizations to yield an open repertoire of communicative gestures” (Arbib 2005: 107).

2.3. *The Involvement of the Mirror System in Imitation*

A very interesting issue, strictly related to the properties of mirror neurons and very important for human learning and evolution of language, is that of imitation. In adult monkeys there are no clear reports of this behavior, in particular of imitation learning (Visalberghi & Fragaszy 2002). Thus, in these species, the mirror neuron system, besides providing individuals with the capacity of understanding actions, can be perhaps exploited for imitative processes such as neonatal imitation and action facilitation (Ferrari *et al.* 2009, 2012). In humans, however, the imitative function, and specifically imitation learning, is very well developed and is critical particularly during language acquisition in children. Many brain imaging studies carried out in the last decade, showed that the parieto-premotor mirror system is activated when subjects are required to observe and imitate simple finger movements, goal-related motor acts and complex action sequences (Buccino *et al.* 2004; see for review Iacoboni 2009). Thus, observation and imitation share the same neural circuits. In fact, during imitation learning, a crucial step is represented by the recognition of the motor acts belonging to the action to be imitated. This step is very likely accomplished by mirror neurons. A second step concerns the internal reconstruction of the sequence of motor acts to be reproduced. Is this a property of mirror neurons or is performed by other cortical regions? An fMRI experiment in which subjects were required to observe novel guitar chords performed by an expert player and then to imitate them (Buccino *et al.* 2004), show, beyond that of the parieto-frontal mirror neuron system, a strong activation of the middle frontal cortex (area 46) during the phase interleaved between observation and imitation. In this phase the subject has to decompose the action into its basic motor elements and then subsequently recombining them into a new action matching the observed one. This is probably the role of prefrontal cortex, known to be crucial in action planning, attention and working memory.

3. Matching Sounds with Actions

The acoustic signal generated by a biological event is the result of a motor act. This motor act corresponds to a specific motor representation stored in the cortical motor system. When this signal, either a phoneme, a word or a sentence, is listened by another individual, a series of neuronal processes are taking place in the brain, starting from the auditory pathway but not ending into the primary auditory cortex and higher order acoustic areas. In fact, many neurophysiological investigations clearly show that part of this information reaches premotor cortex and connects with a specific motor representation (for a review, see Pulvermüller & Fadiga 2010). The activation of this motor representation, in turn, is crucial for the production of the same acoustic signal.

The described process can be considered similar to that occurring in a linguistic interaction between a sender and a receiver. Although the linguistic signal has a specific pattern and characteristics that distinguish it from an action sound or a physical sound, nevertheless also in this domain the matching mechanism can operate similarly to that occurring during listening to an action sound. As a matching mechanism allows the understanding of an action meaning, the same mechanism allows the understanding of a phoneme. Thus, a syllable is not understood simply because of the perception of its acoustic features, but because these features retrieve an invariant motor representation. Interestingly, during the first year of life, infants demonstrate an increased sensitivity to phonological properties and during the babbling phase it is likely that the articulation of sounds and the first produced phonemes play an important part not only in sensorimotor association but also in strengthening and mapping neuroanatomical connections between acoustic areas and cortical motor representations involved in sounds production (Pulvermüller & Fadiga 2010).

This proposal of a 'pragmatic' and embodied understanding of spoken words resembles very much the motor theory of speech perception (Liberman & Mattingly 1985), postulating that "the objects of speech perception are the intended phonetic gestures of the speaker, represented in the brain as invariant motor commands" (p. 2). From these considerations two important implications follow: (i) Speech (and maybe language) understanding can be rooted on a motor ground and (ii) the mechanism matching the auditory and motor representation of an utterance can subserve both phonological matching and the retrieval of the associated meaning. This would constitute a motor-based mechanism for automatic comprehension of semantics.

A series of behavioral and neuroscientific studies support these two claims. One of the first demonstrations of a phonological motor-based matching system is that of Fadiga *et al*. (2002), who recorded motor evoked potentials (MEPs) from the tongue muscles by stimulating with TMS the left motor cortex of normal volunteers instructed to listen to acoustically presented words and pseudo-words, containing either a double 'f' or a double 'r', that require a different tongue muscles involvement to be pronounced (stronger during words containing double 'r'). The TMS pulse was given at the time in which the double consonant was produced by the speaker. The results showed that listening to words and pseudo-words containing the double 'r' determined a significant increase of the amplitude of MEPs recorded from the tongue muscles with respect to listening to words and pseudo-words containing the double 'f' and bitonal sounds. Furthermore, the tongue muscle activation during word listening was higher than that during listening to pseudo-words. These data strongly suggest that phonology and perhaps, partly, semantics, are processed within the motor system. In a related TMS experiments, Watkins *et al.* (2003) stimulated the face motor field of subjects listening to speech or viewing speech-related lip movements, as compared to listening to non-verbal sounds and viewing eyes and brow movements. The results show that, compared to control conditions, listening to and viewing speech enhanced the size of MEPs.

In a more recent experiment (D'Ausilio *et al*. 2009) TMS pulses were given to the lip and tongue cortical motor fields while subjects listened to phonemes produced with the lips (b and p) and the tongue (d and t) and performed a phoneme discrimination task. The TMS pulses were applied just before stimulus presentation to selectively prime the cortical activity specifically in the lip or tongue representations, with the hypothesis to obtain a bias in perception. Results showed that indeed the stimulation of a given motor representation led to a better performance in recognizing speech sounds produced with the concordant effector compared with discordant sounds produced with a different effector.

These and other studies clearly support the view that the motor system contributes to the identification and discrimination of speech sounds. However, the investigations reviewed so far have described the involvement of the motor system at the phonological level. Is there any evidence that during listening to a word the premotor and primary motor cortex are involved, thus contributing to the comprehension of the meaning?

There is some evidence suggesting that the meaning of a word can be indeed processed in these areas. Functional MRI studies demonstrated that pure listening to action verbs or action-related sentences produces an effector specific, somatotopic, activation of the motor cortex (Hauk *et al*. 2004; Tettamanti *et al*. 2005), similar to that found by Buccino *et al*. (2001) during observation of motor acts performed with the mouth, the hand or the foot. In all these cases, as in the above reported TMS studies, listening to action-related verbal material produced an enhancement of motor cortical activation. On the contrary, Buccino *et al*. (2005), using TMS, found a suppression of motor activity while stimulating either the hand or the foot/leg motor area of the left hemisphere, while participants were listening to sentences expressing hand and foot actions. They found a decrease of MEPs amplitude when the stimulated field corresponded to the effector involved in the listened action. A complementary reaction time study, in which subjects had to respond with the hand or the foot to the same sentences, confirmed this suppression effect. If, on one side, these data demonstrate an effector specific modulation of the motor system during listening to action-related material, on the other side, this modulation is consistent with an inhibition effect. Following one of the authors' interpretations, it is plausible that the motor representation elicited by listening to the sentence interfered with the motor program activated in order to respond with the required same effector.

Interestingly, although it has been suggested that the activation of the inferior frontal areas during action-verb listening is not related to a process of meaning comprehension (Hickok 2009), other studies contradict this view. In fact patients with motor neuron disease or lesions in the left inferior frontal cortex have deficits in action-verb understanding and in semantic understanding of pictures depicting actions (Bak *et al*. 2001, 2006).

It must be noted however, that several words are not related to verbs and describe, for example, categories of objects or abstract concepts that are unrelated to actions. In these cases, brain imaging studies have shown activations in areas of the temporal lobe, thus suggesting that the motor links in the semantic processing is limited to words related to actions (see Pulvermüller & Fadiga 2010).

Altogether, these and other findings support the idea of a strong embodiment of speech during acoustic processing, both at phonological and semantic level. However, other theoretical approaches contrast with this view (see, for example, Hickok 2009; Lotto *et al*. 2009). Their arguments are against a main role of the motor system in general and of the mirror system in particular, in speech perception. They report, for example, that infants could discriminate speech sounds that they could not yet produce (Eimas *et al*. 1971). Furthermore, Broca's aphasia does not seem to prevent normal receptive speech ability (Damasio 1992; Goodglass 1993). Furthermore, there are data of double dissociation in Broca's aphasics showing that there are patients impaired in syllables discrimination but as good as normal in word comprehension and other patients that, although impaired in speech discrimination, can be good in repeating heard speech (see Hickok & Poeppel 2007 and Lotto *et al*. 2009 for review). Note, however, that other studies aimed at investigating how deep were the impairments in speech comprehension in these patients, found that some deficits were evident especially in single-word comprehension (Utman *et al*. 2001; Yee *et al*. 2008; see Pulvermüller & Fadiga 2010).

About the role of the motor system in the comprehension of semantics, the critics of the 'motor' approach, although recognizing that its involvement has been shown by several works, maintain that, instead of being central to language understanding, it could play a post-recognition epiphenomenal role (see Hickok 2009).

Altogether, these data suggest that the proposal concerning the involvement of the motor system in speech perception warrants more attention concerning how central is its role and clearly this requires further investigations to better understand the underlying neural mechanisms.

4. Vocal and Gestural Communication

It is broadly accepted that several aspects of human language rely on basic elements that are shared with other animals. However, it is still a matter of debate which feature belongs exclusively to language (Hauser *et al*. 2002). Vocal communication in nonhuman primates has been often compared to human language in an effort to understand possible commonalities and the basic components from which human language might have emerged (Ghazanfar & Hauser 1999; Seyfarth & Cheney 2010). Undoubtedly, monkeys can emit utterances in different contexts to signal urgent events such as the presence of predators or food, or to threat a conspecific during a dispute. The seminal study by Cheney & Seyfarth (1982) showed that vervet monkeys are capable of referential communication, providing information through alarm calls to conspecifics about the different types of predators approaching. Recipients seem to understand the different meaning of the call or, as demonstrated in other studies, are able to take into account the possible causes of the alarm calls (Zuberbühler 2000).

The signs of flexibility in the vocalization systems of many nonhuman primate species is supported by their capacity to modify their utterances based on their assessment of the possible consequence they produce on other individuals. This phenomenon, named the audience effect, demonstrates the capacity of monkeys and apes to modify their vocalizations depending on the social context (Caine *et al*. 1995; Tomasello & Zuberbühler 2002). More recently, it has been demonstrated that wild chimpanzees emit alarm calls with low or high frequency based on the state of knowledge of the group member about a danger (Crockford *et al*. 2012). These data suggest that in some nonhuman primates it has been developed a capacity to control vocalization to inform group members depending on the information available to them about environmental dangers.

In terms of learning, nonhuman primates vocalization shows a certain degree of flexibility. For example some species display population-specific vocal signals and in the course of ontogeny there are modifications in the structure of their vocalization and in their capacity to use it in the appropriate context (Tomasello & Zuberbühler 2002; Egnor & Hauser 2006). However, cross-fostering studies demonstrate that the capacity to learn new vocalizations is very limited (Owren *et al*. 1993). Furthermore, this limitation has also been supported by several experimental studies showing that monkeys and apes cannot learn completely new vocal patterns (see Yamaguchi & Izumi 2008). This is illustrated, for example, by the failure in teaching human spoken language to apes, even though some of the basic cognitive components that are found in language (e.g., use of signs or symbols for communication) are present, as it has been clearly demonstrated by the old experiments involving symbolic communication teaching by means of keyboards (Gardner & Gardner 1969; Patterson 1978; Savage-Rumbaugh *et al*. 1986).

One of the main reasons of the limited flexibility in the vocal patterns may stem from a limited capacity in voluntarily controlling the vocal apparatus. Although behavioral studies have demonstrated that macaque monkeys can achieve a significant level of voluntary vocal control when submitted to operant conditioning tasks (Sutton *et al*. 1973; Aitken &Wilson 1979; Hihara *et al*. 2003; Yamaguchi & Izumi 2008), the success rate of vocal training studies remains variable (Yamaguchi & Myers 1972; Pierce 1985). In a recent experiment we demonstrated that when macaques are subjected to an intensive training aimed at emitting coo calls, they are capable to reach a significant level of success (Coudè *et al*. 2011) However, it also emerged that often the monkeys were able to modify correctly the mouth configuration but showed difficulties in sound emission, thus suggesting that in macaque monkeys there is only a partial voluntary control and coordination of the mouth together with the larynx muscles.

4.1. Gestures in Monkeys and Apes

In one of our previous review we proposed the terms gesture to describe both goal-related actions (e.g., grasping an object with the hand) and communicative oro-facial and brachio-manual movements devoid of an explicit target (Fogassi & Ferrari 2007). Some scholars make different distinctions about what is gesture, emphasizing the motor ineffectiveness, the persistence and the role of gaze etc. (Call & Tomasello 2007). We are not going to cover all the theoretical accounts that define a gesture but instead we would like to make an attempt in understanding the possible link at the neurological level between the control of a movement directed to a target (i.e. a motor act) and the intentional movement that is directed toward another individual which has a communicative meaning. There are several converging evidence from neuroscience, ethology, and developmental psychology that many of the gestures displayed by nonhuman primates began their existence as actions devoid of a communicative function, but over time they became co-opted and transformed into communicative devices that accomplished similar functions (Fogassi & Ferrari 2007; Liebal & Call 2012). Probably, only at a later stage in primate and hominin evolution the vocalization system has been integrated in such gestural system, and in our species we can testify the presence of such ancestral linkage (Corballis 2003).

It's outside the scope of our review to provide a full account of gestural communication in nonhuman primates as it has been already extensively covered elsewhere (Liebal *et al*. 2004; Liebal & Call 2012). However, what is relevant for the current work is to describe some aspects of this type of communication in relation to neurological mechanisms and language evolution. Gestures can involve the oro-facial and/or the brachio-manual system in conjunction with body postures. Although some of the facial gestures are elicited by less urgent contingent situation, they often involve face-to-face exchanges, involuntary acts and autonomic responses. Some of these gestures have been extensively studied by comparative investigations that could reconstruct, with reliable approximation, their possible relatedness and origin among the different species (van Hooff 1962, 1967). Some of these, we believe, might be particularly relevant for the current topic. For example, lipsmacking has been described in several Old World primate species. It is characterized by regular cycles of vertical jaw movement, often involving a parting of the lips, but sometimes occurring with closed, puckered lips and sometimes alternated with tongue protrusions. Importantly, as a communication signal, the lipsmack is always directed at another individual to signal affiliative and benevolent intentions and it is displayed during face-to-face interactions in which both individuals might lipsmack at each other (Maestripieri 1996; Ferrari *et al*. 2009; Morrill *et al*. 2012). Interestingly, this behavior is one of the first to emerge in the course of ontogeny and undergoes into changes both in the pattern and in the frequency with which it is emitted (Ferrari *et al*. 2009). We have recently found that in rhesus macaques, infants produce lipsmacking at a slower frequency than adults and that these cyclic movements become faster and less variable with age (Morrill *et al*. 2012). We also found that other cyclic movements like chewing do not follow similar developmental patterns. We proposed that the development of lipsmacking follows a trajectory that resembles that of babbling in humans.

From an evolutionary perspective it has been proposed that some communicative gestures such as lips-smacking and pucker face very likely evolved from movements aimed to remove and eat particles, such as skin parasites, from the fur of group mates during grooming sessions. This suggestion is corroborated by the observation that the beginning of a grooming session can be preceded or accompanied by a lips-smacking action without ingestion (see Van Hooff 1962, 1967; Maestripieri 1996). Through a process of ritualization, an ingestive action could have lost its behavioral meaning related to feeding and achieved an affiliative meaning. Together, these evolutionary and ontogenetic accounts seem to be in agreement with the proposal by MacNeilage (1998) that the rhythmic cyclic mandibular open-close alternation produced during infant babbling might have evolved from rhythmic mouth movement of our ancestral primates.

Concerning brachio-manual gestural communication, apes use them in a richer and more elaborated way than monkeys (Call & Tomasello 2007). In the last few years there has been an increasing body of research, in part stimulated by the idea that brachio-manual gestures have probably played a role in language evolution (Liebal & Call 2012). Apes, for example, are capable to use several types of gestures, often in combination, to request for food (Leavens *et al*. 2004, 2005; Gomez 2007). In captivity, chimpanzees and also some monkeys point to request food or objects and, in the case of chimpanzees, they are sensitive to the attentional state of the human experimenter when they point. Although they do not gesture to share information or to inform others, it has been pointed out that they might use brachio-manual gestures in many flexible ways. Under human rearing conditions some apes have been reported to use declarative gestures, thus showing the potential to expand their cognitive and contextual use of the communicative gestures (Lyn *et al*. 2011).

Another important aspect of gestural communication is the possibility to perform sequences and to combine them with other communicative signals such as face gestures and vocalization. The issue of sequencing has been investigated in a few studies on apes (Liebal *et al*. 2004; Pollick & de Waal 2007). Chimpanzees can use hands gesture sequences producing signals in one or more modalities (visual, auditory, or tactile) (Liebal *et al*. 2004). Although the sequences could involve the repetition of the same gesture, it has been noted that several were composed by three or more different gestures. Most of them were used in a play context but other sequences were often used in more diverse situations. Sometimes the use of the sequence was the result of the failure to gather the recipient attention (Liebal *et al*. 2004). Other studies demonstrate that chimpanzees can often display attention-gather hand gestures in combination with sound production (Leavens *et al*. 2004). More recently, the use of these combinations has been extensively studied in chimpanzees and bonobos (Pollick & de Waal 2007). The results showed that brachio-manual gestures were more flexible across contexts than the facial/vocal communicative ones. The former appear to be less tied to incipient events that may induce high emotional responses than facial gestures and vocalization. These authors conclude that very likely the flexibility of brachio-manual gestures might have played a central role as a prerequisite model for language evolution. This would be also supported by other data showing that apes can imitate brachio-manual gestures and by the report of population-specificity of some of these gestures. However, one of the limitations in these studies is that they have been carried out on animals living in captivity where both the exposure to humans and the weak ecological value in displaying the full-range of communicative behaviors could limit our knowledge on the natural combinatorial use of different gestures and vocalization in the perspective of language evolution.

5. Anatomo-Functional Homologies between Monkey Premotor Cortex and Human Broca's Area

The relation between the mirror neuron system and language is corroborated by comparative data. We will briefly summarize the main evidence of such homology. First, cytoarchitectonic data suggest that dysgranular area 44 (the posterior part of Broca's area) and monkey area F5 share similar features (Petrides & Pandya 1994; Rizzolatti & Arbib 1998) (see Fig. 1 below). Although some authors have emphasized the anatomical characteristics of monkey area 44 (located, according to Petrides *et al*. (2005), in the fundus of the inferior limb of arcuate sulcus) as a possible precursor of human 44, neurophysiological recordings provided evidence that the whole rostral part of ventral premotor cortex in monkeys has typical features (see below) that are shared with those recently described for Broca's area.

Second, area F5 contains motor neurons related to the execution of both hand and/or mouth actions. Similarly, brain imaging experiments in humans demonstrated that Broca's area, traditionally considered as a 'speech' area, is also involved in hand movement tasks such as complex finger movements, mental imagery of grasping actions, and hand imitation tasks (see Rizzolatti *et al.* 2009).

Third, a recent neurophysiological study showed that in the lateral part of area F5 there are neurons specifically activated during conditioned vocalization (Coudé *et al.* 2011, see below), suggesting its possible role in the evolution of voluntary controlled phonation. These findings are relatively surprising because it is known that ventral premotor cortex contains a representation of the larynx (Hast *et al.* 1974; Simonyan & Jurgens 2003).

Fourth, both areas are considered to be the rostral pole of the mirror neuron system. In fact, the evidence reviewed above demonstrates that Broca's area, as F5 mirror neurons, is activated when subjects observe goal-related hand and mouth motor acts done by another individual (see Rizzolatti *et al.* 2009). Furthermore, in accord with the presence of F5 mirror neurons responding to the sound of motor acts (Kohler *et al.* 2002), also the human left motor cortex is activated when subjects listen to sounds associated with motor acts (Aziz-Zadeh *et al.* 2004; Gazzola *et al.* 2006; Ricciardi *et al.* 2009).

Fifth, in agreement with the presence of oro-facial communicative mirror neurons in F5, Broca's area activates when subjects observe silent speech (Buccino *et al.* 2004b).

Summing up, in monkey premotor cortex there are several features that can pre-adapt this cortical sector for the evolution of a sophisticated communicative system. The core of these features consists in encoding the production and perception of both oro-facial and forelimb gestures in the same cortical area. This double control, once integrated with that of vocalization, would have constituted the basis for a communicative system with an increased complexity and efficiency, and a higher level of flexibility in transferring information to conspecifics.

5.1. *Toward the Integration of Vocalization with Gestures*

Nonhuman primate vocal behavior was traditionally assumed to be predominantly emotional (Seyfart & Cheney 1997; Fitch 2000; Premack 2004) and mainly consisting of involuntary or reflexive responses. However, several investigations partly contrast this view showing that monkeys can achieve a significant level of voluntary vocal control when submitted to operant conditioning tasks (Sutton *et al.* 1973; Aitken & Wilson 1979; Hihara *et al.* 2003).

From a neurophysiological perspective vocal production in nonhuman primates is considered to be controlled by the brainstem and by mesial cortical areas which, besides other functions, are also involved in emotional behavior (West & Larson 1995; Izumi *et al.* 2001; Jurgens 2002). We have recently challenged this view with a study (Coudé *et al.* 2011) in which we recorded from ventral premotor cortex of macaques trained to control simple calls (i.e. coo-calls). The results showed that the rostro-lateral part of PMv (mostly area F5) contains neurons that fire during conditioned but not during spontaneous vocalization. Interestingly, in the majority of these neurons the discharge begun before sound onset, thus suggesting their causal relation with vocal production. Furthermore, these neurons have been found intermingled with other neurons controlling mouth and hand motor acts, and with mirror neurons. The link of these neurons with the motor control of vocalization has been also supported by the electrical microstimulation of this sector, which in some cases elicited larynx contraction.

Interestingly, although it is known that acoustic input related to motor acts reaches area F5, we did not find neurons that, beyond their responses during vocalization, responded also during listening to the same or a similar call. This could be due to the fact that while in monkeys during face-to-face gestural communication there is a frequent exposure to oro-facial gestures of their conspecifics with reciprocal exchanges, the same was not true, in this study, for vocal calls. In other words, these monkeys had a feedback from their own voluntary controlled call production, but they were never exposed, except during specific testing, to voluntary calls produced by others and to their associated facial expressions. Another possible explanation for the absence of mirror neurons for vocal calls is that they exist only for emotional vocalizations, in other brain areas. On the other hand, other studies clearly showed that in the superior temporal and prefrontal cortices of the macaque there are neurons coding listened species-specific vocalizations (Rauschecker *et al*. 1995; Ghazanfar *et al*. 2005; Romanski *et al*. 2005), apparently in absence of vocal production-related discharge. Although these perceptual responses could be simply the result of a high order sensory elaboration, it is also possible that in monkeys the acoustic input reaching frontal areas is not coupled yet with the motor representation of vocalizations, at difference with what occurs in other species, such as humans and songbirds (Pulvermüller *et al*. 2006; Prather *et al*. 2008).

These findings suggest that in monkey may exist, in the lateral part of the cortex, a primitive system for the voluntary control of phonation, anatomically embedded in that controlling hand and mouth goal-directed motor acts. Thus, it is plausible to propose that in a further evolutionary step, these two systems could have been integrated. This is, indeed, what has been found, at a behavioral level, in chimpanzees, in which communicative brachio-manual gestures are often accompanied by vocalization (see previous section). At the neurological level, the investigations on chimpanzees or other apes is by far more complex. However, in the last decades structural and functional neuroimaging studies in apes have provided relevant information concerning the neural control of vocalization and communicative gestures in relation to the possible homology with Broca's area. First, it has been shown that in apes there is an asymmetry of the inferior frontal gyrus (Cantalupo *et al*. 2001, 2009; Keller *et al*. 2009) even though it remains to be clarified whether this effect is due to the contribution of gray and/or white matter (see Schenker *et al*. 2010). Second, it has been shown in chimpanzees that right-lateralized skillful hand behaviors correlate with left asymmetry of the hand field of precentral gyrus (Hopkins *et al*. 2010). Third, a PET study in chimpanzees demonstrated the activation of the homolog of Broca's area during the production of communicative vocal and hand gestures (Taglialatela *et a*l. 2008). This latter finding is important, because it suggests that, with respect to monkeys, the lateral frontal cortex (area 44 and 45) of chimpanzees (Fig. 1) can control brachio-manual communicative gestures, possibly in conjunction with vocal production, and that this function is lateralized. Fourth, although indirectly related to communication, a recent PET study reported that a parieto-frontal system is activated in chimpanzees during execution (performed without seeing its own hand) and observation of grasping actions (Hecht *et al*. 2011), giving support to the evolutionary continuity between monkey and human mirror system. Whether the chimpanzee's mirror system is involved in communication still needs to be investigated.

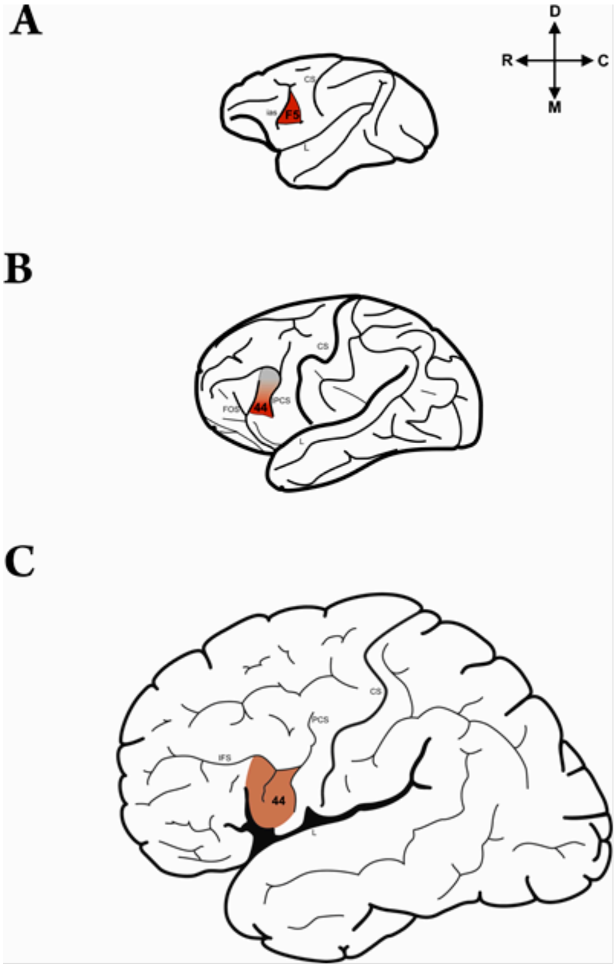


Figure 1: Location (in red color) of area F5 in the macaque monkey (A) and the proposed homologue areas in chimpanzee (B) and man (C). In the chimpanzee brain area 44 has been reconstructed based on the anatomical description of the sulci and contours made by several authors (Bailey et al. 1950; Keller et al. 2009; Schenker et al. 2010; Keller et al. 2011). Area 44 in the human brain has been defined according to Brodmann (1909). CS: central sulcus; ias: inferior arcuate sulsus; FOS: fronto-orbicular sulcus; IFS: inferior frontal sulcus; IPCS: inferior precentral sulcus; L: lateral sulcus; PCS: precentral sulcus.

5.2. *Hypothetical Transitions from Gestures to Language*

The achieved integration, at the apes phylogenetic level, between communicative brachio-manual gestures and vocalization, very likely has been preserved in the next evolutionary steps, even in the presence of a clear predominance of the phonatory system. Although it is speculative to mark when the phonatory system reached an autonomous function, in modern humans, the tight observed connection between spoken language and gestures (McNeill 1992; Goldin-Meadow 1999) seems to testify an old common origin. This link is not just an evolutionary relic without a meaningful function, but enables individuals to possess an enriched multimodal communicative system. Although in some situations only vocal communication can be used, it is also true that under specific acoustic impairment conditions (i.e. in deaf people), gestures can assume the main role in communication (Capirci & Volterra 2008). A support to this role is given also by brain imaging studies showing an activation of the inferior frontal cortex in deaf people during production of meaningful signs (Petitto *et al*. 2000).

Interestingly, a wealth of work by Gentilucci and colleagues provided strong evidence for a reciprocal influence between hand actions/gestures and verbal production (Gentilucci & Corballis 2006). For example, they demonstrated that execution or observation of two different types of grasping (precision grip or whole hand prehension) while subjects were pronouncing syllables affected both lips kinematics and voice formants (Gentilucci *et al*. 2004). In another study they showed that in the contemporaneous production or observation of a hand symbolic gesture and a word, the kinematic parameters of the former were inhibited by the latter, while voice spectra were enhanced by the gesture (Bernardis & Gentilucci 2006). From the findings of this latter study they conclude that spoken word and symbolic gesture can be coded as a single signal by a unique communication system and that this signal may represent the intention to engage a closer interaction with a hypothetical interlocutor.

Summing up, there is much evidence that language and gestures share a common motor code, thus supporting several theories proposing that at cortical level some of the properties and organization of the motor system have been exploited within the vocal domain (Rizzolatti & Arbib 1998; Arbib 2005; Fogassi & Ferrari 2007). In line with this, several brain imaging studies show a great degree of overlap of these two systems, in particular in the frontal lobe. The spatial resolution limitation of these techniques does not allow, at present, to assess whether these two functions activate different anatomical subsectors within the frontal lobe. Even neurologically it is well known that higher order praxic and linguistic deficits can appear associated, but this is not consistent among patients, thus suggesting a partial independence of the two functions (De Renzi 1989).

Once a primitive communicative system based on an association between gestures and vocalization took place, a further step in both the motor and sensory development of this system probably occurred through the acquisition of a more sophisticated phonatory mechanism, which allowed the association of a gesture with a specific sound. At this stage of language evolution the possibility of creating a theoretically infinite set of combinations rendered the phonatory system alone more efficient than the previous vocal-gestural system. This stage was crucial for the development of a speech-based communicative system.

5.3. *Action Sequences and Syntax*

The idea of a possible similarity between the organization of actions and syntactic structure is worth to investigate, in particular under the evolutionary perspective discussed in the previous sections. According to some linguists, syntax function can be defined as a regulator of language (Pinker & Jackendoff 2005). One of the mechanisms belonging to this function is that of combining elements, i.e. words, hierarchically, into meaningful phrases. Similarly to syntactic structure, as shown by classical behavioral and psychophysical studies (Jeannerod 1988; Bernstein 1996; Rosenbaum *et al*. 2007), the 'action' is considered as the top level of the motor organization, and is formed by a sequence of motor acts. Motor acts are the basic constituents of the motor system endowed with a meaning (the motor goal), playing a role similar to that of words within a phrase. The hierarchical sequencing of motor acts into a specific action (for example, (a) grasping a piece of food, (b) bringing it to the mouth, and (c) biting it) aims to a superordinate behavioral goal (eating the food). If the order of the motor act is changed (e.g., biting the food with the mouth, bringing the hand to the mouth, and grasping the food with the hand: c–b–a) the action goal can change (take the food out of the mouth). Similarly, the meaning of a phrase is given by the sequential organization of words. By changing the position of the words in a sentence, its meaning changes or is lost.

With respect to the hierarchical organization of motor sequences, two main series of studies attempted to address this issue at the single neuron level. The first series assessed the responses of neurons in mesial cortices (pre-supplementary motor area, pre-SMA/F6 and supplementary motor area, SMA proper/F3) and prefrontal cortex while monkeys executed sequences of movements (Tanji 2001; Tanji & Hoshi 2008), such as turning, pulling and pushing, or specific sequences of reaching movements, or a series of movements in a maze. These studies showed that the recorded neurons could code either the sequence, the order of a movement inside a sequence or the final location of a trajectory. Another series of studies, carried out in our laboratory (Fogassi *et al*. 2005; Bonini *et al*. 2011), assessed the responses of parietal and premotor neurons during execution and observation of natural action sequences. The results showed that grasping neurons of areas PFG or F5 can discharge differently depending on the specific action sequence in which the grasping is embedded (see Fig. 2 below). Notably, this differential response is shown also by mirror neurons during observation of grasping embedded in different action sequences performed by another individual. This latter series of data suggests that motor neurons of parietal and ventral premotor cortex are organized in motor chains, each coding a specific action goal. Within this organization, a neuron coding a given motor act can discharge differently according to the action sequence to which this act belongs.

Summing up, both the order of a motor series and the organization of a natural action sequence can be coded by cortical single neurons. The premotor-parietal motor system plus the prefrontal cortex can provide a substrate for sequential organization and hierarchical combination of motor elements. We posit that such an organization has been exploited in other domains including some aspects of the syntactic structure of language (see also Fogassi & Ferrari 2007).

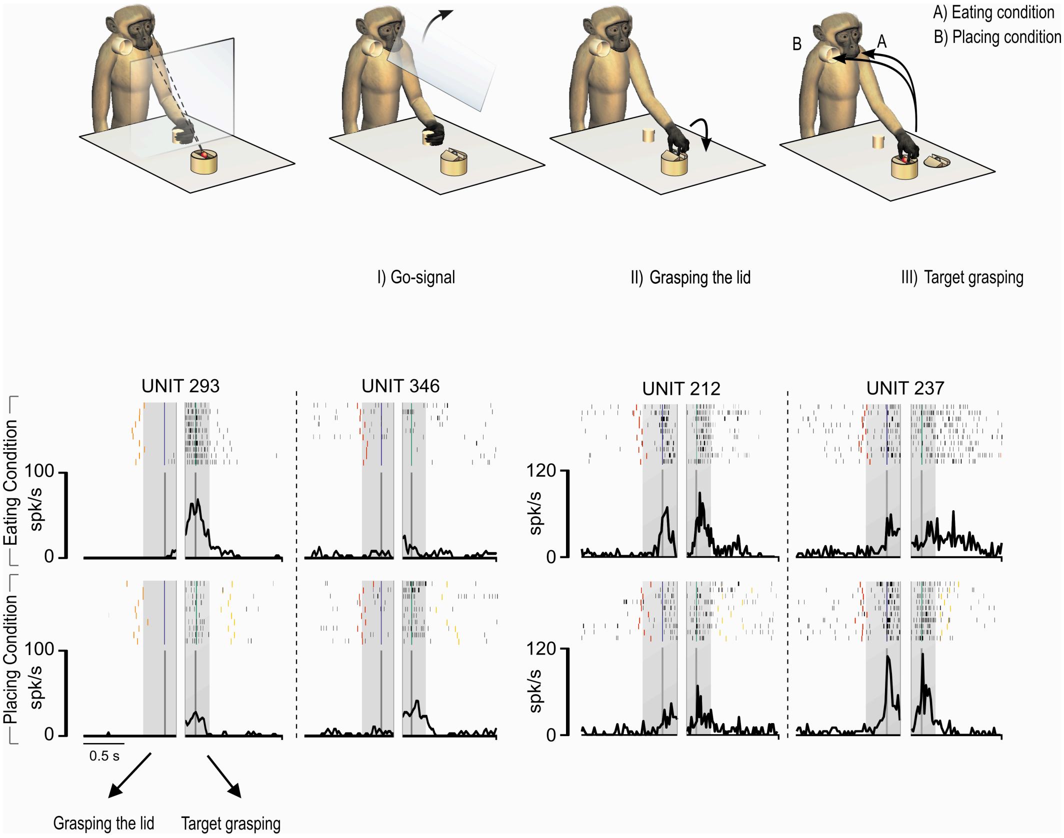


Figure 2: Top. Action sequence tasks. A transparent screen allowed the monkey to establish from the beginning which of the two action goals (eating or placing) had to be performed. Once the screen was removed (I) the monkey had to grasp a lid that covered a container (II), and then, after the removal of the lid, grasped an object inside the container in order to eating it (A) or placing it into another container (B) located near the mouth. Bottom. Activity of four neurons recorded from parietal area PFG during the task. Rasters and histograms representing neuronal activity is aligned on the first (left) and second (right) grasping act performed by the monkey. Ordinate represents neuronal discharge in spikes/sec. Abscissae indicate time course. Modified from Bonini et al. (2011).

In humans it is more difficult to directly demonstrate the existence of a mechanism involved in organizing action sequences, although a few investigations attempted to explore this issue. In an fMRI study (Frey & Gerry 2006) subjects had to just observe or to observe with the purpose to reproduce the constructions, by another individual, of multiparts objects, that required structured motor sequences. During the task the activated regions were inferior frontal and parietal cortices (part of the mirror system), dorsal premotor cortex and pre-SMA, plus subcortical motor structures. In the observation/imitation study by Buccino *et al.* (2004a), described in a previous section, in which subjects had to observe, in order to imitate, sequences of guitar chords performed by a model, the activation involved the parieto-frontal mirror system, the SMA and prefrontal cortex. Altogether, these areas seem to play a role in encoding the meaning of motor acts, in parsing the action sequence into its discrete motor elements and, vice versa, in constructing new actions starting from single motor acts.

Since the first reports of Broca's patients, the IFG region has been repetitively considered as having a crucial role in syntactic processing. However, although not denying this role, other studies provided information that make this picture more complex. In fact there are other areas such as, for example, superior temporal cortex, that are activated during syntactic processing. Moreover, the activation of IFG by syntactic construction involves also concurrent semantic coding. Lastly, this region appears to be more activated when the level of complexity of syntactic comprehension of the sentence is higher (see Gernsbacher & Kaschak 2003; Grodzinky & Friederici 2006). In addition, as it is clear also from the extension of the lesion in Broca's aphasics, an important role is also played by nearby regions, such as, for example, the middle prefrontal cortex, located just medially to IFG, that probably provides this latter region with the working memory mechanism necessary for building and understanding long and complex sentences. Altogether, these reports indicate that within the IFG region, areas contribute to processing linguistic information beyond syntactic structure. This would be in agreement with the proposal that in the IFG region (BA 44, 45 and 47) there are subsectors involved in phonology, syntax and semantics, that appear to be roughly organized along a caudal-to-rostral anatomical sequence, with a certain overlap between them (Bookheimer 2002; Hagoort 2005).

This type of organization of IFG region does not contradict its possible derivation from a system responsible for action organization. Indeed this latter requires, for its optimal functioning, a hierarchical structure, a precise order that makes the action sequence meaningful, and a coded semantics, that is observed both at the level of whole action and of the motor acts composing it, similarly to what it occurs for the meaning of a sentence and a word, respectively. In agreement with this, neuroimaging studies showed that IFG activation is higher when the meaning of a sentence depends not only from word meanings but also from the syntactic structure in which the words are ordered (Dapretto & Bokheimer 1999). On the other hand, other studies suggest that Broca's area is involved in processing hierarchical structures in multiple domains of human cognition (Thompson-Schill *et al*. 2005; Koechlin & Jubault 2006). Thus, although the transition from action to language could have been long and may have required a complex adjustment of the mechanisms involved in sequence organization, nonetheless the existence of a motor substrate endowed with a motor meaning, organized in chunks and accessible by visual and acoustic higher order input, seems an important prerequisite for both language construction and its comprehension.

References

Aitken, P. G. & W. A. Wilson Jr. 1979. Discriminative vocal conditioning in rhesus monkeys: Evidence for volitional control? *Brain Language* 8, 227–240.

Arbib, M. 2005. From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral Brain Sciences* 28, 105–167.

Aziz-Zadeh, L., M. Iacoboni, E. Zeidel, S. Wilson & J. Mazziotta. 2004. Left hemisphere action facilitation in response to manual action sounds. *European Journal of Neuroscience* 19, 2609–2612.

Bailey, P., G. von Bonin & W. S. McCulloch. 1950. *The Isocortex of the Chimpanzee*. Urbana, IL: University of Illinois Press.

Bak, T. H., D. G. O'Donovan, J. H. Xuereb, S. Boniface & J. R. Hodges. 2001. Selective impairment of verb processing associated with pathological changes in Brodmann areas 44 and 45 in the motor neurone disease–dementia–aphasia syndrome. *Brain* 124, 103–120.

Bak, T. H., D. Yancopoulou, P. J. Nestor, J. H. Xuereb, M. G. Spillantini, F. Pulvermüller & J. R. Hodges. 2006. Clinical, imaging and pathological correlates of a hereditary deficit in verb and action processing. *Brain* 129, 321–332.

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

Bernstein, N. A. 1996. On dexterity and its development. In M. L. Latash & M. T Turvey (eds.), *Dexterity and its Development*. Mahwah, NJ: Lawrence Erlbaum.

Bonini L., S. Rozzi, F. U. Serventi, L. Simone, P. F. Ferrari & L. Fogassi. 2010. Ventral premotor and inferior parietal cortices make distinct contribution to action organization and intention understanding. *Journal of Neuroscience* 20, 1372–1385.

Bonini, L., F. U. Serventi, L. Simone, S. Rozzi, P. F. Ferrari & L. Fogassi. 2011. Grasping neurons of monkey parietal and premotor cortices encode action goals at distinct levels of abstraction during complex action sequences. *Journal of Neuroscience* 31, 5876–5886.

Bookheimer, S. 2002. Functional MRI of language: New approaches to understanding the cortical organization of semantic processing. *Annual Review of Neuroscience* 25, 151–188.

Brodmann, K. 1909. *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien Dargerstellt auf Grund des Zellenbaues*. Leipzig: Barth.

Buccino, G., F. Binkofski, G. R. Fink, L. Fadiga, L. Fogassi, V. Gallese, R. J. Seitz, K. Zilles, G. Rizzolatti & H. J. Freund. 2001. Action observation activates premotor and parietal areas in a somatotopic manner: An fMRI study. *European Journal of Neuroscience* 13, 400–404.

Buccino, G., S. Vogt, A. Ritzl, G. R. Fink, K. Zilles, H.-J. Freund & G. Rizzolatti. 2004a. Neural circuits underlying imitation learning of hand actions: An event-related fMRI study. *Neuron* 42, 323–334.

Buccino, G., F. Lui, N. Canessa, I. Patteri, G. Lagravinese, F. Benuzzi, C. A. Porro & G. Rizzolatti. 2004b. Neural circuits involved in the recognition of actions performed by non-conspecifics: An fMRI study. *Journal of Cognitive Neuroscience* 16, 114–126.

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

Caggiano, V., L. Fogassi, G. Rizzolatti, P. Thier & A. Casile. 2009. Mirror neurons differentially encode the peripersonal and extrapersonal space of monkeys. *Science* 324, 403–406.

Caggiano, V., L. Fogassi, G. Rizzolatti, J. K. Pomper, P. Thier, M. A. Giese & A. Casile. 2011. View-based encoding of actions in mirror neurons of area f5 in macaque premotor cortex. *Current Biology* 21, 144–148.

Caine, N. G., R. L. Addington & T. L. Windfelder. 1995. Factors affecting the rates of food calls given by red-bellied tamarins. *Animal Behaviour* 50, 53–60.

Call, J. & M. Tomasello 2007. The gestural communication of apes and monkeys. Mahwah, NJ: Lawrence Erlbaum.

Cantalupo, C. & W. D. Hopkins. 2001. Asymmetric Broca's area in great apes. *Nature* 414, 505.

Cantalupo, C., J. Oliver, J. Smith, T. Nir, J. P. Taglialatela & W. D. Hopkins. 2009. The chimpanzee brain shows human-like perisylvian asymmetries in white matter. *European Journal of Neuroscience* 30, 431–4318.

Capirci O. & V. Volterra. 2008. Gesture and speech. The emergence and development of a strong changing partnership. *Gesture* 8, 22–44.

Casile, A. & M. A. Giese. 2006. Nonvisual motor training influences biological motion perception. *Current Biology* 16, 69–74.

Caspers, S., K. Zilles, A. R. Laird & S. B. Eickhoff. 2010. ALE meta-analysis of action observation and imitation in the human brain. *NeuroImage* 50, 1148–1167.

Cheney D. L. & R. M. Seyfarth. 1982. How vervet monkeys perceive their grunts. *Animal Behaviour* 30, 739–751.

Cochin, S., C. Barthelemy, S. Roux & J. Martineau. 1999. Observation and execution of movement: Similarities demonstrated by quantified electro-encephalograpy. *European Journal of Neuroscience* 11, 1839–1842.

Corballis, M. C. 2003. From mouth to hand: Gesture, speech, and the evolution of right-handedness. *Behavioral and Brain Sciences* 26, 199–260.

Coudè, G., P. F. Ferrari, F. Rodà, M. Maranesi, E. Borelli, V. Veroni, F. Monti, S. Rozzi & L. Fogassi. 2011. Neurons controlling voluntary vocalization in the macaque ventral premotor cortex. *PLoS One* 6(11), e26822.

Craighero, L., L. Fadiga, G. Rizzolatti & C. Umiltà. 1999. Action for perception: A motor-visual attentional effect. *Journal of Experimental Psychology Human Perception Performance* 25, 1673–1692.

Crockford, C., R. M. Wittig, R. Mundry & K. Zuberbühler. 2012. Wild chimpanzees inform ignorant group members of danger. *Current Biology* 22, 142–146.

Damasio, A. R. 1992. Aphasia. *The New England Journal of Medicine* 326, 531–539.

D'Ausilio, A., F. Pulvermüller, P. Salmas, I. Bufalari, C. Begliomini & L. Fadiga. 2009. The motor somatotopy of speech perception. *Current Biology* 19, 381–385.

Dapretto, M. & S. Y. Bookheimer. 1999. Form and content: Dissociating syntax and semantics in sentence comprehension. *Neuron* 24, 427–432.

De Renzi, E. 1989. Apraxia. In: F. Boller & J. Grafman (eds.), *Handbook of Neuropsychology*, vol. 2, 245–263. Amsterdam: Elsevier.

Di Pellegrino, G., L. Fadiga, L. Fogassi, V. Gallese & G. Rizzolatti. 1992. Understanding motor events: A neurophysiological study. *Experimental Brain*

Research 91, 176–80.

Egnor, S. E. & M. D. Hauser. 2006. Noise-induced vocal modulation in cotton-top tamarins (Saguinus oedipus). *American Journal of Primatology* 68, 1183–1190.

Eimas, P. D., E. R. Siqueland, P. Jusczyk & J. Vigorito. 1971. Speech perception in infants. *Science* 3968, 303–306.

Evarts, E. V. 1968. Relation of pyramidal tract activity to force exerted during voluntary movement. *Journal of Neurophysiology* 31, 14–27.

Fadiga, L., L. Fogassi, G. Pavesi & G. Rizzolatti. 1995. Motor facilitation during action observation: A magnetic stimulation study. *Journal of Neurophysiology* 73, 2608–2611.

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

Ferrari, P. F., V. Gallese, G. Rizzolatti & L. Fogassi. 2003. Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. *European Journal of Neuroscience* 17, 1703–1714.

Ferrari, P. F., A. Paukner, C. Ionica & S. J. Suomi. 2009. Reciprocal face-to-face communication between rhesus macaque mother and their newborn infants. *Current Biology* 19, 1768–1772.

Ferrari, P. F., S. Rozzi & L. Fogassi. 2005. Mirror neurons responding to observation of actions made with tools in monkey ventral premotor cortex. *Journal of Cognitive Neuroscience* 17, 212–226.

Ferrari, P. F., R. E. Vanderwert, A. Paukner, S. Bower, S. J. Suomi & N. A. Fox. 2012. Distinct EEG amplitude suppression to facial gestures as evidence for a mirror mechanism in newborn monkeys. *Journal of Cognitive Neuroscience* 24, 1165–1172.

Filimon, F., J. D .Nelson, D. J. Hagler & M. I. Sereno. 2007. Human cortical representations for reaching: mirror neurons for execution, observation, and imagery. *NeuroImage* 37, 1315–1328.

Fitch W. T. 2000. The evolution of speech: A comparative review. *Trends in Cognitive Sciences* 4, 258–267.

Fogassi, L. & P. F. Ferrari. 2007. From mirror neurons to embodied language. *Current Directions in Psychological Science* 16, 136–141.

Fogassi, L, P. F. Ferrari, B. Gesierich, S. Rozzi, F. Chersi & G. Rizzolatti. 2005. Parietal lobe: From action organization to intention understanding. *Science* 308, 662–667.

Frey, S. H. & V. E. Gerry. 2006. Modulation of neural activity during observational learning of actions and their sequential orders. *Journal of Neuroscience* 6, 13194–13201.

Fuster, J. 2008. *The Prefrontal Cortex*. London: Academic Press.

Gallese, V., L. Fadiga, L. Fogassi & G. Rizzolatti. 1996. Action recognition in the premotor cortex. *Brain* 119, 593–609.

Gallese, V., L. Fadiga, L. Fogassi & G. Rizzolatti. 2002. Action representation and the inferior parietal lobule. In W. Prinz & B. Hommel (eds.), *Common Mechanisms in Perception and Action: Attention and Performance, Vol. XIX*, 334–355. Oxford: Oxford University Press.

Gangitano, M., F. M. Mottaghy & A. Pascual-Leone. 2004. Modulation of premotor mirror neuron activity during observation of unpredictable grasping movements. *European Journal of Neuroscience* 20, 2193–2202.

Gardner, R. A. & B. T. Gardner. 1969. Teaching sign language to a chimpanzee. *Science* 165, 664–672.

Gazzola, V. L. Aziz-Zadeh & C. Keysers. 2006. Empathy and the somatotopic auditory mirror system in humans. *Current Biology* 16, 1824–1829.

Gentilucci, M., S. Stefanini, A. C. Roy & P. Santunione. 2004. Action observation and speech production: Study on children and adults. *Neuropsychologia* 42, 1554–1567.

Gentilucci, M. & M. C. Corballis. 2006. From manual gesture to speech: A gradual transition. *Neuroscience and Biobehavioral Reviews* 30, 949–960.

Georgopoulos, A. P., J. F. Kalaska, R. Caminiti & J. T. Massey. 1982. On the relations between the direction of two-dimensional arm movements and cell discharge in primate motor cortex. *Journal of Neuroscience* 2, 1527–1537.

Gernsbacher, M. A. & P. Kaschak. 2003. Neuroimaging studies of language. Production and comprehension. *Annual Review in Psychology* 54, 91–114.

Ghazanfar, A. A. & M. D. Hauser. 1999. The neuroethology of primate vocal communication: Substrates for the evolution of speech. *Trends in Cognitive Sciences* 3, 377–384.

Ghazanfar, A. A., J. X. Maier, K. L. Hoffman & N. K. Logothetis. 2005. Multisensory integration of dynamic faces and voices in rhesus monkey auditory cortex. *Journal of Neuroscience* 25, 5004–5012.

Giese, M. A. & T. Poggio. 2003. Neural mechanisms for the recognition of biological movements. *Nature Reviews Neuroscience* 4, 179–192.

Goldin-Meadow, S. 1999. The role of gestures in communication and thinking. *Trends in Cognitive Sciences* 3, 419–429.

Gomez, J. C. 2007. Pointing behaviors in apes and human infants: A balanced interpretation. *Child Development* 78, 729–734.

Goodglass, H. 1993. *Understanding Aphasia*. San Diego, CA: Academic Press.

Gregoriou, G. G., E. Borra, M. Matelli & G. Luppino 2006. Architectonic organization of the inferior parietal convexity of the macaque monkey. *Journal of Comparative Neurology* 496, 422–451.

Grèzes, J., N. Costes & J. Decety. 1998. Top down effect of the strategy to imitate on the brain areas engaged in perception of biological motion: A PET investigation. *Cognitive Neuropsychology* 15, 553–582.

Grodzinsky, Y. & A. D. Friederici. 2006. Neuroimaging of syntax and syntactic processing. *Current Opinion in Neurobiology* 16, 240–246.

Hagoort, P. 2005. On Broca, brain, and binding: A new framework. *Trends in Cognitive Sciences* 9, 416–423.

Hast, M. H., J. M. Fischer, A. B. Wetzel & V. E. Thompson. 1974. Cortical motor representation of the laryngeal muscles in Macaca mulatta. *Brain Research* 73, 229–240.

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

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

Hecht, E. E., L. Davis & L. A. Parr. 2011. Do chimpanzees 'mirror' others' actions? A functional neuroimaging study of action execution and observation. *Society for Neuroscience Abstracts* 932.05.

Hickok, G. 2009. Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *Journal of Cognitive Neuroscience* 21, 1229–1243.

Hickok, G. & D. Poeppel 2007. The cortical organization of speech processing. *Nature Reviews Neuroscience* 8, 393–402.

Hihara, S., H. Yamada, A. Iriki & K. Okanoya. 2003. Spontaneous vocal differentiation of coo-calls for tools and food in Japanese monkeys. *Neuroscience Research* 45, 383–389.

Hopkins, W. D., J. P. Taglialatela, J. L. Russell, T. M. Nir & J. Schaeffer. 2010. Cortical representation of lateralized grasping in chimpanzees (Pan troglodytes): A combined MRI and PET study. *PLoS ONE* 5(10), e13383.

Iacoboni, M. 2009. Neurobiology of imitation. *Current Opinion in Neurobiology* 19, 661–665.

Izumi, A., K. Kuraoka, S. Kojima & K. Nakamura. 2001. Visually guided facial actions in rhesus monkeys. *Cognitive, Affective & Behavioral Neuroscience* 1, 266–269.

Jeannerod, M. 1988. *The Neural and Behavioural Organization of Goal-Directed Movements*. Oxford: Oxford University Press.

Jeannerod, M. 1994. The representing brain: Neural correlates of motor intention and imagery. *Behavioral and Brain Sciences* 17, 187–245.

Jürgens, U. 2002. Neural pathways underlying vocal control. *Neuroscience and Biobehavioral Review* 26, 235–258.

Keller, S. S., N. Roberts & W. J. Hopkins. 2009. A comparative magnetic resonance imaging study of the anatomy, variability, and asymmetry of Broca's area in the human and chimpanzee brain. *Journal of Neuroscience* 29, 14607–14616.

Keller, S. S., M. Deppe, M. Herbin & E. Gilissen. 2011. Variability and asymmetry of the sulcal contours defining Broca's area homologue in the chimpanzee brain. *Journal of Comparative Neurology* 520(6), 1165–1180.

Koechlin, E. & T. Jubault. 2006. Broca's area and the hierarchical organization of human behavior. *Neuron* 50, 963–974.

Kohler, E., C. Keysers, M. A. Umiltà, L. Fogassi, V. Gallese & G. Rizzolatti. 2002. Hearing sounds, understanding actions: Action representation in mirror neurons. *Science* 297, 846–848.

Kraskov, A., N. Dancause, M. M. Quallo, S. Shepherd & R. N. Lemon. 2009. Corticospinal neurons in macaque ventral premotor cortex with mirror properties: A potential mechanism for action suppression? *Neuron* 64, 922–930.

Leavens, D. A., W. D. Hopkins & R. K. Thomas. 2004. Referential communication by chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology* 118, 48–57.

Leavens, D. A., W. D. Hopkins & K. A. Bard. 2005. Understanding the point of chimpanzee pointing. *Current Directions in Psychological Science* 14, 185–189.

Liberman, A. M. & I. G. Mattingly. 1985. The motor theory of speech perception revised. *Cognition* 21, 1–36.

Liebal, K., J. Call & M. Tomasello. 2004. Use of gestures sequences in chimpanzees (*Pan troglodytes*). *American Journal of Primatology* 64, 377–396.

Liebal, K. & J. Call. 2012. The origins of non-human primates' gestures. *Philosophical Transactions of the Royal Society B* 367, 118–128.

Lotto, A. J., G. S. Hickok & L. L. Holt. 2009. Reflections on mirror neurons and speech perception. *Trends in Cognitive Sciences* 13, 110–114.

Loula, F, S. Prasad, K. Harber & M. Shiffrar. 2005. Recognizing people from their movement. *Journal of Experimental Psychology: Human Perception and Performance* 31, 210–220.

Lui, F., G. Buccino, D. Duzzi, F. Benuzzi, G. Crisi, P. Baraldi, P. Nichelli, C. A. Porro & G. Rizzolatti. 2008. Neural substrates for observing and imagining non-object-directed actions. *Social Neuroscience* 3, 261–275.

Lyn, H., P. M. Greenfield, S. Savage-Rumbaugh, K. Gillespie-Lynch & W. D. Hopkins. 2011. Nonhuman primates do declare! A comparison of declarative symbol and gesture use in two children, two bonobos, and a chimpanzee. *Language and Communication* 31, 63–74.

MacNeilage, P. F. 1998. The frame/content theory of evolution of speech production. *Behavioral and Brain Sciences* 21, 499–546.

Maestripieri, D. 1996. Gestural communication and its cognitive implications in pigtail macaques (Macaca nemestrina). *Behaviour* 133, 997–1022.

McNeill, D. 1992. *Hand and Mind*. Chicago, IL: University of Chicago Press.

Merleau-Ponty, M. 1962. *Phenomenology of Perception* (Transl. C. Smith). London: Routledge.

Morrill, R. J., A. Paukner, P. F. Ferrari & A. A. Ghazanfar. 2012. Lip-smacking develops like speech. *Developmental Science* 15, 557–568.

Nishitani, N. & R. Hari. 2000. Temporal dynamics of cortical representation for action. *Proceedings of the National Academy of Sciences of USA* 97, 913–918.

Owren, M. J., J. A. Dieter, R. M. Seyfarth & D. L. Cheney. 1993. Vocalizations of rhesus (Macaca mulatta) and Japanese (M. fuscata) macaques cross-fostered between species show evidence of only limited modification. *Developmental Psychobiology* 26, 389–406.

Patterson, F. G. 1978. The gestures of a gorilla: Language acquisition in another pongid. *Brain and Language* 5, 72–97.

Perrett, D. I., M. H. Harries, R. Bevan, S. Thomas, P.J. Benson, A.J. Mistlin, A. K. Chitty, J. K. Hietanen & J. E. Ortega. 1989. Frameworks of analysis for the neural representation of animate objects and actions. *Journal of Experimental Biology* 146, 87–113.

Petitto, L. A., R. J. Zatorre, K. Gauna, E. J. Nikelski, D. Dostie & A. C. Evans. 2000. Speech-like cerebral activity in profoundly deaf people processing signed languages: Implications for the neural basis of human language. *Proceedings of the National Academy of Sciences of USA* 97, 13961–13966.

Petrides, M. & D. N. Pandya. 1994. Comparative architectonic analysis of the human and the macaque frontal cortex. In F. Boller & J. Grafman (eds.), *Handbook of Neuropsychology*, 17–58. Amsterdam: Elsevier Science.

Petrides, M., G. Cadoret & S. Mackey. 2005. Orofacial somatomotor responses in the macaque monkey homologue of Broca's area. *Nature* 435, 1235–1238.

Pierce, J. D. 1985. A review of attempts to condition operantly alloprimate vocalizations. *Primates* 26, 202–213.

Piaget, J. 1951. *The Psychology of Intelligence*. London: Routledge/Kegan Paul.

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

Poggio, T. & S. Edelman. 1990. A network that learns to recognize three-dimensional objects. *Nature* 343, 263–266.

Pollick, A. S. & F. B. de Waal. 2007. Ape gestures and language evolution. *Proceedings of the National Academy of Sciences of USA* 104, 8184–8189.

Prather, J. F., S. Peters, S. Nowicki & R. Mooney. 2008. Precise auditory-vocal mirroring in neurons for learned vocal communication. *Nature* 451, 305–310.

Premack, D. 2004. Psychology. Is language the key to human intelligence? *Science* 303, 318–320.

Price, C. J., R. J. Wise, E. A. Warburton, C. J. Moore, D. Howard, K. Patterson, R. S. Frackowiak & K. J. Friston. 1996. Hearing and saying. The functional neuro-anatomy of auditory word processing. *Brain* 119, 919–931.

Pulvermüller, F. & L. Fadiga. 2010. Active perception: Sensorimotor circuits as a cortical basis for language. *Nature Reviews Neuroscience* 11, 351–360.

Pulvermüller, F., M. Huss, F. Kherif, F. Moscoso del Prado Martin, O. Hauk & Y. Shtyrov. 2006. Motor cortex maps articulatory features of speech sounds. *Proceedings of the National Academy of Sciences of USA* 103, 7865–7870.

Rauschecker, J. P., B. Tian & M. Hauser. 1995. Processing of complex sounds in the macaque nonprimary auditory cortex. *Science* 268, 111–114.

Ricciardi, E., D. Bonino, L. Sani, T. Vecchi, M. Guazzelli, J. V. Haxby, L. Fadiga & P. Pietrini. 2009. Do we really need vision? How blind people 'see' the actions of others. *Journal of Neuroscience* 29, 9719–9724.

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

Rizzolatti, G. & G. Luppino, 2001. The cortical motor system. *Neuron* 31, 889–901.

Rizzolatti, G. & C. Sinigaglia. 2009. The functional role of the parieto-frontal mirror circuit: Interpretations and misinterpretations. *Nature Review Neuroscience* 11, 264274.

Rizzolatti, G., L. Fogassi & V. Gallese. 2009. The mirror neuron system: A motor-based mechanism for action and intention understanding. In M. Gazzaniga (ed.), *The Cognitive Neuroscience IV*, 625–640. Cambridge, MA: MIT Press.

Rizzolatti, G., R. Camarda, L. Fogassi, M. Gentilucci, G. Luppino & M. Matelli. 1988. Functional organization of inferior area 6 in the macaque monkey. II. Area F5 and the control of distal movements. *Experimental Brain Research* 71, 491–507.

Roland, P. E., B. Larsen, N. A. Lassen & E. Skinhøj. 1980. Supplementary motor area and other cortical areas in organization of voluntary movements in man. *Journal of Neurophysiology* 43, 118–36.

Romanski, L. M., B. B. Averbeck & M. Diltz. 2005. Neural representation of vocalizations in the primate ventrolateral prefrontal cortex. *Journal of Neurophysiology* 93, 734–747.

Rosenbaum, D. A., R. G. Cohen, S. A. Jax, D. J. Weiss & R. van der Wel. 2007. The problem of serial order in behavior: Lashley's legacy. *Human Movement Science* 26, 525–554.

Rozzi, S., R. Calzavara, A. Belmalih, E. Borra, G. G. Gregoriou, M. Matelli & G. Luppino. 2006. Cortical connections of the inferior parietal cortical convexity of the macaque monkey. *Journal of Neuroscience* 16, 1389–1417.

Rozzi, S., P. F. Ferrari, L. Bonini, G. Rizzolatti & L. Fogassi. 2008. Functional organization of inferior parietal lobule convexity in the macaque monkey: Electrophysiological characterization of motor, sensory and mirror responses and their correlation with cytoarchitectonic areas. *European Journal of Neuroscience* 28, 1569–1588.

Savage-Rumbaugh, S., K. McDonald, R. A. Sevcik, W. D. Hopkins & E. Rubert. 1986. Spontaneous symbol acquisition and communicative use by pygmy chimpanzees (Pan paniscus). *Journal of Experimental Psychology General* 115, 211–235.

Schenker, N. M., W. D. Hopkins, M. A. Spocter, A. R. Garrison, C. D. Stimpson, J. M. Erwin, P. R. Hof & C. C. Sherwood. 2010. Broca's area homologue in chimpanzees (Pan troglodytes): Probabilistic mapping, asymmetry, and comparison to humans. *Cerebral Cortex* 20, 730–742.

Seyfarth, R. M. & D. L. Cheney. 1997. Behavioral mechanisms underlying vocal communication in nonhuman primates. *Animal Learning & Behavior* 25, 249–267.

Seyfarth, R. M. & D. L. Cheney. 2010. Primate vocal communication. In M .L. Platt & A. A. Ghazanfar (eds.), *Primate Neuroethology*, 84–97. Oxford: Oxford University Press.

Shmuelof, L. & E. Zohary. 2008. A mirror representation of others' actions in the human anterior parietal cortex. *Nature Neuroscience* 11, 1267–1269.

Simonyan, K. & U. Jurgens. 2003. Efferent subcortical projections of the laryngeal motorcortex in the rhesus monkey. *Brain Research* 974, 43–59.

Sutton, D., C. Larson, E. M. Taylor & R. C. Lindeman. 1973. Vocalization in rhesus monkeys: Conditionability. *Brain Research* 52, 225–231.

Taglialatela, J. P., J. L. Russell, J. A. Schaeffer & W. D. Hopkins. 2008. Communicative signaling activates 'Broca's' homolog in chimpanzees. *Current Biology* 18, 343–348.

Tanji, J. 2001. Sequential organization of multiple movements: Involvement of cortical motor areas. *Annual Review of Neuroscience* 24, 631–651.

Tanji, J. & E. Hoshi. 2008. Role of the lateral prefrontal cortex in executive behavioral control. *Physiological Reviews* 88, 37–57.

Tettamanti, M., G. Buccino, M. C. Saccuman, V. Gallese, M. Danna, P. Scifo, F. Fazio, G. Rizzolatti, S. F. Cappa & D. Perani. 2005. Listening to action-related sentences activates fronto-parietal motor circuits. *Journal of Cognitive Neuroscience* 17, 273–281.

Thompson-Schill, S. L., M. Bedny & R. F. Goldberg. 2005. The frontal lobes and the regulation of mental activity. *Current Opinion in Neurobiology* 15, 219–224.

Tomasello, M. & K. Zuberbühler. 2002. Primate vocal and gestural communication. In A. Collin, M. Bekoff & G. Burghardt (eds.), *The Cognitive Animal*, 293–299. Cambridge, MA: MIT Press.

Umiltà, M. A., E. Kohler, V. Gallese, L. Fogassi, L. Fadiga, C. Keyzers & G. Rizzolatti. 2001. I know what you are doing: A neurophysiological study. *Neuron* 31, 155–165.

Umiltà, M. A., L. Escola, I. Intskirveli, F. Grammont, M. Rochat, F. Caruana, A. Jezzini, V. Gallese & G. Rizzolatti. 2008. When pliers become fingers in the monkey motor system. *Proceedings of the National Academy of Sciences of USA* 105, 2209–2213.

Utman, J. A., S. E. Blumstein & K. Sullivan. 2001. Mapping from sound to meaning: reduced lexical activation in Broca's aphasics. *Brain and Language* 79, 444–472.

Van Hooff, J. A. R. A. M. 1962. Facial expressions in higher primates. *Symposium of the Zoological Society of London* 8, 97–125.

Van Hooff, J. A. R. A. M. 1967. The facial displays of the catarrhine monkeys and apes. In D. Morris (ed.), *Primate Ethology*, 7–68. London: Weidenfield & Nicolson.

Visalberghi, E. & D. M. Fragaszy. 2002. 'Do monkeys ape?' Ten years after. In K. Dautenhahn & C. Nehaniv (eds.), *Imitation in Animals and Artifacts*. Cambridge, MA: MIT Press.

Yamaguchi, S. I. & R. E. Myers. 1972. Failure of discriminative vocal conditioning in Rhesus monkey. *Brain Research* 37, 109–111.

Yamaguchi, C. & A. Izumi. 2008. Vocal learning in nonhuman primates: Importance of vocal contexts. In N. Masataka (ed.), *The Origins of Language: Unraveling Evolutionary Forces.* Tokyo: Springer Verlag Japan.

Yee, E., S. E. Blumstein & J. C. Sedivy. 2008. Lexical semantic activation in Broca's and Wernicke's aphasia: Evidence from eye movements. *Journal of Cognitive Neuroscience* 20, 592–612.

Watkins, K. E., A. P. Strafella & T. Paus. 2003. Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia* 41, 989–994.

West, R. A. & C. R. Larson. 1995. Neurons of the anterior mesial cortex related to faciovocal activity in the awake monkey. *Journal Neurophysiology* 74, 1856–1869.

Weinrich, M., S. P. Wise & K. H. Mauritz. 1984. A neurophysiological study of the premotor cortex in the rhesus monkey. *Brain* 107, 385–414.

Zuberbühler, K. 2000. Causal cognition in a non-human primate: Field playback experiments with Diana monkeys. *Cognition* 76, 195–207.

Leonardo Fogassi and Pier Francesco Ferrari
Università di Parma
Dipartimento di Neuroscienze and Istituto Italiano di Tecnologia (RTM)
V. Volturno 39
43100 Parma
Italy
leonardo.fogassi@unipr.it and pierfrancesco.ferrari@unipr.it

From Gesture to Speech

Maurizio Gentilucci, Elisa De Stefani
& Alessandro Innocenti

One of the major problems concerning the evolution of human language is to understand how sounds became associated to meaningful gestures. It has been proposed that the circuit controlling gestures and speech evolved from a circuit involved in the control of arm and mouth movements related to ingestion. This circuit contributed to the evolution of spoken language, moving from a system of communication based on arm gestures. The discovery of the mirror neurons has provided strong support for the gestural theory of speech origin because they offer a natural substrate for the embodiment of language and create a direct link between sender and receiver of a message. Behavioural studies indicate that manual gestures are linked to mouth movements used for syllable emission. Grasping with the hand selectively affected movement of inner or outer parts of the mouth according to syllable pronunciation and hand postures, in addition to hand actions, influenced the control of mouth grasp and vocalization. Gestures and words are also related to each other. It was found that when producing communicative gestures (emblems) the intention to interact directly with a conspecific was transferred from gestures to words, inducing modification in voice parameters. Transfer effects of the meaning of representational gestures were found on both vocalizations and meaningful words. It has been concluded that the results of our studies suggest the existence of a system relating gesture to vocalization which was precursor of a more general system reciprocally relating gesture to word.

Keywords: Broca's area; gesture; human kinematics; mirror neurons; voice spectra; word

1. Introduction

The term gesture is used for describing social interactions involving especially movements of human hands and arms. Kendon (1982, 1988) classifies gestures along a continuum of 'linguisticity', observing that from gesticulation to sign languages the obligatory presence of speech declines, the presence of semantic properties increases and the idiosyncratic gestures are replaced by socially regulated signs. In other words, the formalized, linguistic component of the expression present in speech is replaced by signs going from gesticulation to sign

ISSN 1450–341

languages. Hand and arm movements are distinguished, namely gesticulation (i.e. idiosyncratic spontaneous movements of the hands and arms during speech); language-like gestures (i.e. like gesticulation, but grammatically integrated in the utterance); pantomimes (i.e. gestures without speech used in theater to communicate a story); emblems (e.g. insults and praises); sign language (i.e. a set of gestures and postures for a full fledged linguistic communication system).

Later, McNeill (2000) enriched this continuum and he divided it into four continua by using characteristics as 'relationship to speech', 'relationship to conventions', 'relationship to linguistic properties', and 'character of the semiosis'. McNeill (1992) has identified a number of different types of gestures that speakers routinely use when they talk: iconics (i.e. gestures depicting a concrete object or event and bearing a close formal relationship to the semantic content of speech); metaphorics (i.e. as iconics but depicting an abstract idea); deictics (i.e. gestures pointing to something or somebody either concrete or abstract); beats (i.e. gestures with only two phases (up/down, in/out) indexing the word or phrase it accompanies as being significant). Iconic gestures and abstract deictic gestures are called also representational (McNeill 1992; Kita 2000). McNeill (1992) is concerned with gestures similar to gesticulation as defined in Kendon's continuum (Kendon 1988; McNeill 1992). Gesticulation is the most frequent type of gesture in daily use and it covers many variants and usages. It is made chiefly with the arms and hands but is not restricted to these body parts; the head can take over as a kind of third hand if the anatomical hands are immobilized or otherwise engaged, and the legs and feet too can move in a gesture mode. McNeill (1992) claimed that there was no body 'language', but that instead gestures complement spoken language. Gesticulations would be distinct from 'emblems' because they are obligatory associated with speech while emblems and pantomimes may be delivered in utter silence (see McNeill 1992, 2000; Goldin-Meadow 1999; Kendon 2004).

There are two alternative views about the relationship between gesture and speech. The first posits that gesture and speech are two different communication systems (Levelt *et al.* 1985; Hadar *et al.* 1998; Krauss & Hadar 1999). According to this view, gesture works as an auxiliary support when the verbal expression is temporally disrupted or word retrieval is difficult. The other view (McNeill 1992; Kendon 2004) posits that gesture and speech form a single system of communication, since they are linked to the same thought processes even if they differ in expression modalities. According to the view held by McNeill (1992) and Kendon (2004), we have hypothesized that manual gestures and speech share the same control circuit (Bernardis & Gentilucci 2006; Gentilucci *et al.* 2006; Gentilucci & Corballis 2006). This hypothesis can be supported by evidence that speech itself may be a gestural system rather than an acoustic system, an idea captured by the motor theory of speech perception (Liberman *et al.* 1967) and articulatory phonology (Browman & Goldstein 1995). According to this view speech is regarded, not as a system for producing sounds, but rather one for producing mouth articulatory gestures.

We will review neurophysiological and behavioral data in order to support the point that this circuit controlling gestures and speech evolved from a circuit involved in the control of arm and mouth movements related to ingestion. We will suggest that both these circuits contributed to the evolution of the spoken language moving from a system of communication based on arm gestures (Gentilucci & Corballis 2006). These circuits are also responsible for the relations between spoken language and gesture during conversation. That is, these are specific instantiations of more general relations between the control of arm and mouth actions (Willems & Hagoort 2007).

2. Anatomical and Physiological Consideration

The link between gesture and speech (and in general language) supporting the view that gesture and speech are controlled by a same system can be the result of the activity of systems evolved from two classes of neurons recorded in monkey premotor area F5.

Based on cytoarchitectural and histochemical data, the agranular frontal cortex of macaque monkey has been parceled by Matelli and colleagues (Matelli *et al.* 1985, 1991) in the areas shown in Figure 1a. Area F1 corresponds basically to Brodmann's area 4 (primary motor cortex), and the other areas correspond to sub-divisions of Brodmann's area 6. Areas F2 and F7, which lie in the superior part of area 6, are referred to as 'dorsal premotor cortex', whereas areas F4 and F5, which lie in the inferior area 6, are referred to as 'ventral premotor cortex' (Matelli & Luppino 2000).

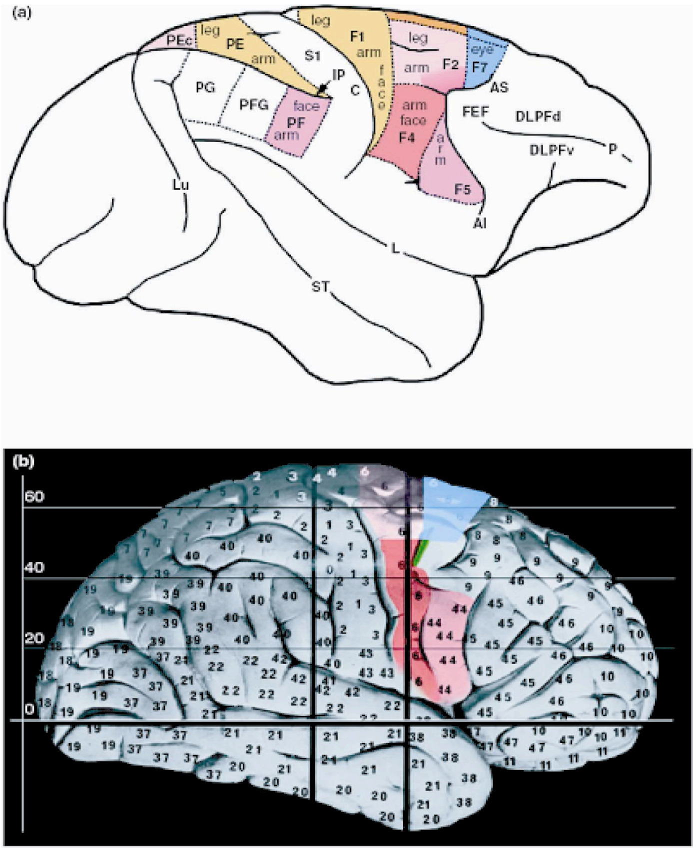


Figure 1: Lateral view of the monkey (a) and human (b) cortex

Neurophysiological studies showed that in area F5, which occupies the most rostral part of ventral premotor cortex, there is a motor representation of distal movements (Rizzolatti *et al*. 1988; Kurata & Tanji 1986; Hepp-Reymond *et al.* 1994). Functional and multi-architectonic data have demonstrated that this area is not a single entity but it consists of three main sectors: F5c, designated as 'convexity', is located on the postarcuate convexity cortex; F5p designated as 'posterior' is located on the posterior bank of the arcuate sulcus dorsally and F5a, designated as 'anterior', on the posterior bank of the same sulcus ventrally (Fig. 2a–b; see Belmalih *et al.* 2009; Gerbella *et al.* 2011).

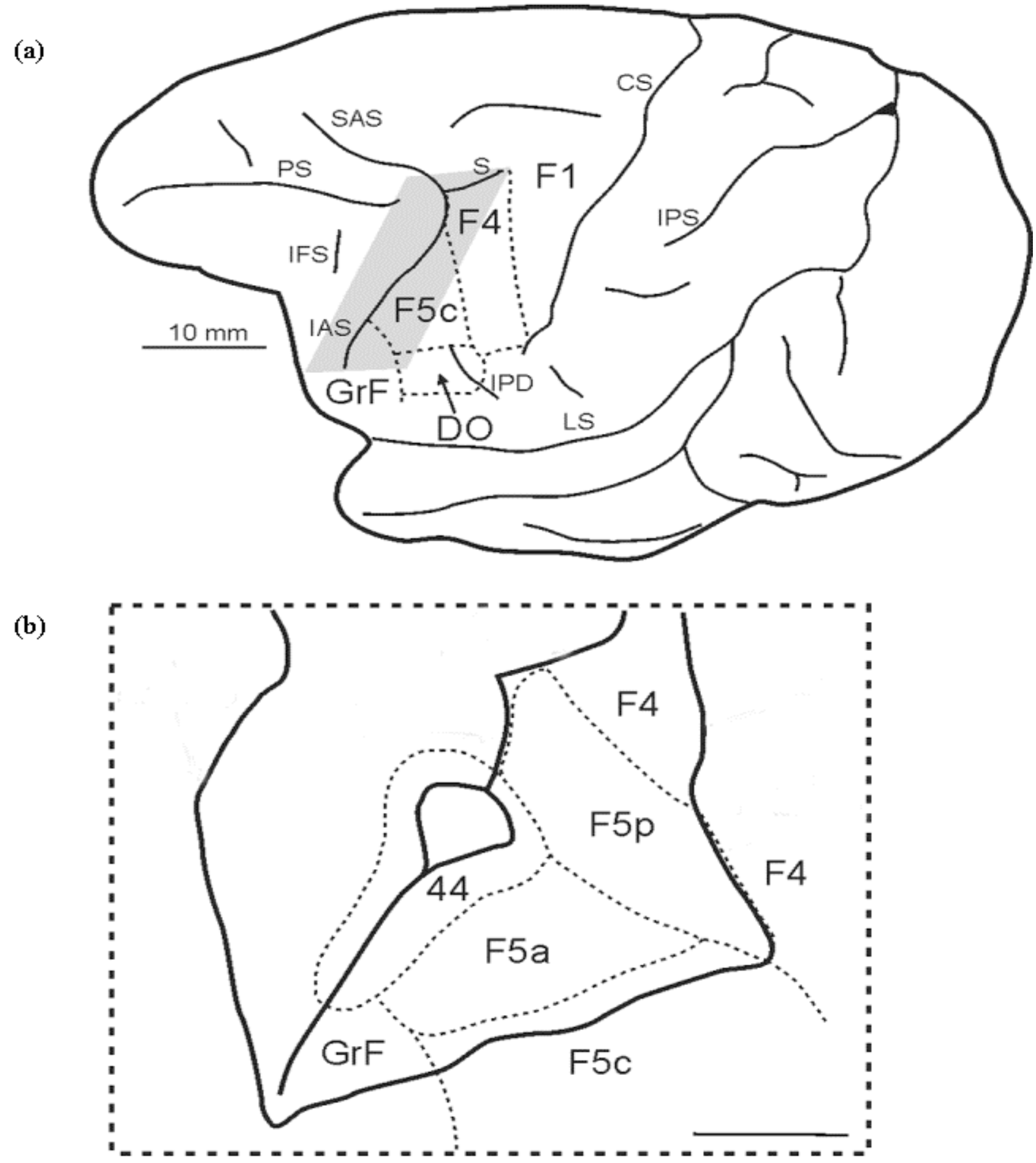


Figure 2: Architectonic maps of the macaque PMv as proposed by Belmalih et al. (2009)

In area F5 two classes of neurons were recorded, which might have been instrumental in the development of a system controlling speech and gestures. The first class of neurons frequently recorded in the posterior part of the inferior postarcuate bank (F5 sector of the arcuate bank in Rizzolatti & Luppino 2001; sector F5c in Gerbella *et al.* 2011) commands grasp actions with hand and mouth (Rizzolatti *et al.* 1988). A typical neuron of this class discharges when the animal grasps a piece of food with its mouth or when the animal grasps the same piece of food with the hand contralateral or ipsilateral to the recorded cortical side. Frequently, the discharge of this class of neurons is selective for a specific type of grasp (for example, a neuron discharges when a precision grasp is used, but not for a power one), and it can be even elicited by the visual presentation of a graspable object, provided that its size is congruent with the type of grasp coded by the neuron ('canonical neurons'; see Murata *et al.* 1997; Rizzolatti *et al.* 1988). Rizzolatti *et al.* proposed that these neurons are involved in coding the aim of the grasp action, i.e. to take possession of an object. From a functional point of view, these neurons can be involved in planning a strategy in order to perform successive grasp actions. For example, they can command the grasp of an object with the hand while preparing the mouth to grasp the same object. From an evolutionary point of view, this circuit of commands might have evolved a system of hand–mouth double command, becoming instrumental in the transfer of a manual gesture communication system, from movements of the hand to movements of the mouth. That is, this system might have been used in language evolution (Gentilucci & Corballis 2006) according to the proposal that language evolved from manual gestures rather than from vocalizations. Indeed, whereas vocalizations of non-speaking primates are mainly related to emotional states, manual actions can provide more obvious iconic links with objects and, consequently, they might have been initially used to represent the physical world (Hewes 1973; Donald 1991; Corballis 1992, 2002; Givòn 1995; Armstrong *et al.* 1995; Rizzolatti & Arbib 1998; Armstrong 1999; Arbib 2005; Ruben 2005; Gentilucci & Corballis 2006). This coupling between hand and mouth, used to transfer a gesture communication system from movements of the hand to movements of the mouth, could also evolve in a system functionally relating gesture and speech.

The second class (the so called 'mirror neurons', MNs) becomes active when the animal executes a transitive action (i.e. acted upon an object) with the hand and when it observes the same action performed by another individual (Gallese *et al.* 1996). In addition, Ferrari *et al.* (2003) recorded discharges in the premotor area F5 of monkeys both from mirror neurons during lip-smacking (the most common communicative facial gesture in monkeys) and from other mirror neurons during mouth movements related to eating. This suggests that non-vocal facial gestures may be indeed transitional between visual gesture and speech.

Finally, mirror neurons in monkey have been also recorded in the rostral part of the inferior parietal lobule (Gallese et al. 2002), and neurons only activated by the observation of movements of different body effectors were recorded in the superior temporal sulcus region (Perrett *et al.* 1989).

According to Rizzolatti and colleagues (Gallese *et al.* 1996; Rizzolatti *et al.* 1996), the mirror neuron activity is involved in representing actions. This motor representation, by matching observation with execution, makes it possible for individuals to understand observed actions. In this way, individuals are able to recognize the meaning and the aim of actions performed by another individual. Therefore, providing a link between an actor and an observer, similar to the one existing between a sender and a receiver of a message, mirror neurons may have played a role in the development of a gestural communication system. Thanks to this mechanism, actions done by other individuals become messages that are understood by an observer. In the present review, we will focus on the system of double motor commands to hand and mouth rather than the mirror system and how it evolved in humans in order to transfer a communication system based on arm actions to mouth postures and it acquired the capability to interact with speech. The role of the mirror system in the construction of a communication system in humans has been reviewed elsewhere (Gentilucci & Corballis 2006; Gentilucci & Dalla Volta 2008; Gentilucci *et al.* 2008).

3. Relations between Execution of Hand/Arm Actions and Speech

A system of double grasp commands to hand and mouth seems to be still active in modern humans, as resulted by the following behavioral studies. Gentilucci *et al.* (2001) showed that, when participants were instructed to open their mouth while grasping objects with their hand, the size of mouth opening increased with the size of the grasped object. The kinematic analysis showed that, concurrently to an increase in kinematic parameters of the finger shaping during the grasp of the large as compared to the small object, there was an increase in the parameters of lip opening even if the participants were required to open their mouth of a fixed amount. Conversely, when the participants opened their fingers while grasping objects with their mouth, the size of the hand opening increased also with the size of the object. Control experiments showed that neither the simple observation of the object nor the proximal component of the reach was responsible for the effect (Gentilucci *et al*. 2001). Recent evidence suggests that even postures of distal effectors affect grasp. Gentilucci & Campione (2011) found that hand postures, in addition to hand actions, influenced the control of mouth grasp. In two experiments, participants reached different objects with their head and grasped them with their mouth, after assuming different hand postures. In one experiment the hand could mimic the holding of a large or small object or it could be relaxed, whereas in the other experiment the hand fingers could be extended or flexed or relaxed (Fig. 3A–B). The latter experiment was a control experiment whose results could be compared with those of experiments 1 and 2 in which the effects of postures of the mouth (open/closed) and toes (extended/flexed) on hand grasp were studied. In both experiments, the kinematics of lip shaping during grasp varied congruently with the posture assumed by the hand, i.e. it was larger or smaller when it could be explicitly (experiment 1) or implicitly (experiment 2) associated with the grasping of large or small objects, respectively.

Figure 3: Experimental set-up, stimuli, procedure and examples of trajectories in Gentilucci & Campione (2011)

In the successive experiment 3 participants were required to open or to close their mouth, or to maintain it relaxed (Fig. 3C). Then, they performed a manual grasp, maintaining that mouth posture. Maximal finger aperture was larger when the mouth was opened as compared to when it was closed. An intermediate aperture was observed in the relaxed mouth condition. The results of these experiments extend the effects of motor interactions with objects to postures of effectors; specifically, the posture of one effector (the mouth or the hand) can be a template for the configuration that will be assumed by the other grasping effector (the hand or the mouth) during shaping. Finally, a control experiment verified whether similar relation also exists between foot and hand. Indeed, previous experiments did not verify whether the reciprocal interactions between postures and actions were specific for hand and mouth or they could be extended to other distal effectors, as, for example, the foot.

Participants executed a manual grasp of an object while their right toes were extended or flexed or relaxed. No significant effect of the foot posture was found on maximal finger aperture. This result disproves a link between hand movements and foot postures: on the contrary, a link was preferentially found between hand and mouth. However, evidence (Baldissera *et al.* 2006) does suggest that the control of hand movements can be associated to the control of foot movements (i.e. during coupled hand and foot oscillations a synchronism between these effectors was observed). To explain this apparent contradiction, we can consider that in modern humans the grasping foot has lost the capacity of activating different interactions with objects of different size and shape. For this reason, despite a clock-movement synchronization, hand and foot do not interact with each other, like hand and mouth, which, on the contrary, are both capable of activating different interactions with objects. Also, from an anatomical point of view (see Buccino *et al.* 2001) premotor area where foot is represented is separated from premotor area where the hand is represented. On the contrary, hand and mouth areas are adjacent and partially overlap.

If this system, coupled with the mirror system, is also used to share a communication gestural repertoire of the hand with the speaking mouth, the execution of transitive actions should affect speech, and specifically the production of phonological units. Gentilucci & colleagues required participants to reach and grasp small and large objects while pronouncing syllables (Gentilucci *et al.* 2001). They found that when grasping the large objects as compared to the small objects, the lip opening and parameters of the vowel vocal spectra increased. Conversely, the pronunciation of a vowel during the entire execution of the grasp affected maximal finger aperture (Gentilucci & Campione 2011). Specifically, the vowel /a/ induced an increase in maximal finger aperture if compared to /i/. The vocalization /ɔ/ induced an intermediate effect. The vowel /a/ is characterized by higher Formant 1 (F1; depending on internal mouth aperture) and lower Formant 2 (F2; depending on tongue protrusion; Leoni & Maturi 2002). In contrast, /i/ is characterized by lower F1 and higher F2. The vowel /ɔ/ has intermediate values. In sum, configurations of the internal mouth related to vocalizations seem to be responsible for effects on finger shaping during grasping. This coupling can be precursor of more complex interactions between gestures and words.

4. Interactions between Gestures and Words

Chieffi *et al.* (2009) studied the relations between production of deictic gestures (HERE, i.e. a pointing directed towards the agent's body, and THERE, i.e. a pointing directed towards a remote point far from the agent's body) and the simultaneous pronunciation of the words QUA 'here' and LÁ 'there'. The authors found facilitation/interference when the meaning of word was congruent/incongruent with the gesture direction; that is, the gesture was quicker in the case of congruence with word meaning. This can be explained by considering that direction was stressed by the word. The reverse occurred in the case of incongruent meaning; that is, the direction was ambiguous because the direction coded by the word was opposite. Consequently, gesture was slowed down. A non-alternative explanation is a priming effect of the word on arm velocity. This suggests an interaction at a higher level due to the presentation of linguistic stimuli.

The relations between gestures and words were also studied when communicative signals like CIAO, NO, STOP were produced (Bernardis & Gentilucci 2006; Gentilucci *et al.* 2006; Barbieri *et al.* 2009). The main finding of these studies was that the social intention, i.e. the intention to interact directly with a conspecific (depending on the communicative meaning of the signal) was transferred from gestures (i.e. emblems) to words, modifying some voice parameters. In turn, following this transfer, the mouth controller modified the hand/arm kinematics by slowing down it. This could be consequent to the fact that the transferred aspects of the social intention coded in the gesture became redundant.

Summing up, in all studies (Bernardis & Gentilucci 2006; Gentilucci *et al.* 2006; Barbieri *et al.* 2009; Chieffi *et al.* 2009) gesture and speech interacted with each other by reciprocally transferring aspects of the signal meaning. Obviously, these aspects differed according to the type of signal.

5. Are the Relations between Hand Postures and Vocalizations Precursors of Relations between Gestures and Speech?

In a previous study, Gentilucci & Campione (2011) found that when subjects pronounced the open vowel /a/, which is characterized by a larger aperture of the internal mouth, the finger shaping of a simultaneous grasp was larger than when they pronounced the closed vowel /i/, which is characterized by a smaller internal mouth aperture. In a subsequent study, Gentilucci *et al.* (2012) reasoned that if the relation between hand actions and vocalizations is precursor of the relation between gesture and speech, same or similar effects of meaningful gestures on both simple vocalizations and words should be found. In this study unimanual/bimanual gestures LARGE and SMALL were contemporaneously presented with a vignette close to the actor in which, in experiment 1, either the vowel ‘A’ (/a/) or ‘I’ (/i/) was printed, in experiment 2 the word GRÀNDE ‘large’ or PÌCCOLO ‘small’, and in experiment 3 the pseudo-words SCRÀNTA or SBÌCCARA (Fig. 4).

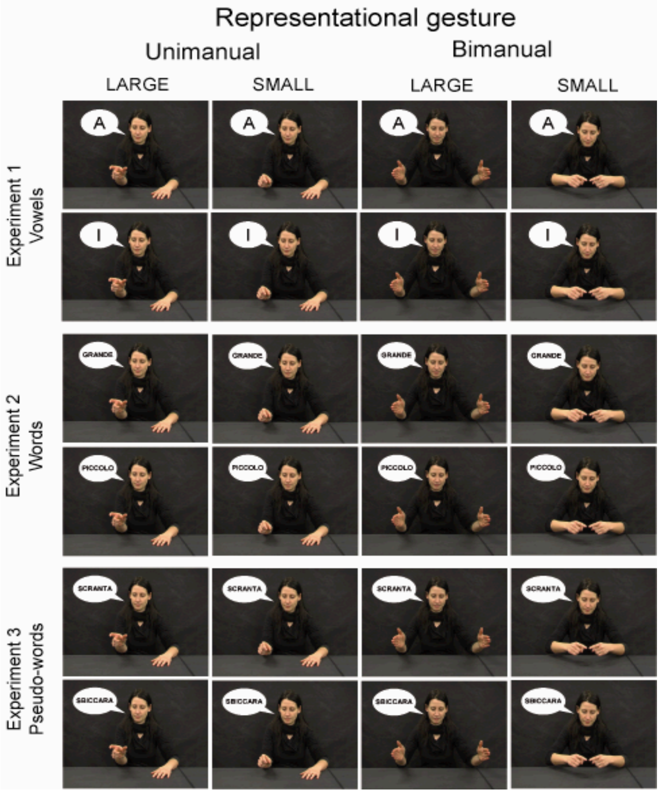


Figure 4: Stimuli presented in experiments 1–3. The panels show all the combinations between gestures and printed vowels (‘A’ and ‘I’; experiment 1) or gestures and printed words (GRANDE, PICCOLO; experiment 2) or gestures and printed pseudo-words (SCRANTA, SBICCARA)

Unimanual gestures affected formant 1 (F1) of voice spectra of the two vowels pronounced alone. This parameter, which is directly related to internal mouth aperture (Leoni & Maturi 2002), increased after gesturing LARGE as compared to SMALL (Fig. 5). F1 of the vowels /a/ and /i/ included in the words GRÀNDE 'large' and PÌCCOLO 'small', respectively, were greater when gesturing LARGE in bimanual condition as compared to the other conditions (Fig. 5). In contrast, F1 of vowels included in the pseudo-words increased when gesturing LARGE in both unimanual and bimanual conditions (Figure 5).

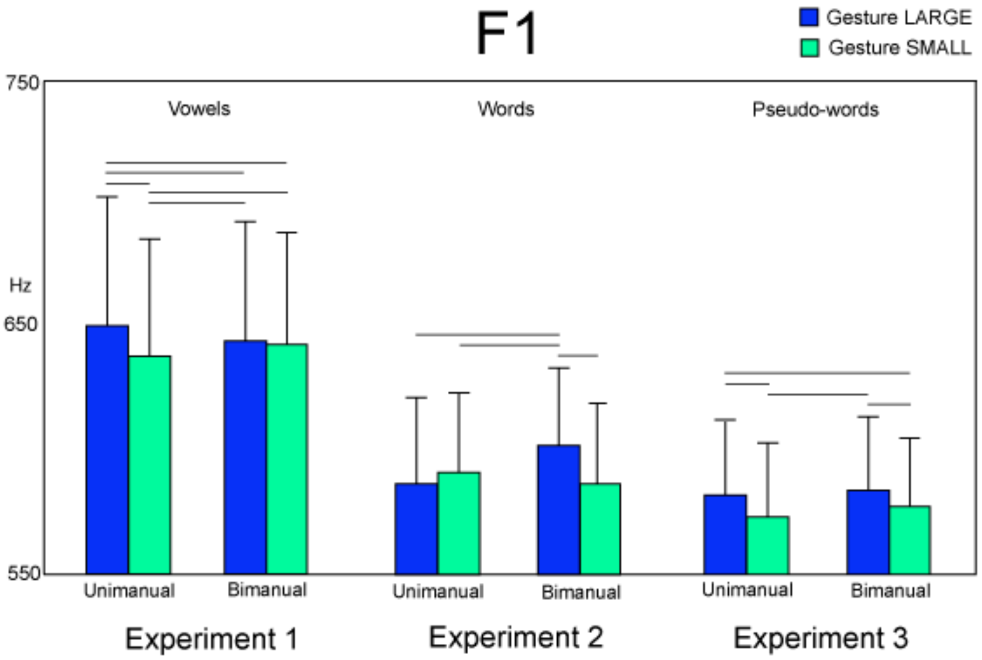


Figure 5: Effects of gestures on Formant 1 (F1) of vowels (experiment 1) or vowels included in words (experiment 2) or vowels included in pseudo-words (experiment 3) pronounced after production of the gestures LARGE and SMALL in unimanual and bimanual conditions. Horizontal bars represent significance or trend to significance in the ANOVAs, whereas vertical bars represent SE.

Summing up, the control of pronunciation of vowels alone was associated to the control of unimanual hand gestures only, according to the hypothesis that the internal mouth and the right hand are controlled by the same system and the two effectors are directly connected (Gentilucci *et al.* 2001; Gentilucci & Campione 2011). Moreover, they support the existence of a more general system reciprocally relating word and gesture meanings. Word meaning was responsible for categorization of all the gestures in LARGE or SMALL. This process was used to construct a size representation common to all the gesturing effectors in which the absolute size was computed. Consequently, the bimanual gesture LARGE was the only categorized as LARGE because the represented size was much greater than the sizes represented by the other gestures: these, conversely, were categorized as SMALLs. In turn, the meaning of the categorized gesture affected word pronunciation. Finally, a size representation not yet independent of the gesturing effectors was activated when pronouncing pseudo-words.

Summing up, we found similarity in the relations between gestures and vocalizations and between gestures and words. However, the differences between the two systems should be discussed. The system gesture/vocalization seems to be simpler since it couples right hand postures with mouth postures. In contrast the system gesture/word seems to be more complex and distributed. This system is involved in a process of abstraction since a size categorization is performed in which the absolute distance rather than that relative to the effector is taken into account.

The gesture LARGE induced an increase in F1 of /i/ of the word PICCOLO and the gesture SMALL induced a decrease in the F1 of /a/ of the word GRANDE; that is, the gesture did not selectively affect the vowels of words whose meaning could be or not associated to the gesture meaning. This result may be explained following the hypothesis that the system relating words and gestures derives from a system relating assumed postures of the hand and simple vowel pronunciation, i.e. due to internal mouth posture. This effect was not selective for vowels (i.e. /a/ vs /i/) and probably this property was conserved in the evolution of the system. This produced predominance of gesture meaning on word meaning in order that the gesture could modulate the meaning of a word. For example the word PICCOLO 'small' could be differently interpreted and pronounced if accompanied by the gesture LARGE or SMALL, respectively. Specifically, the word could be interpreted as less small if accompanied by the gesture LARGE and conversely smaller if accompanied by the gesture SMALL.

Kelly *et al.* (2004) carried out an Event Related Potential (ERP) experiment in which participants saw an actor producing a representational gesture expressing the property like width or height. If the gesture was preceded by a spoken word expressing a different property, a stronger deflexion was observed in ERPs (N400 effect). In many language studies, N400 effect was found when semantic process is harder to integrate into the previous context (for a review, see Kutas & van Petten 1994). Consequently, Kelly *et al.* (2004) interpreted their results as consequent to semantic processing of the gesture. Other studies (Wu & Coulson 2005; Holle & Gunter 2007; Kelly *et al.* 2007; Ozyurek *et al.* 2007) confirmed an N400 effect for incongruence between word and gesture. The data of the study by Gentilucci *et al.* (2012) are in agreement with the idea about a semantic processing of the gesture. Indeed, from a functional point of view the gestures were categorized according to the meaning of the words, and, in turn the meaning of the gestures modulated the meaning of the word.

6. Final Anatomical Considerations

Previously, Gentilucci *et al.* (2006) proposed that Broca's area in Inferior Frontal Gyrus (IFG) plays a role in the reciprocal control between gesture and speech. On the basis of the results by the Gentilucci *et al.*'s study (2012) we extend this proposal; we suggest the existence of two partially overlapping circuits involved in the reciprocal control between gesture and speech. The first is related to the control of vocalization and unimanual gestures (both transitive actions and meaningful intransitive gestures). This circuit can be remnant of the circuit controlling the grasp with the hand and the mouth and it may be located in pars orbitalis of IFG (area BA44, Fig. 1b). This area is involved in encoding phonological representations in terms of mouth articulation gestures (Demonet *et al.* 1992; Zatorre *et al.* 1992; Paulesu *et al.* 1993), in manipulation of complex objects (Binkofski *et al.* 1999), and is part of the human mirror circuit (Gazzola & Keysers 2009; Kilner *et al.* 2009; for a review, see Rizzolatti & Craighero 2004). The second circuit is more involved in the relations between gesture and speech concerning the semantics of the signals. This circuit is enlarged as compared to the first one and may also comprises pars triangularis and/or pars orbicularis of IFG (areas BA45, BA47; Fig. 1b), sectors which are more related to semantics than phonology (Bookheimer 2002). In previous neuro-imaging studies, Willems *et al.* (2007) and Xu *et al.* (2009) found a common circuit comprising pars opercularis, triangularis, and orbitalis of IFG which was activated by the processing of speech or gesture. It might allow a common access of words and gestures to semantics in order to integrate the two signals. In the present study, categorization of unimanual and bimanual gestures on the basis of word meaning might take place in this circuit. In addition, in this circuit transferring aspects of gesture meaning (i.e. the size) to the word might also occur and, consequently, its pronunciation might change. In sum, an enlarged circuit, whose primary (and precursor) nucleus allows a direct communication between vocalization and unimanual gestures (both actions and meaningful gestures), was involved in controlling gestures and words.

References

Arbib, Michael A. 2005. From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences* 28, 105–124.

Armstrong, David F., William C. Stokoe & Sherman E. Wilcox. 1995. *Gesture and the Nature of Language.* Cambridge: Cambridge University Press.

Armstrong, David F. 1999. *Original Signs: Gesture, Sign, and the Source of Language.* Washington, DC: Gallaudet University Press.

Baldissera, Fausto G., Paolo Cavallari & Roberto Esposti. 2006. Synchrony of hand-foot coupled movements: Is it attained by mutual feedback entrainment or by independent linkage of each limb to a common rhythm generator? *BMC Neuroscience* 26, 7–70.

Barbieri, Filippo, Antimo Buonocore, Riccardo Dalla Volta & Maurizio Gentilucci. 2009. How symbolic gestures and words interact with each other. *Brain and Language* 110, 1–11.

Belmalih, Abdelouahed, Elena Borra, Massimo Contini, Marzio Gerbella, Stefano Rozzi & Giuseppe Luppino. 2009. Multimodal architectonic subdivision of the rostral part (area F5) of the macaque ventral premotor cortex. *The Journal of Comparative Neurology* 512, 183–217.

Bernardis, Paolo & Maurizio Gentilucci. 2006. Speech and gesture share the same communication system. *Neuropsychologia* 44, 178–190.

Binkofski, Ferdinand, Giovanni Buccino, Stefan Posse, Rudiger J. Seitz, Giacomo Rizzolatti & Hans-Joachim Freund. 1999. A fronto-parietal circuit for object manipulation in man: Evidence from an fMRI-study. *The European Journal of Neuroscience* 11, 3276–3286.

Bookheimer, Susan. 2002. Functional MRI of language: New approaches to understanding the cortical organization of semantic processing. *Annual Review of Neuroscience* 25, 151–188.

Browman, Catherine P. & Louis Goldstein. 1995. Dynamics and articulatory phonology. In Robert F. Port & Timothy van Gelder (eds.), *Mind as Motion*, 175–193. Cambridge, MA: MIT Press.

Buccino, Giovanni, Ferdinand Binkofski, Gereon R. Fink, Luciano Fadiga, Leonardo Fogassi, Vittorio Gallese, Rudiger J. Seitz, Karl Zilles, Giacomo Rizzolatti & Hans J. Freund. 2001. Action observation activates premotor and parietal areas in a somatotopic manner: An fMRI study. *European Journal of Neuroscience* 13, 400–404.

Chieffi, Sergio, Claudio Secchi & Maurizio Gentilucci. 2009. Deictic word and gesture production: Their interaction. *Behavioural Brain Research* 203, 200–206.

Corballis, Michael C. 1992. On the evolution of language and generativity**.** *Cognition* 44, 197–226.

Corballis, Michael C. 2002. *From Hand to Mouth: The Origins of Language*. Princeton, NJ: Princeton University Press.

Démonet, Jean-Francois, Francois Chollet, Stuart Ramsay, Dominique Cardebat, Jean-Luc Nespoulous, Richard Wise, André Rascol & Richard S. Frackowiak. 1992. The anatomy of phonological and semantic processing in normal subjects. *Brain* 115, 1753–1768.

Donald, Merlin. 1991. *Origins of the Modern Mind*. Cambridge, MA: Harvard University Press.

Ferrari, Pier F., Vittorio Gallese, Giacomo Rizzolatti & Leonardo Fogassi. 2003. Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. *European Journal of Neuroscience* 17, 1703–1714.

Gallese, Vittorio, Luciano Fadiga, Leonardo Fogassi & Giacomo Rizzolatti. 1996. Action recognition in the premotor cortex. *Brain* 119, 593–609.

Gallese, Vittorio, Luciano Fadiga, Leonardo Fogassi & Giacomo Rizzolatti. 2002. Action representation and the inferior parietal lobule. In Wolfgang Prinz & Bernhard Hommel (eds.), *Attention and Performance XIX: Common Mechanisms in Perception and Action*, 334–355. Oxford: Oxford University Press.

Gazzola, Valeria & Christian Keysers. 2009. The observation and execution of actions share motor and somatosensory voxels in all tested subjects: Single-subject analyses of unsmoothed fMRI data. *Cerebral Cortex* 19, 1239–1255.

Gentilucci, Maurizio, Francesca Benuzzi, Massimo Gangitano & Silvia Grimaldi. 2001. Grasp with hand and mouth: A kinematic study on healthy subjects. *Journal of Neurophysiology* 86, 1685–1699.

Gentilucci, Maurizio, Paolo Bernardis, Girolamo Crisi & Riccardo Dalla Volta. 2006. Repetitive transcranial magnetic stimulation of Broca’s area affects verbal responses to gesture observation. *Journal of Cognitive Neuroscience* 18, 1059–1074.

Gentilucci, Maurizio & Michael C. Corballis. 2006. From manual gesture to speech: A gradual transition. *Neuroscience and Biobehavioral Reviews* 30, 949–960.

Gentilucci, Maurizio & Riccardo Dalla Volta. 2008. Spoken language and arm gestures are controlled by the same motor control system. *Quarterly Journal of Experimental Psychology* 61, 944–957.

Gentilucci, Maurizio, Riccardo Dalla Volta & Claudia Gianelli. 2008. When the hands speak. *Journal of Physiology Paris* 102, 21–30.

Gentilucci, Maurizio & Giovanna C. Campione. 2011. Do postures of distal effectors affect the control of actions of other distal effectors? Evidence for a system of interactions between hand and mouth. *PLoS One* 6, e19793.

Gentilucci, Maurizio, Giovanna C. Campione, Elisa De Stefani & Alessandro Innocenti. 2012. Is the coupled control of hand and mouth postures precursor of reciprocal relations between gestures and words? *Behavioural Brain Research* 233, 130–140.

Gerbella, Marzio, Abdelouahed Belmalih, Elena Borra, Stefano Rozzi & Giuseppe Luppino. 2011. Cortical connections of the anterior (F5a) subdivision of the macaque ventral premotor area F5. *Brain Structure and Function* 216, 43–65.

Givòn, Talmi. 1995. *Functionalism and Grammar*. Amsterdam: John Benjamins.

Goldin-Meadow, Susan. 1999. The role of gesture in communication and thinking. *Trends in Cognitive Science* 3, 419–429.

Hadar, Uri, Dafna Wenkert-Olenik, Robert M. Krauss & Nachum Soroker. 1998. Gesture and the processing of speech: Neuropsychological evidence. *Brain and Language* 62, 107–126.

Hepp-Reymondd, Marie-Claude, Erhard J. Hüsler, Marc A. Maier & Hui-Xin Qi. 1994. Force-related neuronal activity in two regions of the primate ventral premotor cortex. *Canadian Journal of Physiology and Pharmacology* 72, 571–579.

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

Holle, Henning & Thomas C. Gunter. 2007. The role of iconic gestures in speech disambiguation: ERP evidence. *The Journal of Neuroscience* 19, 1175–1192.

Kelly, Spencer D., Corinne Kravitz & Michael Hopkins. 2004. Neural correlates of bimodal speech and gesture comprehension. *Brain and Language* 89, 253–260.

Kelly, Spencer D., Sarah Ward, Peter Creigh & James Bartolotti. 2007. An intentional stance modulates the integration of gesture and speech during comprehension. *Brain and Language* 101, 222–233.

Kendon, Adam. 1982. The study of gesture: Some remarks on its history. *Recherches Séiotiques/Semiotic Inquiry* 2, 25–62.

Kendon, Adam. 1988. How gestures can become like words. In Fernando Poyatos (ed.), *Crosscultural Perspectives in Nonverbal Communication*, 131–141. Toronto: C. J. Hogrefe Publishers.

Kendon, Adam. 2004. *Gesture: Visible Action as Utterance.* Cambridge: Cambridge University Press.

Kilner, James M., Alice Neal, Nikolaus Weiskopf, Karl J. Friston & Chris D. Frith.

2009. Evidence of mirror neurons in human inferior frontal gyrus. *The Journal of Neuroscience* 29, 10153–10159.

Kita, Sotaro. 2000. How representational gestures help speaking. In Donald McNeill (ed.), *Language and Gesture,* 162–185. Cambridge: Cambridge University Press.

Krauss, Robert M. & Uri Hadar. 1999. The role of speech-related arm/hand gestures in word retrieval. In Lynn Messing & Ruth Campbell (eds.), *Gesture, Speech, and Sign*, 93–116. Oxford: Oxford University Press.

Kurata, Kiyoshi & Jun Tanji. 1986. Premotor cortex neurons in macaques: Activity before distal and proximal forelimb movements. *The Journal of Neuroscience* 6, 403–411.

Kutas, Marta & Cyma Van Petten. 1994. Psycholinguistics electrified: Event-related brain potential investigations. In Morton A. Gernsbacher (ed.), *Handbook of Psycholinguistics*, 83–143. San Diego, CA: Academic Press.

Leoni, Federico A. & Pietro Maturi. 2002. *Manuale di Fonetica,* 3rd edn. Rome: Carocci.

Levelt, Willem J. 1985. Pointing and voicing in deictic expressions. *Journal of Memory and Language* 24, 133–164.

Liberman, Akiva M., Franklin S. Cooper, Donald P. Shankweiler & Michael Studdert-Kennedy. 1967. Perception of the speech code. *Psychological Review* 74, 431–461.

Matelli, Massimo, Giuseppe Luppino & Giacomo Rizzolatti. 1985. Patterns of cytochrome oxidase activity in the frontal agranular cortex of the macaque monkey. *Behavioural Brain Research* 18, 125–136.

Matelli, Massimo, Giuseppe Luppino & Giacomo Rizzolatti. 1991. Architecture of superior and mesial area 6 and the adjacent cingulate cortex in the macaque monkey. *The Journal of Comparative Neurology* 311, 445–462.

Matelli, Massimo & Giuseppe Luppino. 2000. Parietofrontal circuits: Parallel channels for sensory-motor integrations. *Advances in Neurolology* 84, 51–61.

McNeill, Donald. 1992. *Hand and Mind: What Gestures Reveal about Thought*. Chicago, IL: University of Chicago Press.

McNeill, Donald (ed.). 2000. *Language and Gesture*. Cambridge: Cambridge University Press.

Murata, Akira, Luciano Fadiga, Leonardo Fogassi, Vittorio Gallese, Vassilis Raos & Giacomo Rizzolatti. 1997. Object representation in the ventral premotor cortex (area F5) of the monkey. *Journal of Neurophysiology* 78, 2226–2230.

Ozyürek, Asli, Roel M. Willems, Sotaro Kita & Peter Hagoort. 2007. On-line integration of semantic information from speech and gesture: Insights from event-related brain potentials. *Journal of Cognitive Neuroscience* 19, 605–616.

Paulesu, Eraldo, Chris D. Frith & Richard S. Frackowiak. 1993. The neural correlates of the verbal component of working memory. *Nature* 362, 342–345.

Perrett, David I., Mark H. Harries, Ruth Bevan, S. Thomas, P. J. Benson, A. J. Mistlin, A. J. Chitty, J. K. Hietanen & J. E. Ortega. 1989. Frameworks of analysis for the neural representation of animate objects and actions. *The Journal of Experimental Biology* 146, 87–113.

Rizzolatti, Giacomo, Rosolino Camarda, Leonardo Fogassi, Maurizio Gentilucci, Giuseppe Luppino & Massimo Matelli. 1988. Functional organization of inferior area 6 in the macaque monkey. II. Area F5 and the control of distal movements. *Experimental Brain Research* 71, 491–507.

Rizzolatti, Giacomo, Luciano Fadiga, Vittorio Gallese & Leonardo Fogassi. 1996. Premotor cortex and the recognition of motor actions. *Cognitive Brain Research* 3, 131–141.

Rizzolatti, Giacomo & Michael A. Arbib. 1998. Language within our grasp. *Trends in Neurosciences* 21, 188–194.

Rizzolatti, Giacomo & Giuseppe Luppino. 2001. The cortical motor system. *Neuron* 31, 889–901.

Rizzolatti, Giacomo & Laila Craighero. 2004. The mirror-neuron system. *Annual Review of Neuroscience* 27, 169–192.

Ruben, Robert J. 2005. Sign language: Its history and contribution to the understanding of the biological nature of language. *Acta Oto-Laryngologica* 125, 464–467.

Willems, Roel M. & Peter Hagoort. 2007. Neural evidence for the interplay between language, gesture, and action: A review. *Brain and Language* 101, 278–289.

Willems, Roel M., Aslı Özyürek & Peter Hagoort. 2007. When language meets action: The neural integration of gesture and speech. *Cerebral Cortex* 17, 2322–2333.

Wu, Ying C. & Seana Coulson. 2005. Meaningful gestures: Electrophysiological indices of iconic gesture comprehension. *Psychophysiology* 42, 654–667.

Xu, Jiang, Patrick J. Gannon, Karen Emmorey, Jason F. Smith & Allen R. Braun. 2009. Symbolic gestures and spoken language are processed by a common neural system. *Proceedings of the National Academy of Sciences of the United States of America* 106, 20664–20669.

Zatorre, Robert J., Alan C. Evans, Ernst Meyer & Albert Gjedde. 1992. Lateralization of phonetic and pitch discrimination in speech processing. *Science* 256, 846–849.

Maurizio Gentilucci, Elisa De Stefani, and Alessandro Innocenti
Università di Parma
Dipartimento di Neuroscienze
Rete Multidisciplinare Tecnologica
& Istituto Italiano di Tecnologia
Via Volturno 39
43100 Parma
Italy
maurizio.gentilucci@unipr.it, elidestefani@yahoo.it, and alessandro.innocenti@unimi.it

Influence of Language on Colour Perception: A Simulationist Explanation

Loïc P. Heurley, Audrey Milhau, Gabrielle Chesnoy-Servanin,
Laurent P. Ferrier, Thibaut Brouillet & Denis Brouillet

“How can perception be altered by language?” is the fundamental question of this article. Indeed, various studies have pointed out the influence of colour-related knowledge on object and colour perception, evoked by linguistic stimuli. Here the relevance of the simulationist approach is assumed in order to explain this influence, where the understanding of colour-related words or sentences involves a process of colour simulation that is supported by a neuronal network partially similar to the network involved in colour perception. Consequently, colour-related knowledge and colour perception can interact through a process of pattern interference. In support of this idea, studies are discussed showing priming effects between colour simulation and colour perception, but two limitations are also raised. Firstly, these works all used between-category colour discrimination tasks that allow the intervention of lexical processes that can also explain priming. Secondly, these works control the congruency link between prime and target at the level of ‘colour category’, and no demonstration is made of an influence at the level of specific hues. Consequently, the simulationist view of language/perception interactions seems an interesting way to thinking but more experimens are needed in order to overcome some limitations.

Keywords: colour perception; colour simulation; interaction; knowledge; language; priming

1. Introduction

For many years, perception was considered to be a process extracting information from the environment (Pylyshyn 1981; Fodor 1983). In parallel to this classic cognitivist approach, many researchers have developed the idea that perception is a process of construction of information that can integrate various influences (e.g., Neisser 1967; Stein & Meredith 1995). Firstly, bottom-up signals that come from all senses, for example judgments about taste and smell can be altered by colour perception (Spence *et al.* 2010). Secondly, perception may be constrained by the carrying out of actions (Gibson 1979). Thirdly, emotion conjured up by a stimulus or by mood can disrupt or facilitate the perception of

biolinguistics
© The Author (2012)
ISSN 1450–341
Biolinguistics 6.3–4: 354–382, 2012
http://www.biolinguistics.eu

emotional stimuli (Bocanegra & Zeelenberg 2009). Finally, a lot of studies have shown an influence of memory and knowledge on perception, mainly *via* language (e.g., label effect; see Pohl 2004 for a review).

The aim of this article is to suggest a possible explanation of language/ perception interaction based on simulationist approach of memory retrieval. We specifically focus our attention on the influence of colour-related words (e.g., banana, tomato) on colour perception because this case of language/perception interaction is strongly supported empirically. More specifically, we discuss two lines of research: (i) research on object recognition and (ii) on colour perception (section 2). In a third section, we present core ideas of sensorimotor models of memory including the simulationist conception of memory retrieval. Based on this approach, we develop the possibility that colour-related words can influence colour perception because each recruits common neuronal substrate. Indeed, in a sensorimotor model of memory colour simulation (i.e. a specific form of retrieval) would recruit sensory areas of the brain in order to represent sensory attributes of knowledge. In the fourth section, we discuss recent neuro-imaging evidences supporting this major assumption. However, based on these studies, we specified this basic idea. More precisely, we suggest that the neuronal substrate of colour simulation and colour perception would be best conceived as two partially overlapping neuronal networks that can interact only partially. This particular extension led us to expect two forms of priming: (i) from colour simulation to colour perception and (ii) from colour perception to colour simulation. In line with these predictions, we report a few studies demonstrating these possibilities (section 5). In a sixth section, we suggest that these studies have, at least, two limitations. We develop these limitations and try to propose possible experiments in order to overcome them in further research. Finally, in the last section (section 7), we sum up the potential simulationist explanation of language/perception interactions and develop two empirical ideas in order to best support this approach.

2. Influences of Colour Related Knowledge

Researchers working on knowledge organization and memory have been interested in ‘sensori-knowledge’ that we define as the knowledge of objects that are strongly associated with specific sensory attributes, such as their shape, their colour, their location, their sound, their movement, and so on (e.g., Pulvermüller 2003; Martin 2007; Barsalou 2008). In this flow of research, colour is a feature that has received much attention. This section develops studies suggesting the storage of colour in knowledge (section 2.1) and the influence of this knowledge (activated by processing linguistic stimuli) on perception (sections 2.2 and 2.3).

2.1. Colour Associations: A Stored Information

The memory of colour was first studied by pre-cognitivist researchers. For instance, Bruner & Postman (1949) studied colour associated with playing cards. Authors presented cards with congruent or incongruent colours: A red clover or a black clover, a red heart or a black heart. The participant's task was to judge the colour of the cards. Results indicated that in the incongruent condition, participants perceived a colour between red and black, suggesting an influence of the colour stored in memory (i.e. colour associated with the card) and thus the colour storage (see also Duncker 1939; Bruner & Postman 1949; Bruner *et al.* 1951; Harper 1953; Delk & Fillenbaum 1965). More recently, Reilhac & Jimenez (2006) brought to light the memory of colour by using black-and-white line drawings of objects strongly associated with yellow (e.g., banana) or green colours (e.g., artichoke). In this experiment, authors asked participants to categorize a drawing target presented subsequently to a drawing prime. Primes and targets either belonged to the same category or to a different one (i.e. animals or vegetables) and/or were of the same typical colour or of a different one (e.g., prime: *Frog* or *lettuce,* target: *Crocodile*). Only drawings with an important degree of colour agreement were used (estimated in a pilot study). Results showed priming when the drawings (i) belonged to the same category and (ii) more importantly when they were associated with the same typical colour. This last result suggests that colour association is stored in memory, more precisely in semantic memory. The study of Nijboer *et al.* (2007) focusing on scene recognition reinforces this conclusion. Generally, the results in the recognition of visual scenes show a facilitation effect for the recognition of natural scenes (e.g., beach *vs*. town) when they are presented in their usual colours; thus suggesting the part played by colour in the recognition of visual scenes (Oliva & Schyns 2000). Nijboer *et al.* (2007) found that this facilitation effect is absent in a subject suffering from a developmental colour agnosia. According to Nijboer *et al.* (2007), this participant would have a specific impairment of colour association stored at a semantic level. Consequently, the facilitation observed in healthy participants' results would be due to the retrieval of colour representations stored at a semantic level (absent in the agnosic patient).

To resume, colour seems to be a sensory feature of particular concepts or knowledge (i.e. that represent objects strongly associated with a colour) that can be available and automatically activated when participants have access to these concepts. In the next two subsections we discuss experiments reporting influence of these concepts on perception and we will concentrate on experiments using linguistic stimuli.

2.2. *Influence on Perception of Coloured Objects*

In the field of the perception of objects, various studies have indirectly tested the possible influence of colour on perception. Indeed, there are two theories that oppose each other concerning the part played by colour in the recognition of objects. On the one hand, the 'edge-based theory' assumes that recognition is mainly or only based on the shape of the object (Biederman 1987), whereas on the other hand the 'surface-plus-edge-based theory' defends that colour is an important cue for the recognition of objects. To settle between these two theories, different studies have measured the time taken to recognize objects presented in colour in comparison with objects presented in their greyscale version (i.e. without colour). The results obtained were recently synthesized by Bramão *et al.* (2011). Their review suggests that colour is an important cue only for the recognition of objects that are strongly associated with a particular colour (e.g., banana). These objects have been named 'high diagnostic colour' objects (i.e. HDC objects).

In one of their experiments, Tanaka & Presnell (1999, experiment 2) first presented two words each denoting an object, one appeared on the left-hand side of the screen and the other one on the right-hand side. In the second step, a drawing of an object was presented at the centre of the screen, in a coloured or greyscale version. Participants had to decide whether the object corresponded to the word presented earlier on the right or on the left-hand side (i.e. recognition task). Results showed that the decision time was faster when objects were HDC and appeared in colour compared with the conditions in which (i) the same objects were presented in greyscale and (ii) compared with the coloured or greyscale version of 'low diagnostic colour' objects (i.e. non HDC objects such as a chair). This result is important because it suggests that the recognition of objects involves information about colour and this supports the 'surface-plus-edge-based' theory. This result is also important because it suggests that reading words denoting objects strongly associated with a colour facilitates perception of congruent coloured objects.

Therriault *et al.* (2009) extended this previous research. In their second experiment, they asked participants to read a word that represented either an HDC object (e.g., pumpkin) or a non-HDC object (e.g., chair, car). Then participants had to judge whether the subsequently presented drawing represented the same object or a different one (i.e. verification task). The latter drawing appeared in a congruent colour (e.g., an orange pumpkin), in an incongruent colour (e.g., a blue pumpkin), or in greyscale (i.e. control condition). Results demonstrated that response times in the congruent condition were faster than in the greyscale condition, and the latter were faster than in the incongruent condition. In others words, the activation of colour when participants read words can either facilitate or disrupt the recognition of objects. In a third experiment, Therriault *et al.* (2009) obtained similar results when attention wasn't directly focused on colour.

In short, these various experiments and also others (see Bramão *et al.* 2011 for a review) suggest that colours stored in memory can facilitate or disrupt perception of objects presented in colour. Moreover, they demonstrate that this influence is produced by reading the linguistic stimuli that denotes colour-related objects, suggesting an interaction between memory, language, and object perception.

2.3. *Influence on Colour Perception*

The retrieval of colour-related objects influences not only the perception of objects but also the perception of colours. For instance, Mitterer & de Ruiter (2008) conducted an experiment in which participants had to discriminate the colour of an object. Participants saw three black-and-white line drawings: A carrot (i.e. a prototypically orange object), a banana (i.e. prototypically yellow), and a sock (i.e. without a particular colour association). These specific drawings appeared in various hues from a 'good orange' to a 'good yellow'. The results obtained with an ambiguous hue (i.e. between yellow and orange) were very informative. When a banana drawing was in this ambiguous hue, it was more often seen yellow, whereas when the drawing represented a carrot, the ambiguous hue was more often seen orange (compared with the sock condition). This result demonstrates convincingly that knowledge can influence colour perception.

Olkkonen *et al.* (2008) reach the same conclusion with a very different method. In their experiments, participants were presented with two types of stimuli: (i) fruit and vegetables with a typical colour (e.g., lemon, cucumber) and (ii) disks (i.e. control condition). Participants had to adjust the colour of the objects in order to make them appear in grey. Participants had four keys allowing them to add a bit of blue, yellow, green or red. Results showed that for fruit and vegetables that posses a typical colour, participants added more of the colour opposed to the typical colour than in the control condition (to perceive in grey). In practical terms, when participants perceived a yellow banana and that they had to colour it in grey, they added more blue (i.e. the opposite colour to yellow in the colorimetric model used) in comparison to a yellow disk that they had to colour in grey (see also Hansen *et al.* 2006; Nijboer *et al.* 2011; Witze *et al.* 2011 for similar results).

Another work conducted by Naor-Raz *et al.* (2003) also shows an influence of colour knowledge on colour perception but by using linguistic stimuli. In their second experiment, Naor-Raz *et al.* (2003) used a Stroop-like procedure in which participants had to name the colour in which the words were printed. Words used represented objects strongly associated with a colour (e.g., banana) and could appear written in the congruent colour (e.g., banana written in yellow) or in an incongruent colour (e.g., banana written in purple). Results show that colour-naming times were faster for the incongruent conditions compared with the congruent ones. Despite a slowdown of response times opposed to previous data, these results also demonstrated the possibility that colour-related knowledge can influence colour perception (see also Klein 1964; Yee *et al.* 2012).

Consequently, studies presented here give evidence that retrieval of knowledge can influence colour perception and not only object perception. We report studies using pictures but also words that suggest, once again, that this influence may be mediated by reading linguistic stimuli (also Kubat *et al.* 2009).

The aim of this first section was to develop studies indicating an influence of colour-related knowledge on colour perception. To sum up, results reported here suggest that the retrieval of colour-related knowledge while words are processed can influence both recognition of coloured objects and perception of colours. More precisely, colour or colour association seems to be a component of some knowledge able to interact with the processing of perceived colour. This influence can be found when participants have to implicitly process colour such as for objects or scene recognition or explicitly when participants have to discriminate colour of target stimuli. Nevertheless, the underlying process that allows this interaction remains largely unexplored. One possible candidate lies in the retrieval process assumed by the sensorimotor models of memory. In the next section, we develop the core ideas of these models as well as the specific retrieval process that it assumes.

3. Simulationist Approach of Cognition and Memory

Experiments presented above show an influence of colour-related knowledge elicited by language on perception (both object and colour perception); but why such an influence? Here, we will develop on the idea that a possible answer can be found in a recent approach of cognition that assumes both a distributed memory across the entire brain and a particular form of retrieval that can be called 'simulation'. More precisely, it is possible to postulate that the interaction may be explained by the nature of the format in which colour knowledge is stored and retrieved from memory. In this section, we will first develop the cores ideas of the sensorimotor models of memory. Then, we explain the specificity of the simulation process by which knowledge is retrieved. Accordingly, we develop a simulationist explanation of memory/perception interaction based on the use of the same neuronal substrate.

3.1. Sensorimotor Models of Memory

The organization and format of knowledge that the sensorimotor models stand for is a key specificity of the models in comparison to other memory models developed until now. We will mainly develop on the originator model of Allport (1985); all the other sensorimotor models of memory are strongly similar to this one (e.g., Schacter *et al.* 1998; Martin & Chao 2001; Pulvermüller 2001; Barsalou *et al.* 2003; Versace *et al.* 2009). Allport (1985) explains that there are two ways of conceiving the implementation of knowledge (or concepts) in the brain. The first considers that knowledge corresponds to a precise physical unit (e.g., neuron or a group of neurons). Thus, a physical unit would be dedicated to the implementation of only one concept. The second considers that knowledge corresponds to a specific pattern of activity of a group of neurons. Thus, a same group of neurons can support different concepts as long as different patterns of activity are involved (i.e. distributed knowledge). On the basis of the work of Hebb (1949), he explains that such patterns would form and reinforce themselves as long as co-occurrent activations (pre- and post-synaptic) occur between neurons. More precisely, a neuronal pattern would constitute itself step by step *via* the alteration of synaptic weights (i.e. the correlation learning rule; Hebb 1949). This leads Allport to talk about 'auto-associated patterns' to support the idea that a pattern is a stable unit with specific properties. Nevertheless, until then Allport (1985) only details a model of knowledge already developed by the connectionists (e.g., Hinton 1981, Farah & McClelland 1991, McClelland & Rumelhart 1985).

Allport's main contribution that will characterize the sensorimotor models of memory is the way he considers the implementation of knowledge at the level of the brain. According to him, the auto-associated patterns representing objects in memory would be composed of auto-associated sub-patterns distributed across all the sensory and motor domains. He takes for example a 'telephone' and explains that the pattern representing this object is distributed at a visual, tactile, auditory and motor level coded in the form of visual (e.g., format, size), tactile (e.g., texture), auditory (e.g., ringing) and motor (i.e. motor routine carried out with the object: to pick it up, to hold it, etc.) attributes. He also explains that the attributes associated with the verbal interaction regarding the telephone, such as the phonological, spelling and written attributes of the referent word can be respectively stored at an auditory, visual and motor level. Thus, the encoding process is above all an associative process implying that the different sensory and motor attributes, supported by brain structures all across the brain, have to constitute themselves in a unit (i.e. binding). Moreover, Allport (1985) explains that the auto-associated sub-patterns can also be embedded in other patterns as long as they represent the objects sharing similar attributes. For example, the auto-associated sub-pattern of the ‘yellow’ attribute can be part of the pattern representing not only a ‘banana’ but also a ‘lemon’ or a ‘chick’, etc. In the wake of Allport (1985), different authors have added various ideas in order to precise this model. First, Pulvermüller (2003) adds the idea of functional units of the patterns. According to this author, the different areas composing the network would work as a unit. The functioning of this unit would depend on the integrity of each part of the unit. If one part of the cortex happens to be damaged, the networks based on this area would consequently be more a less affected. Secondly, Barsalou (2009) suggests that the neuronal patterns representing knowledge would not only be composed of sensory and motor attributes but also of emotional attributes.

3.2. *Simulation Process of Retrieval*

Given the way with which knowledge is encoded and organized, the different sensorimotor models of memory agree on the nature of the retrieval process. These models assume that retrieval implies reactivation of neuronal networks representing knowledge, a process named ‘simulation’ (Barsalou 1999). Simulation has at least two main characteristics. Firstly, partial activation of a knowledge tends to reactivate all aspects of this knowledge, including all sensory attributes. This feature is due to the ‘process of pattern completion’ whereby an input that stimulated only a small number of neurons, composing a pattern, tends to reactivate the entire pattern. For instance, the reading of a word (e.g., banana) only involves the activation of the orthographic and phonological attributes of this knowledge. However, through the ‘pattern completion process’ all components of the network are rapidly active including the associated colour yellow (Pulvermüller 2003). This process occurs because every part of the pattern is strongly connected with each other through a Hebbian principle of reinforcement (Hebb 1949). Barsalou *et al.* (2003) underlined the fact that a simulation is never complete and involves bias compared with the original sensory experience. Secondly, and more importantly, activation of various sensory attributes of knowledge seems to involve the use of common neuronal structures with on-line perception, more precisely structures involved in low-level perceptual processing. This specificity is due to the fact that sensory attributes would be stored in areas of the brain also involved in perception. In the case of the word ‘banana’, the knowledge retrieved involves a specific network distributed in part on the neuronal substrate involved in colour perception. Accordingly, it is possible to assume that perception of colour and language can interact because language processing involves a simulation process in order to reactivate colour-related knowledge. The use of such a process involves the use of common neuronal structures or the sharing of the same processes or same resources by language processing and perception. In other words, the format in which colour information is stored (connection between neurons of the visual area) and retrieved (by reactivation of the visual area) implies its automatic interaction with colour perception that uses the same visual areas.

To conclude, remember that the purpose of this section was to develop a process that would explain the influence of colour-related words on colour perception. We suggest that one possible candidate is the specific retrieval process assumed in sensorimotor models of memory usually called 'simulation'. Indeed, sensorimotor models of memory explain that memory would be a property of the entire brain including sensorimotor areas and retrieval of knowledge would occur *via* reactivation of neuronal patterns distributed over the brain. In the case of colour-related words, reading would involve activation of the entire patterns representing referents of words including in the visual area (specifically the colour perception area) supporting the colour attributes of knowledge. The recruitment of the same neuronal structures in order to simulate and perceive colour involves a meeting point for these two functions that can explain how colour-related knowledge can influence low-level colour perception. In the next section and in support of this simulationist explanation, we discuss studies showing the involvement of the same neuronal structures for colour simulation and colour perception (i.e. gyri lingual and fusiform). Moreover, based on these works we also specify more this possible explanation.

4. Colour Perception and Colour Simulation: Similar Neuronal Patterns

In this section, we will first review neuro-imaging studies having looked for the 'colour perception centre' within the visual cortex. Then, we will review neuro-imaging studies that have tried to point out the use of the same neuronal structures to retrieve colour-related knowledge. In a third sub-section, we discuss the implications of a neuropsychological dissociation in which participants have preserved mental colour imagery whereas colour perception is impaired (i.e. achromatopsia). We finally explain that colour perception and colour simulation seem to recruit partially overlapping neuronal networks.

4.1. Cortical Structures Involved in Colour Perception

Various experiments trying to situate the 'cortical colour centre' and conducted on healthy participants suggest different results. First, Zeki *et al.* (1991) situated this centre in the lingual gyrus in the occipital cortex. Its activation was found when participants passively viewed basic coloured stimuli (i.e. Mondrian). Authors have named this structure the hV4 area, human equivalent of the monkey V4 area (Zeki 1983a, 1983b, 1983c). Other studies have found a more extensive activation around the lingual gyrus when participants had to actively discriminate the colour of the stimuli, which suggests that there is a part played by the task applied to coloured stimuli (Corbetta *et al.* 1991; Guylás & Roland

1994). In order to clarify the structures involved in colour perception taking into account the potential role of tasks, Beauchamp *et al.* (1999) carried out an experiment where they compared activations in function of the task. They adapted the Farnsworth-Munsell Test (Farnsworth 1957) usually used to detect achromats (i.e. people with a specific loss of colour vision). The procedure used involved perception of various hues of the same colour presented in a linear gradient from light hue to dark hue or presented in a disorderly manner. Two tasks were used: (i) A passive viewing condition of stimuli and (ii) an active condition where participants had to discriminate ordered stimuli and disorderly ones. When participants passively viewed coloured stimuli, results showed the same activation that Zeki *et al.* (1991) found and also an extended activation in the anterior portion of the fusiform gyrus. When participants carried out a judgment on the stimuli, three major results are founded. Firstly, ventro-occipito-temporal activations both in lingual and fusiform gyri were observed as in the passive condition. Secondly, the activation of the fusiform gyrus is more important. Thirdly, they also observed distributed activations over the cortex, such as activations of V1 and V2 areas (i.e. within the visual cortex) and activations of the left frontal cortex, dorso-lateral part of the occipital cortex and superior parietal lobe in a majority of participants (see also Engel *et al.* 1997 for a similar result; see Gegenfurtner & Kiper 2003 for a review).

Based on these results, Beauchamp *et al.* (1999) explained that the carrying out of a task involving judgment about colour involves the activation of more colour selective regions compared with the passive condition. Consequently, they assumed a 'decentralized view' of colour perception in which there is no specific colour centre. The colour perception results from the activation of a complex distributed network across the cortex mainly concentrated around the lingual and fusiform gyri. Results of Corbetta *et al.* (1991) and Guylás et Roland (1994) that used a discrimination task as well as results of Zeki & Marini (1998) that used coloured photos of objects support this view. Indeed, Zeki & Marini (1998) made an experiment with complex colour pictures of objects and scenes rather than abstract coloured stimuli (i.e. Mondrian). In this experiment, results show that perception of these more complex stimuli also involves an activation of the lingual and fusiform gyri compared with the perception of the same objects in greyscale (see also Bramão *et al.* 2010 for similar results). Moreover, this 'decentralized view' of colour perception is in accordance with results showing that a lesion of the V4 area in a monkey involves only a small disturbance of colour perception suggesting the implication of other structures (Gegenfurtner 2003). This view is also in accordance with complementary results showing the impairment of many areas in 'cerebral achromates', and not only of the fusiform and lingual gyri (Bouvier & Engel 2006).

In short, the 'cortical colour centre' should rather be seen as a distributed network, concentrated around the lingual and fusiform gyri, that can be differentially activated according to the perceptual task and the coloured stimuli used.

4.2. Cortical Structures Involved in Colour Simulation

Various experiments have sought to bring to light that retrieving coloured-related knowledge or low-level colour perception call on to the same structures. The first experiments were conducted by Martin *et al.* (1995). They compared brain activations in two conditions; participants either had to name colour-related objects represented by black-and-white line drawings (e.g., banana) or produce the colour associated with the objects (i.e. yellow). The underlying idea is that a production task would involve the retrieving of colour information while the naming task would not. Results showed an activation of the fusiform gyrus in the production task compared with the naming task. Moreover, they found the same results in a second experiment where drawings were replaced by words denoting the same objects. These results are congruent with neuro-imaging studies on colour perception that have showed the main role of the fusiform gyrus in colour perception (e.g., Beauchamp *et al.* 1999). Wiggs *et al.* (1999) carried out a similar experiment and obtained the same results. Moreover, they showed that the naming task also involved the activation of the fusiform gyrus compared with the control condition where participants viewed a visual noise. This last result suggests that simply naming an object strongly associated with a colour seems to involve the colour perception network. In another study, Chao & Martin (1999) directly compared, for the same participants, activations when (i) they perceived abstract coloured stimuli (i.e. Mondrian), (ii) named black-and-white line drawings, and (iii) produced the associated colour. Results showed that perception and production did not activate the same neuronal structure, but very close structures 2 centimetres (cm) away from one another. Based on this result, Chao & Martin (1999) concluded that perception of colour and retrieval of coloured-related knowledge do not require the same neuronal structure. Nevertheless, Martin (2009) explains that they failed to find the use of exactly the same neuronal structure in this previous study, simply because they used a passive perception task: a passive-viewing Mondrian task which activated only the lingual gyrus (Zeki *et al.* 1991). It is possible that by using more complex stimuli or a more active-viewing condition (Beauchamp *et al.* 1999), perception would be more complex and an activation of the fusiform gyrus would be found in both tasks. Simmons *et al.* (2007) directly tested this hypothesis. In their experiment, authors compared, for the same participants, activations that occurred when they (i) carried out the same 'colour-perception-task' used by Beauchamp *et al.* (1999) and (ii) carried out a retrieval task where participants had to judge the veracity of sentences involving a concept and a colour (e.g., a banana is yellow?). The results showed an activation of a common neuronal substrate in the two tasks: The left fusiform gyrus. Moreover, activations that occurred in the retrieval task were consistent with activations found in previous studies (Chao & Martin 1999; Martin *et al.* 1995; Wiggs *et al.* 1999). This study shows that the retrieval of colour-related knowledge involves the recruitment of the same neuronal structures that are used to perceive colour and more precisely involved in low-level colour perception (i.e. that does not involve object identification or object recognition). Nevertheless, these structures only overlap when the task and the stimuli are relatively complex.

Moreover, consistent with the second experiment of Martin *et al.* (1995), the work of Simmons *et al.* (2007) suggests an activation of the fusiform gyrus when participants process linguistic inputs (i.e. short sentences, see also Goldberg *et al.*

2006 for a similar procedure and results). This idea of an activation of the fusiform gyrus when participants read colour-related words is also reinforced by studies conducted by Kellenbach *et al.* (2001) and Oliver & Thompson-Schill (2003). In their experiments, authors used the same procedure: They asked participants to judge if a word was associated or not with a colour (e.g., banana *vs*. snow). Black, white, and grey hues were defined as non-colours. Results showed an activation of part of the fusiform gyrus close to the active part found in the previous studies (≈1cm).

Recently two other studies extended these results. Firstly, Hsu *et al.* (2011) made an experiment in order to understand why the lingual gyrus activated in colour perception is not found to be activated in retrieval of colour-related knowledge. In order to best understand this absence of activation, Hsu *et al.* (2011) manipulated two variables : (i) level of detail of the colour simulation and (ii) the cognitive style of participants. In their experiment, participants had to carry out a judgment on luminance in the same way as in the work of Simmons *et al.* (2007) and also carry out a conceptual task involving colour simulation. More precisely, participants first saw two words representing two colour-related objects. Then, a third word would appear on the screen and participants had to select which of the two words represented the object associated with the closest colour. In one bloc, the three words belonged to the same colour category and similarities occurred at a specific hue level (e.g., 'SCHOOL BUS' compared with 'BUTTER' and 'EGG YOLK'). Whereas in another bloc, the three words belonged to different colour categories (e.g., 'BEETLE' compared with 'PAPRIKA' or 'BANANA'). The first bloc involved a colour simulation level deeper than the second bloc. Moreover, participants were divided in two groups based on their score to the *VVQ* (i.e. *Verbal and Visual Questionnary*; Kirby *et al.* 1988). In one group, participants had a visual cognitive style whereas in the other, participants had a verbal cognitive style. Results show an activation of the left lingual and fusiform gyri during the conceptual task. The activation of the fusiform gyrus is more important when the level of simulation is deeper and the activation of the lingual gyrus is more important for participants with a visual cognitive style. Moreover, Hsu *et al.* (2011) observed an overlap of structures activated in the colour perception task and in the conceptual task both at the level of the lingual and fusiform gyri.

In other recent work, Hsu *et al.* (2012) extended their previous results. In this experiment, Hsu *et al.* (2012) compared the activation found when participants have to compare colours associated with two words. In one condition, words denote colour objects with the same or a different colour (e.g., 'BANANA' *vs*. 'BUTTER'). In a second condition, words denote achromatic objects identical or different on the level of grey (e.g., 'SNOW' *vs*. 'COAL'). The goal is to compare the activation of the cortex in the 'conceptual colour condition' with the 'conceptual achromatic condition' in order to know which cortex areas are specifically activated when a 'chromatic condition' is compared with an 'achromatic condition'. Indeed, comparing a 'chromatic condition' with an 'achromatic condition' is the usual way to determine which cortex areas are specifically activated in colour perception (i.e. chromaticity effect; e.g., Beauchamp *et al.* 1999; Chao & Martin 1999; Simmons *et al.* 2007). In previous experiments, activation in the conceptual task is obtained by comparing retrieval of colour-related knowledge with retrieval of non colour-related knowledge, like action concepts (e.g., Simmons *et al.* 2007) and not 'achromatic conditions'. The major result is the overlap of colour perception and colour simulation at the lingual gyrus level.

In short, all these studies converge on the evidence that retrieval of colour-related knowledge, through the processing of linguistic stimuli, involves the activation of the same neuronal structures: the lingual and fusiform gyri also used when participants perceive colours or coloured objects (Zeki & Marini 1998, Beauchamp *et al.* 1999). Nevertheless, the implication of the same neuronal structures seems to depend on many variables such as the type of stimuli, cognitive style, type of conditions compared in order to define the neuronal structures activated. Moreover, it is important to note that colour conceptual tasks, as used in previous presented experiments, leave the possibility that participants used mental imagery of colour in order to complete the tasks. For instance, the conceptual task used by Hsu *et al.* (2012) is also used by researchers wanting to measure mental imagery of colour (e.g., De Vreese 1991; Bartolomeo *et al.* 1997; van Zandvoort *et al.* 2007). Consequently, it seems a risk to extend this similarity in brain structures to all tasks involving an access to colour-related knowledge, especially given the current state of knowledge on the question.

4.3. *Achromatopsia with Preserved Colour Mental Imagery?*

If the same structures are used to perceive and retrieve colour, it would not be possible to find people with impairment in colour perception (i.e. achromatopsia) and a preserved ability to retrieve colour (e.g., colour mental imagery). However, Shuren *et al.* (1996) and Bartolomeo *et al.* (1997) report two cases of patients who were impaired on tasks involving colour perception: the Ischihara test (Ishihara 1974) and the Farnsworth-Munsell test (Farnsworth 1957). However, these patients accurately carried out tasks involving the imagination of colours: naming the colour of objects from memory, mentally comparing hues (De Vreese 1991), and giving the most objects as possible associated with a given colour. This dissociation is an issue. Nevertheless, as Simmons *et al.* (2007) explain, the neuropsychological evidence of the dissociation between the access to colour-related knowledge and the perception of colour does not necessarily imply that these two abilities are completely independent from one another. It is possible that the perception of colour and the access to the knowledge related to colour involve brain structures only partly identical.

4.4. *Overlap of Neural Networks*

Studies on colour perception suggest that a neuronal substrate must be conceived as a neuronal network distributed across the brain (i.e. 'decentralized view'; cf. Bouvier & Engel 1997; Beauchamp *et al.* 1999). More precisely, this network seems to be mainly concentrated around the lingual and fusiform gyri, but its distribution is relative to various factors, such as a perceptual task and the type of stimuli perceived (Zeki & Marini 1998; Beauchamp *et al.* 1999; Bramão *et al.* 2010). On the other hand, studies focusing on the neuronal bases of the retrieval of colour-related knowledge, in other words 'colour simulation', point out the involvement of the same neuronal structures. Nevertheless, it is also possible to assume a 'decentralized view' in this case. Indeed, colour simulation seems to be supported by a distributed neuronal network also concentrated around the lingual and fusiform gyri (Martin *et al.* 1995; Chao & Martin 1999; Wiggs *et al.* 1999; Kellenbach *et al.* 2001; Oliver & Thompson-Schill 2003; Goldberg *et al.* 2006; Simmons *et al.* 2007; Hsu *et al.* 2011, 2012), but the distribution of the network changes depending on various factors such as a conceptual task, the type of stimuli, the type of cognitive style, etc. Moreover, results agree with the possibility that these two networks partially overlap each other at the level of the lingual and fusiform gyri because these structures are active in conceptual and perceptual tasks. This overlap also depends on the combination of various factors affecting independently colour perception and colour simulation (i.e. type of coloured stimuli, type of perceptual and conceptual task, etc.). Note that this distributed approach is more consistent with studies demonstrating a dissociation between central colour blindness (i.e. achromatopsia) and colour imagery (Shuren *et al.* 1996; Bartolomeo *et al.* 1997). Indeed, this dissociation implies that it seems exaggerated to consider the perception and simulation of colour as fully supported by well-defined structures, such as the lingual and fusiform gyri, which once damaged must affect both functions simultaneously. In contrast, a design in terms of overlap of distributed networks allows the possibility of relatively independent processing.

Consequently, this conception helps explain the influence of colour simulation on colour perception through a process of 'neuronal pattern interference'. Indeed, colour simulation can affect colour perception because colour simulation recruits a neuronal network similar, in particular points, to the network used when colours are perceived. The idea of pattern interference can already be found in the work of Masson (1995) who tries to explain the semantic priming from a connectionist point of view. The idea of Masson (1995) is that we observe semantic priming because the neuronal pattern involved in the representation of the prime can be, at least, partially similar to the neuronal pattern representing the target. Accordingly, priming emerges from the pre-activation of parts of the neuronal network representing the target that are common with the neuronal network representing the prime. The same idea can be assumed for the case of the influence of colour simulation on colour perception. Indeed, colour simulation induces a pre-activation of certain parts of the neuronal network (i.e. lingual and fusiform gyri) that will be re-used for perception.

In others words, retrieval of colour-related knowledge involves the activation of parts of a neuronal network also involved in low-level colour perception. More importantly, this approach of interference between retrieval of colour-related knowledge and colour perception allows specific predictions about this interaction. Firstly, it is possible to expect a form of priming of colour perception whereas the prime involves only colour simulation. Indeed, until now all the works that show an influence of knowledge on colour perception use procedures in which stimuli simultaneously involve access to knowledge and perception of colour (e.g., coloured drawings and coloured words). These two activities are not separated in time. But the idea that this interference arises from the overlap of network parts also predicts the possibility of interference even when the simulation is not performed at the same time as perception. Consequently, colour simulation can affect colour perception in a priming procedure. Moreover, the direction of influence is not restricted, and colour perception could also influence colour simulation. In the next section, we report experiments that support the possibility to observe an influence even when simulation and perception of colour are supported by temporally separate stimuli (i.e. a form of perceptual priming). Moreover, we also report evidence showing reversed priming because the direction of the influence is not restricted.

5. Bidirectional Priming between Colour Simulation and Colour Perception

The simulationist view of cognition emphasizes that colour-related knowledge can influence colour perception because the retrieval of this type of knowledge involves a simulation process using a neuronal network partially similar to the network involved in colour perception. This conception leads to the idea that colour simulation can affect colour perception even if these two processes are temporally separated. In this fifth section we will subsequently report two lines of evidence supporting this possibility. Firstly, we discuss results showing the possibility to prime colour perception *via* colour simulation and in a second subsection we discuss works showing the possibility to prime colour simulation *via* colour perception.

5.1. Priming of Colour Perception by Colour Simulation

The first interesting work showing a priming of colour perception *via* colour simulation is the work of Richter & Zwaan (2009). Authors conducted an experiment where participants saw a coloured square, immediately followed by a word and followed itself by a second coloured square, in each trial. Participants had to perform two different tasks successively: They (i) had to carry out a lexical decision on the word and (ii) decide if the second square appeared in the same or in a different colour as the first one. In target trials, words were nouns that denoted a colour (e.g., yellow, blue) that corresponded or not to the colour of the last square. Results showed that response times were shorter when the colour of the word was congruent with the colour of the last square in comparison with the condition where the colours were different. This result suggests that colour perception involving low-level processes can be influenced by linguistic stimuli even when simulation is carried out just before colour perception. Nevertheless, this study has a limitation. Indeed, in order to compare the two coloured targets, participants have to store, in short-term memory, the colour of the first square in order to compare it with the second square. Consequently, the influence of words can be situated at a perceptual level (i.e. affects the perception of the second square) or at a memory level (i.e. affects the remembering of the first square). This second possibility is supported by various works showing an influence of colour-related knowledge on the remembering of colour (Duncker 1939; Loftus 1977).

In order to overcome this limitation, we have recently conducted experiments using a perceptual task that precludes the conservation of colour in memory. Moreover, we used words denoting objects associated with a colour and not directly a colour label in order to extend the results of Richter & Zwaan (2009). In the first experiment (Heurley *et al.* 2012), participants saw words presented very briefly and immediately followed by a green or yellow circle. Participants had to discriminate only the colour of the circle (i.e. chromatic discrimination task). Prime words denoted natural things strongly associated with a green or yellow colour (e.g., chick, artichoke). This strong association was controlled in a pilot study where participants had to give the first colour that came to mind when they read colour-related words. Only words with a colour consensus higher than 70% were used. Results showed facilitation when the colour elicited by the words matched the perceived colour compared with the condition where the colours were different. We replicated this facilitation in a second experiment where the participant's task did not explicitly focus attention on colour or involve a colour label answer. Indeed, lexical priming of colour name could explain the facilitation effect in this experiment. In other words, perception of the word 'banana' elicits spontaneously the word 'yellow' facilitating the 'yellow response'. In the second experiment (Heurley *et al.* in press), participants saw two circles in the target's display. These two circles were presented in the same colour or in a different one. In the case where the circles appeared in the same colour, they could either appear in yellow or in green. Participants had to detect whether target circles were presented in the same colour or in different colours. When circles appeared in the same colour, results showed facilitation when the colour of the two circles was congruent with the colour elicited by the prime (e.g., two yellow circles preceded by the word 'BANANA') compared with the condition in which the colour elicited by the prime was different (e.g., 'ARTICHOKE'). In this last experiment, the effect observed can not be explained in terms of lexical priming because response labels were not colour labels. Moreover, the use of a mental imagery strategy, that could explain our results, seems precluded because words were only presented during 150 ms (with a 350 ms SOA) and participants only had to read the words as fast as possible.

Taken together, these studies support the possibility of obtaining perceptual colour priming whereas priming occurs *via* colour simulation. In these experiments, primes involve a colour simulation activating a neuronal network composed of parts that are also used by the neuronal network used in order to perform the perceptual task on the target. Moreover, primes are always words (colour words or words denoting colour-related objects) showing that this effect can be mediated by language comprehension as is assumed by various researchers (Pulvermüller 2001, 2003; Barsalou *et al.* 2003; Zwaan & Kaschak 2009).

5.2. *Priming of Colour Simulation by Colour Perception*

If language influences colour perception through the use of partial overlapping of close neuronal networks, it is also plausible to expect a reverse influence: of perception on language. Indeed, colour perception may pre-activate part of the network that could be used to process linguistic stimuli. The first evidence of this opposite influence is found in a study by Richter & Zwaan (2009) already reported. In their study, participants first saw a coloured square followed by a word denoting the same or a different colour (or a non word), followed itself by a second coloured square (in the same colour as the first square or in a different one). Results showed that lexical decision times for words were influenced by the colour of the first square: response times were faster when the same colour was perceived and simulated compared with the condition where they were different. Thus, Richter & Zwaan (2009) showed an influence of colour perception on language processing. This result seems to suggest that language processing can be influenced by perception. However, in this work, words used denoted directly a colour, thus this can give place to another explanation in lexical terms. For instance, perception of a colour can automatically activate the associated word that facilitates subsequent language processing of a congruent word (i.e. colour labels). In consequence, the facilitation observed can be explained by lexical priming.

Nijboer *et al.* (2006) conducted an experiment where the same reverse influence was found. But in this experiment, authors used colour-related words (e.g., tomato) thus this prevents an explanation in terms of lexical priming. In this experiment, participants had to decide whether a target word was an existing word or a pseudo-word (i.e. lexical decision task). In each trial, participants saw a coloured prime followed by a target word or a pseudo-word. Fifty six colour-related words were mixed with non-related colour words (i.e. 168) and pseudo-words. Primes used could occur in seven prototypical colours (i.e. red, blue, green, yellow, pink, orange, and brown). In target trials, a colour-related word could be preceded by a congruent colour patch (e.g., red patch – TOMATO) or an incongruent colour patch (bleu patch – TOMATO). Results showed that lexical decision times for colour-related words were shorter in the congruent condition than in the incongruent condition. Moreover, Nijboer *et al.* (2006) added a control condition where the colour patch was multicoloured. Lexical decision times in this condition differed from the congruent condition but not from the incongruent condition. These results strongly support the possibility that language processing can be facilitated by an earlier low-level colour perception.

Recently, Gebuis *et al.* (2009) investigated bidirectional influence between perception and language in colour-grapheme synesthetes. In colour-grapheme synesthesia, letter or number leads to a vivid and robust experience of a colour. Accordingly, this effect is interesting for our purpose because this effect can be seen as a particular case of language/perception influence. Gebuis *et al.* (2009) explained that it is well established that in participants with colour-grapheme synesthesia, colour perception can be affected by the perception of numbers. For instance, naming the colour of the ink in which a number is written can be facilitated if the colour associated with the number in the memory of the synesthetes is congruent. However, if the colour associated with the number is different, a slowdown is observed: this is a type of Stroop effect. Accordingly, the access to colour information when numbers are processed seems to be strongly supported. The reverse access: the access to number information when colours are processed also seems possible. In their experiment, Gebuis *et al.* (2009) try to demonstrate this bidirectional link between number and grapheme in synesthesia. Participants saw a number coloured in the colour associated with this number in their memory (congruent condition) or coloured in a different colour (incongruent condition). Then participants had to do two different tasks. In the first one, they had to name aloud the number in order to investigate the influence of colour on number processes. In the second task, they had to name aloud the colour of the ink in order to investigate the influence of number on colour processes. Gebuis *et al.* (2009) observed a priming effect in both tasks showing the bidirectional influence between number and colour in synesthetes. More precisely, perception of a number involved access to colour information (i.e. that can be named in the context of our article: colour simulation) that influenced colour perception, and perception of colour involved access to numerical information that influenced number processing. This last influence suggests that colour perception could automatically reactivate the link in memory between the colour and the number (as a particular colour/object association) influencing number processing.

To sum up, in this fifth section, we reported works suggesting the possible priming of colour simulation on colour perception as well as the reverse priming. These works support the idea that colour-related knowledge and colour perception interact in a bidirectional way *via* the reciprocal influence of underlying neuronal networks (i.e. pattern interference process). Indeed, as predicted by this approach it is possible to find an influence even when colour simulation and colour perception are performed separately in time. More precisely, it is possible to observe a form of colour priming when prime only involves a simulation of colour, not real perception and it is also possible to find semantic priming when prime involves real perception, not semantic processing. However, in the next section, we develop two limitations of these various studies that challenge the simulationist approach.

6. Two Problems for the Simulationist Approach

In the previous sections, we assumed the possibility that colour-related knowledge could influence colour perception through a colour simulation process. Studies supporting this hypothesis show an influence of colour-related knowledge on colour perception in a priming paradigm (Heurley *et al.* 2012, in press, Nijboer *et al.* 2006, Richter & Zwaan 2009). However, we suggest that these works have two limitations. First, recent studies suggest the possibility that specific colour perception tasks can not only involve perceptual processes but also lexical processes. Accordingly, it is possible that the influence of colour-related knowledge on colour perception occurs at a lexical level and not at a perceptual level in previously reported experiments. Secondly, various experiments, discussed in the previous sections, manipulate the congruence link between knowledge and coloured target at a 'colour category' level and not at the level of specific hues. However, it is possible that colour simulation involves simulation of specific hues (e.g., dark green, light green) rather than the general 'colour category' (e.g., green). We develop these two problems in the two following sub-sections and based on the literature, we try to put forward ideas in order to overcome these limitations.

6.1. Lexical Influence?

Studies on categorical perception of colour have shown that response times to discriminate two hues belonging to two different colour categories (e.g., green hue *vs.* yellow hue) are shorter than response times to discriminate two hues belonging to the same colour category (e.g., dark green *vs.* light green; see Goldstone & Hendrickson 2009 for a review). Bornstein & Korda (1984) have developed the idea that this difference could be due to the intervention of two cues in order to discriminate two hues from two different categories (i.e. between-category discrimination): a perceptive cue and a lexical cue. However, only the perceptive cue would be used to discriminate two hues from the same category (i.e. within-category discrimination). This idea is similar to the 'Whorf hypothesis' according to which colour perception could be influenced by language (i.e. colour labels; see Kay & Kempton 1984).

Recently, a number of works have confirmed this hypothesis. For example, Roberson and Davidoff (2000) carried out an experiment where every trial started with the brief appearance of a coloured target on the screen. This was followed by a 5 second delay. Then two colours appeared. Participants had to choose which of the two colours corresponded to the first coloured target (i.e. 'two Alternative-Forced Choice' task). The two coloured choices either belonged to the same category (i.e. within-category discrimination) or to two different categories (i.e. between-category discrimination). During the five second delay, participants either (i) faced a white screen (i.e. control condition), either (ii) had to follow a curbed line with the eyes (i.e. visual interference), either (iii) read a list of words (i.e. verbal interference). Results show a classic effect in the control condition: participants are faster to discriminate the target colour when the two colours belong to two different categories. This effect persists even with a visual interference. However, the effect disappears when the interference is of a verbal nature. Roberson & Davidoff (2000) explain that the verbal task disrupts the retention of the label of the colour of the target. This disruption cancels the lexical cue supposed to create an advantage in the between-category discrimination task, which induces a disappearance effect. This result thus confirms the involvement of a lexical cue in a between-category discrimination task (see Pilling *et al.* 2003 for identical results).

Gilbert *et al.* (2006) carried out another experiment confirming this conclusion. Participants had to discriminate two targets which appeared simultaneously on the screen. At each trial, twelve coloured squares were presented in a circle. One square among the twelve was presented in a different hue. The participants' task was to detect as fast as possible if this different square had appeared on the left or right side of the circle. The hue of the target square could either belong to the same category as the other squares or to a different category. Gilbert *et al.* (2006) observe a 'categorical perception effect' only when the targets were presented on the right side of the circle. Authors explain these results by the fact that the right part of the circle is situated in the right visual field which is linked to the left hemisphere dedicated to language processing. Consequently, this result would confirm the involvement of linguistic processing during colour perception, at least during a between category discrimination task (see also Drivonikou *et al.* 2007).

Recently, Siok *et al.* (2009) carried out an experiment in which they directly focused on the cortical structures activated during a between- and a within-category discrimination task. For this, Siok *et al.* (2009) used the same procedure as Gilbert *et al.* (2006); however, this time participants' cortical activity was recorded by fMRI. The results highlight an activation of the areas involved in the processing of language when participants have to discriminate a square target coloured in a hue belonging to a category different to the other squares (i.e. between-category discrimination task). This last study reinforces, in the same way as the two previous ones, Bornstein & Korda (1984)'s hypothesis according to which between-category discrimination involves two cues: a perceptive and a lexical cue.

Taken together, these results confirm the possibility that colour perception involves lexical processes, at least in a between-category discrimination task. Previous results demonstrating that colour-related knowledge influences colour perception all used between-category discrimination tasks. For instance, in one of our experiments (Heurley *et al.* 2012), the perceptual targets were composed of two coloured circles that appeared in the same colour (i.e. they belonged to the same colour category) or in a different colour (i.e. they belonged to two different colour categories). In the work of Nijboer *et al.* (2006), various primes were used that belonged to different categories, the authors never used different hues of a same colour category. Richter & Zwaan (2008) tried to control the link between the two different colour targets that participants had to compare in order to obtain colours that belonged to the same lexical category. They did a pre-test where participants had to name the colour of various targets used. It seems that, at least for various couples of colours used, a colour would imply different colour labels. Accordingly, it is possible that in these various experiments, the colour perception task used implied lexical processes. This is an issue because lexical cues used in order to perceive colour could be influenced by colour labels associated with colour-related knowledge. Consequently, the influence observed could be the result of a form of lexical priming of the label rather than perceptual priming. In order to ensure this possibility, further experiments would need to be developed in which participants would have to carry out a within-category discrimination task on coloured targets primed by congruent or incongruent colour-related words. Indeed, this type of task only seems recruit low-level perceptual processes. Accordingly, if retrieval of colour-related knowledge involves a colour simulation process that influences colour perception *via* neuronal pattern interference, we would always observe perceptual priming.

6.2. Simulation of Hues?

The second problem concerns what is simulating in colour simulation? All the experiments reported here agree, implicitly, on the idea that colour simulation is the simulation of 'colour categories' and not specific hues. Indeed, all studies control the link of congruence between prime and target at a level of colour category. For instance, Richter & Zwaan (2009) used names of colour categories (e.g., BLUE, YELLOW, GREEN) that matched or mismatched with coloured targets at a level of colour categories (e.g., blue, yellow, or green patches). In Heurley *et al.* (in press) or in Nijboer *et al.* (2006), words used represented objects associated with a colour (e.g., BANANA, TOMATO, ARTICHOKE), that matched or mismatched with coloured patches also at a level of colour categories (e.g., yellow, red or green patches). More precisely, in Heurley *et al.* (2012, in press), we used for instance the word 'LETTUCE' and the word 'ARTICHOKE' presented just before two simultaneous green patches in each trials (i.e. congruent trials) or two simultaneous yellow patches (i.e. incongruent trials). In the congruent trials, there was congruence because the labels of the colour associated with the words and with the patches were similar (i.e. label 'green'). Indeed, the difference of hues associated with the object was not taken into account. For instance, the fact that the hue of green associated with the word 'LETTUCE' is lighter than the hue of green associated with 'ARTICHOKE' was not taken into account because the same green colour was used as target each time. Consequently, the link of congruence between words and patches was controlled at the level of colour category and not at a level of specific hue.

At the perceptual level, discrimination of specific hues of a 'colour category' seems natural. For instance, in Beauchamp *et al.* (1999), participants were able to discriminate various hues of a same colour category in order to discriminate if the various patches were arranged from a dark hue to a light hue or were arranged in a disorderly manner (see also Simmons *et al.* 2007; Hsu *et al.* 2011, 2012). Results of studies on categorical perception of colour also support this capacity. Indeed, participants are able to discriminate various hues of a same colour category in within-category discrimination tasks (Roberson & Davidoff 2000; Pilling *et al.* 2003; Gilbert *et al.* 2006; Drivonikou *et al.* 2007; Siok *et al.* 2009). Finally, a recent study on colour memory also reinforces this possibility. In the experiment of Allen *et al.* (2011) participants had to adjust a test patch in order to look like the patch seen just before. Each trial was divided in three phases. In the 'learning phase', participants saw a coloured circle during 60s and were instructed to memorize the colour. Then in a retention phase, participants had to count during 120s. More precisely, beeps indicated to participants to say out loud a number from 0 to 9 in a random order. The goal of this task was to prevent verbal rehearsal of the colour test. In a test phase, participants had to set the new central patch to look like the test colour they remembered seeing previously. Moreover in order to test the role of working memory, participants were divided in two groups: one group with high working memory capacity and the other with low capacity, tested *via* the 'Reading Span' and the 'Operation Span' tests (Daneman & Carpenter 1980; Unsworth *et al.* 2005). Results showed that the differences between coordinates of the selected colour and colour learning, in the specific colourimetric model used, were relatively low suggesting that participants had a good capacity to remember specific hues of learning colour (see also Olkkonen *et al.* 2008; Witze *et al.* 2012). Let us note that this result was found for the two groups of participants showing the independence of working memory capacity. Taken together, these results support the idea that participants are able to discriminate different hues. Consequently, if colour simulation is really a neuronal simulation of colour perception, colour simulation can involve simulation of hues, not just 'colour categories'.

One result of Hsu *et al.* (2011), already presented, partially supports the possibility to simulate specific hues in word reading. Indeed, in their experiments, participants first saw two words on the screen representing two colour-related objects. Then, a third word would appear on the screen and participants had to choose the object that was associated with the most similar colour. In one of the two blocs, the three words belonged to the same colour category and similarity occurred at a level of specific hue. For instance, participants had to compare the target word 'SCHOOL BUS' with two words also associated with the colour yellow: 'BUTTER' and 'EGG YOLK'. In this case, the answer was 'EGG YOLK' because its yellow hue is closest to the yellow hue of a school bus than to the yellow hue of butter. In this condition, Hsu *et al.* (2011) observed an activation of the fusiform and lingual gyri that are also involved in colour perception. Accordingly, this result seems to suggest the possibility to simulate specific hues in word reading. Nevertheless, a major question remains open: can hue simulation influence hue perception? Indeed, no work seems designed to test this hypothesis whereas it is a possible prediction of the simulationist approach of knowledge/perception interaction. In order to test this new prediction, further experiments would need to control the link of congruence between colour-related knowledge and coloured targets at the level of specific hues.

In this section, we developed two limits of current works supporting the simulationist approach of knowledge/perception interaction in colour perception. Firstly, we developed the idea that perceptual tasks generally used could involve lexical processes that could explain the observed priming without the use of a simulation process. To overcome this possibility, we propose to use within-category colour discrimination tasks in order to ensure the recruitment of lexical processes. Secondly, we developed the possibility that colour simulation could involve simulation of specific hues. Indeed, up to now, experiments have only controlled the congruency link at the level of colour category and absolutely not at the level of specific hues. Consequently, it is not possible to know if colour simulation can influence perception at a hue level. To overcome this limitation, new experiments must be carried out controlling the link of congruency at the level of specific hues.

7. Conclusion

"How can perception be altered by language?" was the fundamental question of this article. Indeed, various studies have pointed out the influence of colour-related knowledge on object and colour perception (Tanaka & Presnell 1999; Naor-Raz *et al.* 2006; Mitterer et De Ruiter 2008; Theurriault *et al.* 2009). Thus, the article does not focus directly on the language function but more precisely on how language can interact with perception, specifically with colour perception. We assumed the relevance of the simulationist approach which defends a specific form of retrieval using the same neuronal substrate than low-level perception.

We developed the hypothesis that the use of a common structure creates a meeting point for language and perception by coordinating the two through the neural bases of a distributed memory. Results from neuroimaging studies studying neuronal bases of colour perception and colour simulation allow us to conclude that these two functions are supported by two distributed neuronal networks (Shuren *et al.* 1997; Bartolomeo *et al.* 1998; Simmons *et al.* 2007; Hsu *et al.* 2011, 2012). These two networks can interact at particular point and also seem independent. Consequently, we develop the idea that colour-related knowledge and colour perception can interact through a process of pattern interference (Masson 1995). Based on this idea, we suggest the possibility that colour simulation and colour perception can interact in a bi-directional way even when colour simulation and colour perception are slightly separated in time. In other words, we expected the possibility to obtain a priming effect between colour simulation and colour perception. In support of this hypothesis, we discussed results showing (i) an influence of words denoting colour or colour-related objects on a subsequent low-level colour processing (Heurley *et al.* 2012, in press, Richter & Zwaan 2009), and (ii) results that suggest a reverse influence, from perception to language processing (Nijboer *et al.* 2006; Gebuis *et al.* 2009; Richter & Zwaan 2009). These various works are in accordance with the simulationist view; nevertheless we also develop two limitations of these studies. Indeed, these works all used between-category colour discrimination tasks that allow the intervention of lexical processes in colour perception, not only perceptual processes (Bornstein & Korda 1984; Pilling *et al.* 2003; Roberson & Davidoff 2003; Gilbert *et al.* 2006; Drivonikou *et al.* 2007). This limitation is important because the priming effect observed can be potentially explained in terms of lexical priming rather than perceptual priming. The second limitation is that these works control the congruency link between prime and target at the level of 'colour category' rather than at a level of specific hue. But, if colour simulation is a neuronal simulation of colour perception, hue simulation can be expected because hue perception is possible (e.g., Beauchamp *et al.* 1999; Simmons *et al.* 2007; Allen *et al.* 2011; Hsu *et al.* 2011, 2012). In order to overcome these two limitations and to support and extend the simulationist explanation, we propose new experimental controls.

To conclude, we would like develop two other untested predictions of the simulationist view. First, it is possible to expect an effect of delay that separates colour simulation and colour perception. Indeed, activation of a neuronal pattern can imply a refractory period of their constitutive neurons that preclude their immediate re-use. Consequently, it is possible that with a short SOA, the influence would not be obtained or a slowdown could appear. A slowdown of response times when the previous simulation is congruent with the actual perception has been obtained in two studies. For instance, in the experiment of Estes *et al.* (2008), a slowdown of response times was observed when prime words were associated with a spatial location identical to the location where subsequent perceptual targets appeared (e.g., the word ATTIC followed by a perceptual target on the top of the screen). Meteyard *et al.* (2007) studied the impact sentence implying a lateral movement towards the right or the left-hand side on the perception of a moving stimulus. Once again, when the simulated and perceived movements were identical, response times were longer. These works allow the possibility to observe the same slowdown for a particular SOA in the case of colour perception.

A second possible prediction concerns the neuronal process underlying the influence of colour simulation on colour perception. Possible insight comes from the work of Kristjánsson *et al.* (2007) who focused on the neuronal correlates of the perceptual priming of colour. Kristjánsson *et al.* (2007) replicated the experiment of Maljkovic & Nakayama (1994) while cortical activity was recorded by fMRI. Three coloured diamond shapes placed as a triangle were presented to the participants. The right or left diamond shape placed at the base of the triangle was presented in a different colour. Participants simply had to say in what form appeared the target coloured differently from the others. Kristjánsson *et al.* (2007), in the same way as Maljkovic & Nakayama (1994), observed a priming effect when the colour of the target was the same during two consecutive trials in comparison to when the colour of the target changed from one trial to another. Moreover, Kristjánsson *et al.* (2007) observed a decrease of the BOLD signal (*Blood Oxygen Level Dependent*) in the visual areas involved in the perception of colour. This 'suppression effect' is often observed in perceptual priming tasks and is seen to be the neuronal correlate to perceptual priming (for a review see Wiggs & Martin 1998). In the case where the priming effect observed when colour simulation precedes colour perception results from perceptual priming, it is possible to expect the same decrease of the BOLD signal in the visual areas involved in the perception of colour. Such a result could bring to light the fact that colour simulation and colour perception recruit the same neuronal substrate as well as the fact that the influence of colour simulation on colour perception is supported by the same neuronal process as colour priming.

Finally, in this article, we developed a functional consequence of grounded and embodied language on the sensorimotor areas of the brain: the possibility that language can influence perception through a simulation process (and also the reverse influence). Not only did we focus our attention on how cognition and language are embodied or grounded but we also focused our attention on the functional consequence of embodying cognition for the cognitive system. In line with recent work of the embodied and grounded approach (Wu & Barsalou 2009; Barsalou 2003; Pulvermüller 2010), we tried to build explanations for cognitive functioning without a symbolic and amodal architecture. According to this approach, memory, knowledge, language and perception function in a coordinated way which can either alter or facilitate perception. We agree that this view needs more work in order to overcome various limitations and to test other possible predictions. However, it seems an interesting way to conceive the language/perception interaction while including the constraints of the neuronal substrate underlying cognitive functioning.

References

Allen, Elizabeth C., Sian L. Beilock & Steven K. Shevell. 2011. Working memory is related to perceptual processing: A case from color perception. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 37(4), 1014–1021.

Allport, D. A. 1985. Distributed memory, modular subsystems ad dysphasia. In S. K. Newman & R. Epstein (eds.), *Current Perspectives in Disphasia*, 32–60. Edinburgh: Churchill Livingstone.

Barsalou, Lawrence W. 1999. Perceptual symbol systems. *Behavioral and Brain Sciences* 22, 577–660.

Barsalou, Lawrence W. 2008. Grounded cognition. *Annual Review of Psychology* 59, 617–645.

Barsalou, Lawrence W. 2009. Simulation, situated conceptualization, and prediction. *Philosophical Transactions of the Royal Society of London: Biological Science* 364, 1281–1289.

Barsalou, Lawrence W., W. Kyle Simmons, Aron K. Barbey & Christine D. Wilson. 2003. Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences* 7(2), 84–91.

Bartolomeo, Paolo, Anne-Catherine Bachoud-Lévi & Gianfranco Denes. 1997. Preserved imagery for colours in a patient with cerebral achromatopsia. *Cortex* 33, 369–378.

Beauchamp, Michael S., James V. Haxby, Jonathan E. Jennings & Edgar A. de Yoe. 1999. An fMRI version of the Farnsworth-Munsell 100-hue test reveals multiple colour-selective areas in human ventral occipitotemporal cortex. *Cerebral Cortex* 9, 257–263.

Biederman, Irvin. 1987. Recognition-by-components: A theory of human image understanding. *Psychological Review* 94(2), 115–147.

Bocanegra, Bruno R. & René Zeelenberg. 2009. Emotion improves and impairs early vision. *Psychological Science* 20(6), 707–713.

Bornstein, Marc H. & Nancy O. Korda. 1984. Discrimination and matching within and between hues measured by reaction times: Some implications for categorical perception and levels of information processing. *Psychological Research* 46, 207–222.

Bouvier, Seth E. & Stephen A. Engel. 2006. Behavioral deficits and cortical damage loci in cerebral achromatopsia. *Cerebral Cortex* 16, 183–191.

Bramão, Ignês, Luis Faísca, Christian, Forkstam, Alexandra Reis & Karl M. Petersson. 2010. Cortical brain regions associated with colour processing: An fMRi study. *The Open Neuroimaging Journal* 4, 164–173.

Bramão, Ignês, Alexandra Reis, Karl M. Petersson & Luis Faísca. 2011. The role of colour information on object recognition: A review and meta-analysis. *Acta Psychologica* 138(1), 244–253.

Bruner, Jerome S. & Leo Postman. 1949. On the perception of incongruity: A paradigm. *Journal of Personality* 18, 108–143.

Bruner, Jerome S., Leo Postman & John Rodrigues. 1951. Expectation and the perception of colour. *American Journal of Psychology* 64, 216–227.

Chao, Linda L. & Alex Martin. 1999. Cortical regions associated with perceiving, naming, and knowing about colours. *Journal of Cognitive Neuroscience* 11(1), 25–35.

Corbetta, Maurizio, Francis M. Miezin, Suzan Dobmeyer, Gordon L. Shulman & Steve E. Petersen. 1991. Shape, colour, and speed: Functional anatomy by positron emission tomography. *The Journal of Neuroscience* 11(8), 2393–2402.

Daneman, Meredyth & Patricia Carpenter. 1980. Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior* 19, 450–466.

De Vreese, Luc P. 1991. Two systems for colour-naming defects: Verbal disconnection vs colour imagery disorder. *Neuropsychologia* 29, 1–18.

Delk, John L. & Samuel Fillenbaum. 1965. Differences in perceived colour as a function of characteristic colour. *American Journal of Psychology* 78, 290–293.

Drivonikou, Gilda V., Paul Kay, Terry Regier, Richard B. Ivry, Aubrey L. Gilbert, Anna Franklin *et al.* 2007. Further evidence that Whorfian effects are stronger in the right visual field than the left. *Proceedings of the National Academy of Sciences of the United States of America* 104, 1097–1102.

Duncker, Karl. 1939. The influence of past experience upon perceptual properties. *American Journal of Psychology* 52, 255–265.

Engel, Stephen, Xuemei Zhang & Brian Wandell. 1997. Colour tuning in human visual cortex measured with functional magnetic resonance imaging. *Nature* 388, 68–71.

Estes, Zachari, Michelle Verges & Lawrence W. Barsalou. 2008. Head up, foot down: Object words orient attention to the objects' typical location. *Psychological Science* 19(2), 93–97.

Farah, Martha J. & James L. McClelland. 1991. A computational model of semantic memory impairment: Modality-specificity and emergent category-specificity. *Journal of Experimental Psychology: General* 120, 339–357.

Farnsworth, Dean. 1957. *The Farnsworth-Munsell 100-hue for the examination of the colour vision*. Baltimore, MD: Munsell Colour Compagny.

Fodor, Jerry. A. 1983. *The Modularity of Mind: An Essay on Faculty Psychology*. Cambridge, MA: MIT Press.

Gebuis, Titia, Tanja. C. W. Nijboer & Maarten J. van der Smagt. 2009. Of colored numbers and numbered colors: Interactive processes in grapheme-color synesthesia. *Experimental Psychology* 56(3), 180–187.

Gegenfurtner, Karl R. 2003. Cortical mechanisms of colour vision. *Natural Review: Neuroscience* 4, 563–572.

Gegenfurtner, Karl R. & Daniel C. Kiper. 2003. Color vision. *Annual Review of Neuroscience* 26, 181–206.

Gibson, James J. 1979. *The Ecological Approach to Visual Perception*. New York: Psychology Press.

Gilbert, Aubrey. L., Terry Regier, Paul Kay & Richard B. Ivry. 2006. Whorf hypothesis is supported in the right visual field but not the left. *Proceedings of the National Academy of Sciences of the United States of America* 103, 489–494.

Goldberg, Robert E., Charles A. Perfetti & Walter Schneider. 2006. Perceptual knowledge retrievaI activates sensory brain regions. *Journal of Neuroscience* 26, 4917–4921.

Goldstone, Robert L. & Andrew T. Hendrickson. 2009. Categorical perception. *Interdisciplinary Reviews: Cognitive Science* 1, 69–78.

Guylás, Balázs & Per E. Roland. 1994. Processing and analysis of form, colour and binocular disparity in the brain: Functional anatomy by positron emission tomography. *European Journal of Neuroscience* 6, 1811–1828.

Hansen, Thorsten, Maria Olkkonen, Sebastian Walter & Karl R. Gegenfurtner. 2006. Memory modulates colour appearance. *Nature Neuroscience* 9(11), 1367–1368.

Harper, Robert S. 1953. The perceptual modifications of coloured figures. *American Journal of Psychology* 66, 86–89.

Hebb, Donald O. 1949. *The Organization of Behavior: A Neuropsychological Theory*. New York: John Wiley & Sons.

Heurley, Loïc P., Thibaut Brouillet, Gabrielle Chesnoy & Denis Brouillet. 2012. Color perception involves color representations firstly at a semantic level and then at a lexical level. *Cognitive Processing*. doi: 10.1007/s10339-012-0527-z.

Heurley, Loïc P., Elisabeth Rolland-Thiers, Laurent P. Ferrier & Denis Brouillet. In press. Influence des connaissances associées à une couleur dans une tâche de discrimination chromatique. *L'Année Psychologique*.

Hinton, Geoffrey. 1981. Implementing semantic networks in parallel hardware. In G. Hinton & J. A. Anderson (eds.), *Parallel Model of Associative Memory*, 191–217. Hillsdale, NJ: Lawrence Erlbaum Associates.

Hsu, Nina S., Steven M. Frankland & Sharon L. Thompson-Schill. 2012. Chromaticity of color perception and object color knowledge. *Neuropsychologia* 50, 327–333.

Hsu, Nina S., David J. M. Kraemer, Robyn T. Oliver, Margaret L. Schlichting & Sharon L. Thompson-Schill. 2011. Color, context, and cognitive style: Variations in color knowledge retrieval as a function of task and subject variables. *Journal of Cognitive Neuroscience* 23(9), 2544–2557.

Ishihara, Shinobu. 1974. *Tests for Colour-Blindness*. Tokyo: Kanehara Shup.

Kay, Paul & Willet Kempton. 1984. What is the Sapir-Whorf Hypothesis? *American Anthropologist* 86(1), 65–79.

Kellenbach, Marion L., Matthew Brett & Karalyn Patterson. 2001. Large, colourful, or noisy? Attribute- and modality-specific activations during retrieval of perceptual attribute knowledge. *Cognitive, Affective, & Behavioral Neuroscience* 1(3), 207–221.

Kirby, John, Phillip Moore & Neville Schofield. 1988. Verbal and visual learning styles. *Contemporary Educational Psychology* 13, 169–184.

Klein, George S. 1964. Semantic power measured through the interference of words with color-naming. *American Journal of Psychology* 77, 576–588.

Kristjánsson, Árni, Patrik Vuilleumier, Sophie Schwartz, Emiliano Macaluso & Jon Driver. 2007. Neural basis for priming of pop-out during visual search revealed with fMRI. *Cerebral Cortex* 17, 1612–1624.

Kubat, Rony, Daniel Mirman & Deb Roy. 2009. Semantic context effects on colour categorization. In Niels A. Taatgen & Hedderik van Rijn (eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society*, 491–495. Austin, TX: Cognitive Science Society.

Loftus, Elizabeth F. 1977. Shifting human color memory. *Memory & Cognition* 5(6), 696–699.

Maljkovic, Vera & Ken Nakayama. 1994. Priming of pop-out: I. Role of features. *Memory & Cognition* 22, 657–672.

Martin, Alex. 2007. The representation of object concepts in the brain. *Annual*

Review of Psychology 58, 25–45.

Martin, Alex. 2009. Circuits in mind: The neural foundations for object concepts. In Michael S. Gazzaniga (ed.), *The Cognitive Neurosciences*, 4th edn., 1031–1045. Cambridge, MA: MIT Press.

Martin, Alex & Linda L. Chao. 2001. Semantic memory and the brain: Structure and process. *Current Opinion in Neurobiology* 11, 194–201.

Martin, Alex, James V. Haxby, Francois M. Lalonde, Cheri L. Wiggs & Leslie G. Ungerleider. 1995. Discrete cortical regions associated with knowledge of colour and knowledge of action. *Science* 270, 102–105.

Masson, Michael E. J. 1995. A distributed memory model of semantic priming. *Journal of Experimenlal Psychology: Learning, Memory and Cognition* 21(1), 3–23.

McClelland, James L. & David E. Rumelhart. 1985. Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General* 114, 159–197.

Meteyard, Lotte, Behador Bahrami & Gabriella Vigliocco. 2007. Motion detection and motion verbs: Language affects low-level visual perception. *Psychological Science* 18(11), 1007–1013.

Mitterer, Holger & Jan Peter de Ruiter. 2008. Recalibrating colour categories using world knowledge. *Psychological Science* 19(7), 629–634.

Naor-Raz, Gallit, Michael J. Tarr & Daniel Kersten. 2003. Is colour an intrinsic property of object representation? *Perception* 32, 667–680.

Neisser, Ulric. 1967. *Cognitive Psychology*. New York: Appleton-Century-Crofts.

Nijboer, Tanja C. W., Titia Gebuis, Susan F. te Pas & Maarten J. van der Smagt. 2011. Interactions between colour and synaesthetic colour: An effect of simultaneous colour contrast on synaesthetic colours. *Vision Research* 51, 43–47.

Nijboer, Tanja C. W., Maarten J. van der Smagt, Martine J. E. Van Zandvoort & Edward H. F. De Haan. 2007. Colour agnosia impairs the recognition of natural but not of non-natural scenes. *Cognitive Neuropsychology* 24(2), 152–161.

Nijboer, Tanja C. W., Martine J. E. van Zandvoort & Edward H. F. de Haan. 2006. Seeing red primes tomato: Evidence for comparable priming from colour and colour name primes to semantically related word targets. *Cognitive Processes* 7, 269–274.

Oliva, Aude & Philippe G. Schyns. 2000. Diagnostic colors mediate scene recognition. *Cognitive Psychology* 41, 176–210.

Oliver, Robyn T. & Sharon L. Thompson-Shill. 2003. Dorsal stream activation during retrieval of object size and shape. *Cognitive, Affective, & Behavioral Neuroscience* 3(4), 309–322.

Olkkonen, Maria, Thorsten Hansen & Karl R. Gegenfurtner. 2008. Color appearance of familiar objects: Effects of object shape, texture, and illumination changes. *Journal of Vision* 8(5), 1–16.

Pilling, Michael, Alison Wiggett, Emre Özgen & Ian R. L. Davies. 2003. Is colour 'categorical perception' really perceptual? *Cognition* 31, 538–551.

Pohl, Rüdiger. F. 2004. Effects of labelling. In Rüdiger F. Pohl (ed.), *A Handbook of Fallacies and Biases in Thinking, Judgment and Memory*, 327–344. New York:

Psychology Press.

Pulvermüller, Friedmann. 2001. Brain reflections of words and their meaning. *Trends in Cognitive Sciences* 5(12), 517–524.

Pulvermüller, Friedmann. 2003. *The Neuroscience of Language*. Cambridge: Cambridge University Press.

Pulvermüller, Friedmann. 2010. Brain-language research: Where is the progress? *Biolinguistics* 4(2–3), 255–288.

Pylyshyn, Zenon W. 1984. *Computation and Cognition*. Cambridge, MA: MIT Press.

Reilhac, Gilles & Manuel Jiménez. 2006. Amorçage de la couleur typique d'un objet lors d'une tâche de catégorisation. *Canadian Journal of Experimental Psychology* 60(4), 285–293.

Richter, Tobias & Rolf A. Zwaan. 2009. Processing of colour words activates colour representations. *Cognition* 111(3), 383–389.

Roberson, Debi & Jules Davidoff. 2000. The categorical perception of colors and facial expressions: The effect of verbal interference. *Memory & Cognition* 28, 977–986.

Schacter, Daniel L., Kenneth A. Norman & Wilma Koutstaal. 1998. The cognitive neuroscience of constructive memory. *Annual Review of Psychology* 49, 289–318.

Shuren, Jeffrey E., Thomas G. Brott, Bruce K. Schefft & Wes Houston. 1996. Preserved color imagery in an achromatopsic. *Neuropsychologia* 34, 485–489.

Simmons, W. Kyle, Vimal Ramjee, Michael S. Beauchamp, Ken McRae, Alex Martin & Lawrence W. Barsalou. 2007. A common neural substrate for perceiving and knowing about colour. *Neuropsychologia* 45, 2802–2810.

Siok, Wai Ting, Paul Kay, William. S. Y. Wang, Alice H. D. Chan, Lin Chen, Kang Kwong Luke *et al.* 2009. Language regions of brain are operative in color perception. *Proceedings of the National Academy of Sciences of the United States of America* 106(20), 8140–8145.

Spence, Charles, Carmel A. Levitan, Maya U. Shankar & Massimiliano Zampini. 2010. Does food colour influence taste and flavor perception in humans? *Chemosensory Perception* 3, 68–84.

Stein, Barry. E. & Alex. M. Meredith. 1993. *The Merging of the Senses*. Cambridge, MA: MIT Press.

Tanaka, James W. & Lynn M. Presnell. 1999. Colour diagnosticity in object recognition. *Perception & Psychophysics* 61(6), 1140–1153.

Therriault, David J., Richard H. Yaxley & Rolf A. Zwaan. 2009. The role of colour diagnosticity in object recognition and representation. *Cognitive Processes* 10(4), 335–342.

Unsworth, Nash, Richard P. Heitz, Josef C. Schrock & Randall W. Engle. 2005. An automated version of the operation span task. *Behavior Research Methods* 37, 498–505.

Van Zandvoort, Martine J. E., Tanja C. W. Nijboer & Edward de Haan. 2007. Developmental colour agnosia. *Cortex* 43, 750–757.

Versace, Rémy, Elodie Labeye, Guillemette Badard & Marylène Rose. 2009. The contents of long-term memory and the emergence of knowledge. *European Journal of Cognitive Psychology* 21(4), 522–560.

Wiggs, Cheri L. & Alex Martin. 1998. Properties and mechanisms of perceptual priming. *Current Opinion in Neurobiology* 8, 227–233.

Wiggs, Cheri L., Jil Weisberg & Alex Martin. 1999. Neural correlates of semantic and episodic memory retrieval. *Neuropsychologia* 37, 103–118.

Witze, Christoph, Hanna Valkova, Thorsten Hansen & Karl R. Gegenfurtner. 2011. Object knowledge modulates colour appearance. *i-Perception* 2, 13–49.

Wu, Ling-ling, & Lawrence W. Barsalou. 2009. Perceptual simulation in conceptual combination: Evidence from property generation. *Acta Psychologica* 132, 173–189.

Yee, Eiling, Sarah. Z. Ahmed & Sharon L. Thompson-Schill. 2012. Colorless green ideas (can) prime furiously. *Psychological Science* 23, 364–369.

Zeki, Semir. 1983a. Colour coding in the cerebral cortex: The reaction of cells in monkey visual cortex to wavelengths and colours. *Neuroscience* 9, 741–765.

Zeki, Semir. 1983b. Colour coding in the cerebral cortex: The responses of the wavelength-selective and colour-coded cells in monkey visual cortex to changes in wavelength composition. *Neuroscience* 9, 767–781.

Zeki, Semir. 1983c. The distribution of wavelength and orientation selective cells in different areas of the monkey visual cortex. *Proceedings of the Royal Society B: Biological Sciences* 217, 449–470.

Zeki, Semir & Ludovica Marini. 1998. Three cortical stages of colour processing in the human brain. *Brain* 121, 1669–1685.

Zeki, Semir, James D. G. Watson, C. J. Lueck, K. J. Friston, C. Kennard & R. S. J. Frackowiak. 1991. A direct demonstration of functional specialization in human visual cortex. *The Journal of Neuroscience* 17(3), 641–649.

Zwaan, Rolf A. & Michael P. Kaschak. 2009. Language in the brain, body and world. In Philip Robbins & Murat Aydede (eds.), *The Cambridge Handbook of Situated Cognition*, 368–381. Cambridge: Cambridge University Press.

Loïc P. Heurley, Audrey Milhau, Gabrielle Chesnoy-Servanin, Laurent P. Ferrier, Thibaut Brouillet, and Denis Brouillet
Université Paul Valéry
Epsylon Laboratory
Site Saint Charles
34000 Montpellier
France

heurleyloic@yahoo.fr, audrey.milhau@gmail.com, gabrielle.chesnoy@chu-st-etienne.fr, laurent.paul.ferrier@gmail.com, thibaut.brouillet@gmail.com, and d.brouillet@yahoo.fr

Digitized Fossil Brains: Neocorticalization

Harry J. Jerison

This report is based on 3D digital scans of endocasts of 110 species of fossil mammals and 35 species of living mammals. It presents direct evidence of the last 60 million years of brain evolution. Endocasts are casts of the cranial cavity. They are brainlike in size and shape, and their surface features can be named as if they were brain structures. Although endocast data are restricted to outer surfaces of brains, a few inferences about inner structure are possible. Neocortex in the forebrain, for example, is identifiable and measurable as cerebral forebrain on the endocast dorsal to the rhinal fissure. An important result in this report is that surface area of neocortex as identified on endocasts appears to have reached a maximum of about 80% of the total endocast surface area in anthropoid primates including humans. This may be a fundamental limitation in brain size. The average neocorticalization percentage for mammals as a whole rose from about 20% to about 50% of the surface area during the 60 million years covered by this analysis. Neocorticalization is associated with the evolution of higher mental processes, including the evolution of language as a hominin specialization. The limitation of the increase in *relative* amount of neocortex is similar in all anthropoids. Neocortex is greater in *absolute* area in living humans because the total size of the hominin brain is so much larger than in other primates.

Keywords: cerebral cortex; encephalization; neuroimaging; primates

1. Introduction

The evolution of the brain is analyzed with fossil endocasts arranged according to geological age. This direct information on the brain's evolution provides a context for the data of comparative neuroanatomy of brains of living species such as those reviewed by Hofman (2012) and by Kaas (2012), although in endocasts many comparative neurological measurements are impossible. Cortex buried in sulci and fissures is not observable, nor are data on white matter versus gray matter. My digitizing procedure provides error-free measurements of endocasts, most unusually of surface areas otherwise unmeasurable on irregular solids. The program also provides familiar measurements of lengths and volume (see http://headus.com.au).

ISSN 1450–341

Biolinguistics 6.3–4: 383–392, 2012
http://www.biolinguistics.eu

The measurements of surface area, as I will review them, show that progressive neocorticalization occurred in mammalian evolution. It would be fruitful if comparative neuroanatomists would routinely add information on the outer surface areas to reports on micro and molecular anatomy of brains they study to improve the correlation between fossil data on endocasts with their data on living brains.

Endocasts are not brains. They are usually rock, plaster, or latex casts molded by the cranial cavity. In birds and mammals they are remarkably brainlike in size and shape. The literature on the evolution of the brain is usually on the comparative neuroanatomy of living brains with living species arranged as if they represent a phylogenetic tree. Fossil endocasts represent a more truly evolutionary pattern of brain evolution, which can help anchor the data of comparative neuroanatomy. Endocasts, however, are limited to data on external surfaces. In this report I present measurements of external surface area in endocasts in the centimeter-gram-second (cgs) system and present an illustrative example of the use of such measurements to analyze the evolution of neocorticalization.

Endocasts are not always perfect pictures of brains, but the natural endocast in Figure 1 shows just how good an image they provide. That endocast is of the 37 Ma (million years ago) artiodactyl *Bathygenys reevesi*, one of the many fossil nonprimate brains in my sample. My conclusions are based on measurements of endocasts of many species of mammals including those of primates. This report is on the evolution of neocortex in mammals during the past 60 million years. Cerebral cortex is a uniquely mammalian trait, and *Bathygenys* is a fine example of the identification of that trait in a fossil mammal's endocast. For biolinguistics as a trait that evolved I will emphasize primate neocortex, of course, but all of the mammals that I have worked with provide as good an example of the identification of neocortex. My judgments are about behavioral capacities controlled by neocortex in all mammals and can be applied in particular to neocortical control of language in the human species.

Figure 1: Natural endocranial cast of Bathygenys reevesi, a late Eocene artiodactyl that lived about 37 million years ago. The cast is University of Texas specimen 40209-431 from the Big Bend region of the Rio Grande River in North America (Wilson 1971; see also Jerison 1979, Macrini 2009)

2. Methods and Results

The utility of one's measurements depends on the quality of the endocast as an image of the brain. Figure 1 of the endocast of *Bathygenys reevesi* shows how good the image can be. Laser scans in turn provide accurate information on volume, length, and surface area of any endocast, but because olfactory bulbs are often distorted in endocasts I exclude them from my computations. My scanning software enables me to mark surface area for measurement. Those for *Bathygenys* are in the legend to Figure 2, which shows how the scanned image is actually clearer than the natural endocast in illustrating the important presence of the rhinal fissure. An endocast is a picture only of the brain's outer surface, but there is a good relationship between the area hidden in the convolutions and the total area. Endocast surface area is a good base for reasoning about total cortical surface and its implications for the evolution of higher mental processes. The surface area of the cerebral cortex is recognized as estimating the total number of neurons, of information processing units.

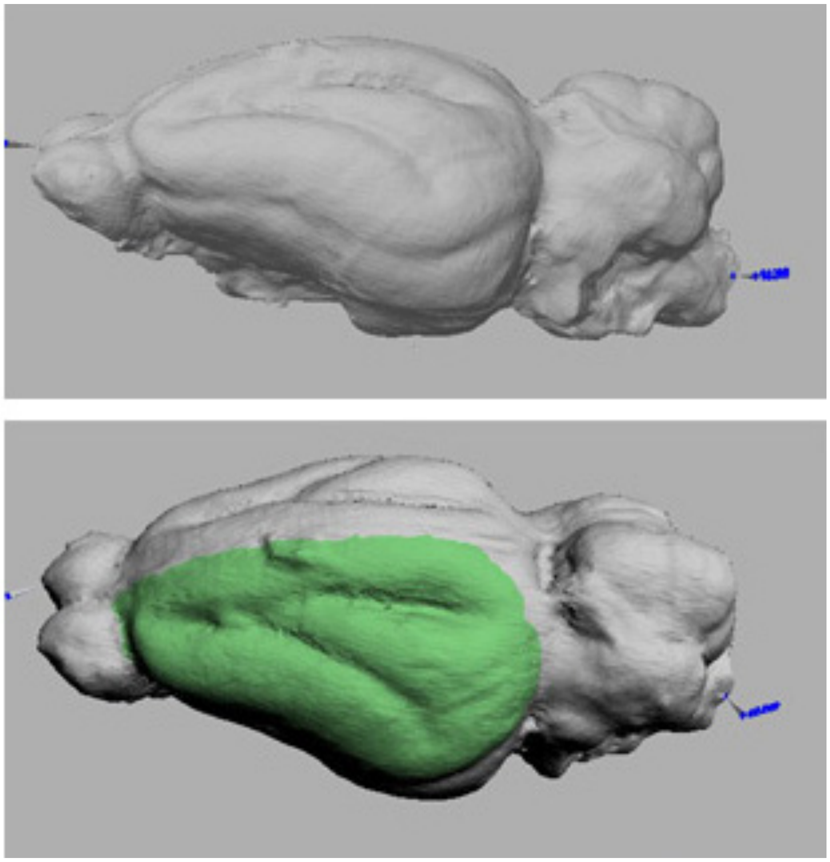


Figure 2: Laser scan of Bathygenys reevesi. Top is lateral view of 3D image; bottom is dorsal view with marked half of the neocortical surface area in green. Rhinal Fissure is very evident. Endocast length = 5.2 cm; volume = 11.9 ml; total surface area = 34 cm^2; olfactory bulb surface area = 2.5 cm^2; total neocortex surface area = 9.4 cm^2. Neocorticalization ratio = 9.4/(34-2.5) = 30%

The measurement of neocorticalization depends partly on one's ability to mark the limits of neocortex. This region is visible on the lateral surface of the brain and endocast. Its ventral border is the rhinal fissure; anterior border is at the entry of the olfactory tract; a dorsal border is at the midline, and posterior border is at the end of forebrain. None of these borders are perfectly visible and depend on the judgment of the investigator. Because of artifacts introduced by measurements of olfactory bulbs in fossils I always subtracted their surface area from total endocast surface area. Neocorticalization was then measured as the surface area of neocortex in cm^2 relative to the entire endocast surface area reduced by olfactory bulbs. In the case of *Bathygenys* it was 30%. Slightly less than 1/3rd of the endocast/brain of *Bathygenys* was devoted to neocortex.

In addition to the measurement of surface area of neocortex relative to the rest of the brain one determines the geological age of one’s fossils. *Bathygenys* marks the Chadronian at the end of the Eocene, which was about 37 million years ago (37 Ma).

A word now about the use of the Rhinal Fissure as a ventral margin of neocortex. Examining many slides in the Wisconsin brain collection this fissure is evident in all 275 of their species of mammals. Figure 3 shows additional evidence in the Armadillo. Rhinal fissure is marked on the lateral surface of the brain. A section about midway through the neocortex of this species (Wisc 60-465) shows the heavily stained layer of nerve cells in the cerebral cortex ventral to the rhinal fissure. This is Lamina II of the cortex and is a landmark identifying paleocortex as opposed to neocortex.

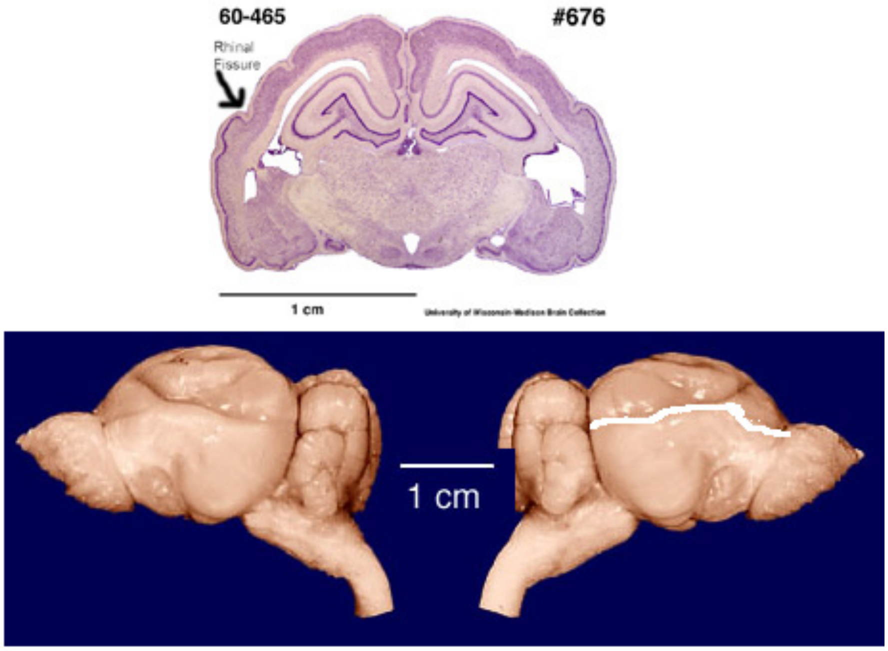


Figure 3: Armadillo (Dasypus novemcinctus) Top: Section through the brain, showing the heavy staining of paleocortical Lamina I of the section ventral to the rhinal fissure. Bottom: Lateral views of whole brain; Rhinal Fissure marked on the right hemisphere of the brain showing its origin at the border of the olfactory tract as it enters the brain

Acknowledging our interest in primate evolution my scan of an Eocene prosimian primate endocast, *Adapis parisiensis*, is shown in Figure 4. Its neocortex is as I marked it. The living galago brain beside it happens to be about the same size, 10 ml or so. The images of the fossil endocast and of the living brain are remarkably similar, but differences are also evident. *Galago* neocortex covers a larger portion of the surface area. Rhinal fissure is partially masked in both specimens by ventral neocortex. An important point can be made about the apparent size of the medulla. It is much thicker on the endocast than on the brain. This reflects the fact that the 'medulla' of the endocast is actually a cast of the foramen magnum. This entry of the spinal cord as medulla into the cranial cavity is enlarged to contain blood vessels and the venous *cisterna magnum* that surround medulla. (I have argued that this cistern is part of a cushioning system that protects medulla from ballistic brain movements within the cranial cavity. It is, therefore related to both brain size and body size. The dimensions of the foramen magnum have been used to estimate body size for analysis of encephalization. Although dimensions of medulla are excellent for this purpose those of the foramen magnum as an independent variable to replace body size in allometric analysis 'confound' a brain size effect with the body size effect.)

In *Adapis* the percentage of neocortex area relative to whole brain area was 51% as I measured it. This degree of neocorticalization is obviously greater than in its contemporary the artiodactyl, *Bathygenys*. I have no data on neocorticalization in the brain of *Galago* which was prepared at the University of Wisconsin, but it is clearly greater than my measure on the *Adapis* endocast. This pattern in differential brain evolution is evident in most of my data. With respect to encephalization, *Adapis* weighed about 1700 grams. *Galago* weighs about 250 grams. The similarity of the two brains in size as well as shape is striking, reflecting the much greater encephalization of the living lemuroid compared to its Eocene relative. The comparisons of brain and endocast in these prosimians support our treatment of endocasts as brains.

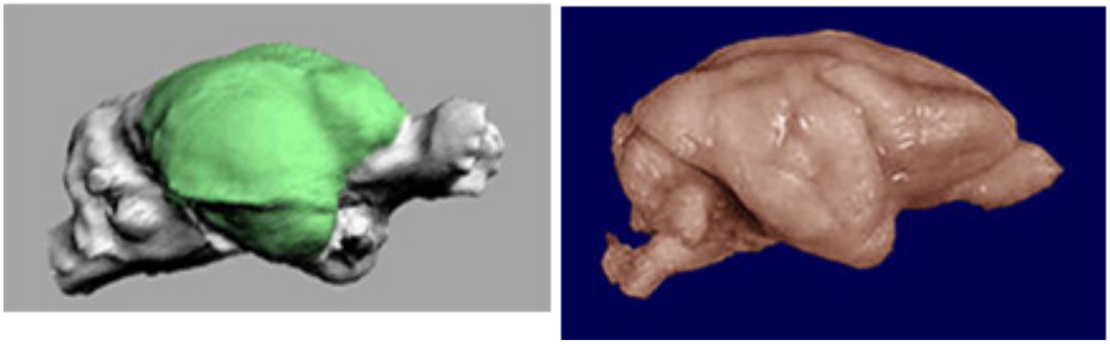


Figure 4: Endocast of late Eocene prosimian, Adapis parisiensis (Field Museum specimen 59259) and living Galago senegalensis (U Wisc 61-686), both about the same size, about 10 ml. (by permission)

Another reservation about the comparability of endocasts and brains is evident in hominin brains in Figure 5. This illustration is in four parts. At the top is a photograph of lateral views of the human brain in the Wisconsin Brain Collection. Left hemisphere is at the left and the right hemisphere at the right. Those familiar with living brains at autopsy will immediately recognize that as in a large majority of human brains the Sylvian Fissure of the left hemisphere is longer than the right. The lengthened Sylvian allows space for Heschl's gyrus buried in the region of the insula. Hemispheric asymmetry in this respect has frequently been measured and analyzed in the human brain. Neither of the endocasts in the lower half of Figure 6 provides an image of the Sylvian fissure. The Falk endocast is significant for other reasons (Falk 2012), showing the layout of cerebral arteries, which are not visible in the endocast at the lower right, prepared from MRI. For our analysis of the evolution of the capacity for language I regret the absence of surface features in endocasts that represent measurable language areas in the brain. Unlike the Sylvian Fissure the Rhinal Fissure and other boundaries of neocortex as present in endocasts enable one to measure and analyze neocorticalization. The new data in this report are on neocorticalization, which is also important for the evolution of language in hominins.

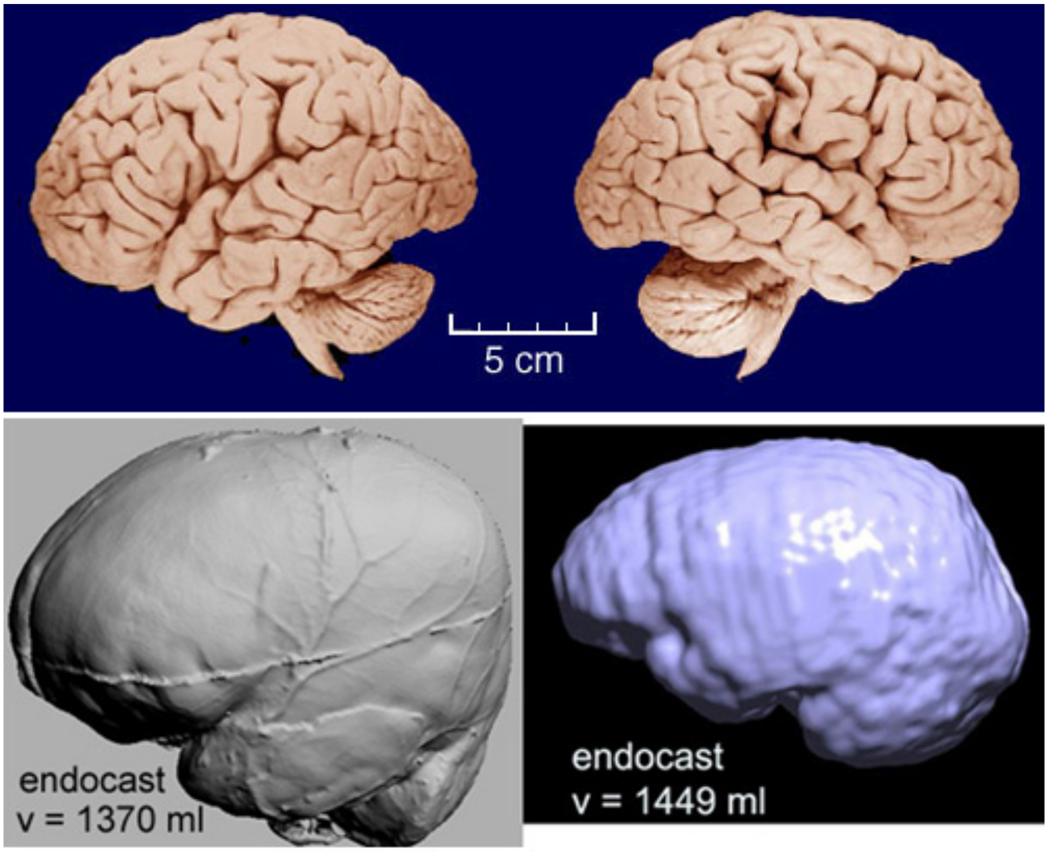


Figure 5: Views of the human brain. Preparation at autopsy at the top from the Wisconsin brain collection. Endocast at bottom left is from Dean Falk's endocast collection and had been prepared from a skull collected for anthropological collections (Falk 1992, by permission). Endocast at the bottom right is from an elderly human prepared from an MRI scan in the author's collection.

3. Results: Neocorticalization

Figure 6 graphs a quantitative analysis of neocorticalization as measured on digitized images of mammalian endocasts. I am in the midst of a monograph for the Field Museum in Chicago with more details on my collection of these images: about 150 species of mammals. Digitization was with laser scans (see http://hjerison.bol.ucla.edu). Many of my scans are of specimens collected by the late Professor Len Radinsky of the University of Chicago who deposited them at the Field Museum to form its Radinsky Collection.

A general point about the data of Figure 6. We see how primates are all 'above average'. I do not label the data on other groups in this graph. I could have identified a dozen species of equids, for example, which are first known in the fossil record of about 55 Ma. A side issue for us, they have been 'average' (near the regression line) for most of their evolutionary history. Marsupials as a group have always been 'below average' as have most of the fossil Neotropical species of South America. These are secondary matters for biolinguistics but appropriate for a broad picture. For this report I present the data most relevant for the evolution of language, in particular in the human species. I place my statistical words in quotation marks to acknowledge the questionable theoretical status of my regression analysis.

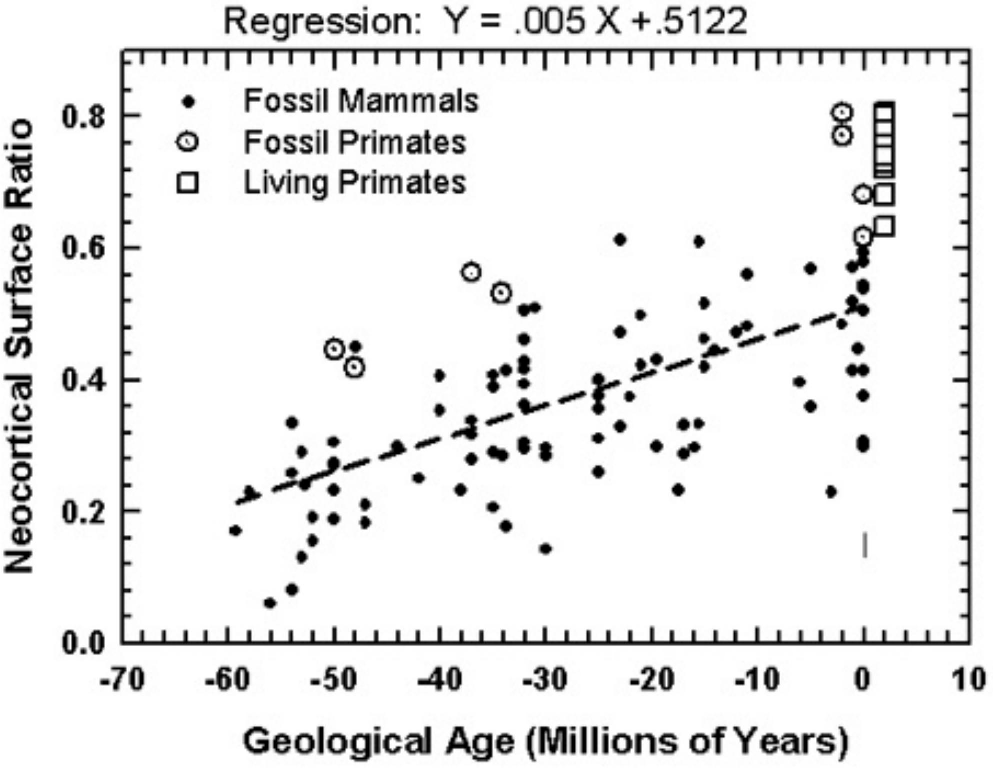


Figure 6: Neocorticalization ratios measured by the surface area of neocortex relative to whole endocast surface area. Fossil sample marked by dots and circles, living sample by squares.

The graph in Figure 6 shows a regression equation fitted to all of the fossil data. The unit **X** is in millions of years before the present (Ma). For example the present geological age has Ma = **X** = 0 and the 'average' neocorticalization is now 0.51, or 51%. That means that at present an average mammal has 51% of its brain's surface area devoted to neocortex. The earliest graphed geological age is 60 Ma. At **X** = –60 the average amount of neocorticalization was (.005)(–60) + .51 = 0.51–0.3 = 0.21. 21% of the surface area of the brain of the average mammals of 60 Ma was devoted to neocortex. The linear regression equation fitted to the data may be interpreted as showing an average of 5% increase in neocortex in mammals per 10 million years.

If we are not carried away by the use of averages this is a fine way of summarizing what we observe. Regression analysis is based on working with a sample taken randomly from its population, and it is well to keep that in mind. The primates in Figure 6 form a fine cohort for our purpose. The four Eocene species were prosimians (Infraorder Strepsirrhini) related to living lemuroids. The four Plio-Pleistocene species were two australopithecines in the hominin lineage and two prosimians. Two points of interest: Both groups of primates were 'above average' in neocorticalization that is described by the regression line. With respect to the theory of regression analysis they do not represent a random selection from the mammals as a class. But it is helpful to recognize that there is something about brain size in primates that makes them a specialized order of mammals with respect to enlarged brains.

4. Discussion

The most unusual feature of the graph for me was the comparison between living and fossil primates. Here are a few of the numbers. The two australopithecines are Taung and Sk 1585 (Swartkrans) and were respectively 80.5% and 77.2% neocorticalized. I have two human skulls collected by Falk (2012) in my living primate sample, and these are 80.0% and 77.7% neocorticalized. I have one chimpanzee and it's 80.7% neocorticalized, which makes it my top performer by 0.2%. I have a Patas monkey from Africa, and it is 78% neocorticalized. I have two New World Saki monkeys and these are 78.0% and 79.0% neocorticalized. The point is that neocorticalization in primates has topped out at about 80%. I have guessed that the growth of mammalian brains is limited with respect to the amount of neocortex that is developed. This is presumably genetically determined.

Assuming a genetic developmental barrier of 80%, one can prepare an oddly inverted picture of the evolutionary enlargement of brain size in hominids. Imagine a scenario in which the beginnings of language appeared in very early hominins, perhaps in australopithecines (Jerison 2001, 2007; Falk 2012). Even in its early stages and like all new traits this must have required at least moderately extensive neural network for its control. The network is now certainly much more extensive. I am impressed by recent publications of colleagues (e.g., Pulvermüller 2010; see also Braitenburg & Schuez 1992) that display the remarkably extensive neocortical neural networks involved with language. Broca's and Wernicke's areas are only a fraction of those revealed by brain scans of living humans performing language tasks. My speculation is straightforward. Assuming a selective advantage for the early adaptations and the later developments of language in our hominin lineage there would have to be space for the language network in the neocortex. To enlarge our neocortex there remains the 80% barrier. Hence, to produce an appropriately enlarged mass of neocortex the whole brain would have to be enlarged. My speculation is that the remarkable extent of human evolutionary encephalization resulted from selection to provide additional neocortex in the face of the 80% barrier.

References

Braitenberg, V. & A. Schuez. 1992. Basic features of cortical connectivity and some considerations on language. In J. Wind, B. Chiarelli, B. H. Bichakjian, A. Nocentini & A. Jonker (eds.), *Language Origin: A Multidisciplinary Approach*, 89–102. Dordrecht: Kluwer.

Falk, D. 1992. *Braindance*. New York: Holt.

Falk, D. 2012. Hominin paleoneurology: Where are we now? In M. A. Hofman & D. Falk (eds.), *Progress in Brain Research*, vol. 195, 255–272. Amsterdam & New York: Elsevier.

Hofman, M. A. 2012. Design principles of the human brain: An evolutionary perspective. In M. A. Hofman & D. Falk (eds.), *Progress in Brain Research*, vol. 195, 373–390. Amsterdam & New York: Elsevier.

Jerison, H. J. 1979. The evolution of diversity in brain size. In M. E. Hahn, C. Jensen & B. C. Dudek (eds.). *Development and Evolution of Brain Size: Behavioral Implications*, 29–57. New York: Academic Press.

Jerison, H. J. 2001. Adaptation and preadaptation in hominid evolution. In P. V. Tobias, M. A. Raath, J. Moggi-Cecchi & G. A. Doyle (eds.), *Humanity from African Naissance to Coming Millennia*, 373–378. Florence: Firenze University Press and Johannesburg: Witwatersrand University Press.

Jerison, H. J. 2007. Fossils, brains, and behavior. In S. Watanabe & M. A. Hofman (eds), *Integration of Comparative Neuroanatomy and Cognition*, 13–31. Tokyo: Keio University Press.

Kaas, J. H. 2012. The evolution of neocortex in primates. In M. A. Hofman & D. Falk (eds.), *Progress in Brain Research*, vol. 195, 91–102. Amsterdam: Elsevier.

Macrini, T. E. 2009. Description of a digital cranial endocast of Bathygenys reevesi (Merycoidodontidae; Oreodontoidea) and implications for apomorphy-based diagnosis of isolated, natural endocasts. *Journal of Vertebrate Paleontology* 29, 1199–1211.

Pulvermüller, F. 2010. Brain-language research: Where is the progress? *Biolinguistics* 4, 255–266.

Wilson, J. A. 1971. Early Tertiary vertebrate faunas, Vieja Group. Trans- Pecos Texas: Agriochoeridae and Merycoidodontidae. *Texas Memorial Museum Bulletin* 18, 1–83.

Harry J. Jerison
University of California
Department of Psychiatry and Biobehavioral Sciences
760 Westwood Plaza
Los Angeles, CA 90095
USA
hjerison@ucla.edu

Gestures Enhance Foreign Language Learning

Manuela Macedonia & Katharina von Kriegstein

Language and gesture are highly interdependent systems that reciprocally influence each other. For example, performing a gesture when learning a word or a phrase enhances its retrieval compared to pure verbal learning. Although the enhancing effects of co-speech gestures on memory are known to be robust, the underlying neural mechanisms are still unclear. Here, we summarize the results of behavioral and neuroscientific studies. They indicate that the neural representation of words consists of complex multimodal networks connecting perception and motor acts that occur during learning. In this context, gestures can reinforce the sensorimotor representation of a word or a phrase, making it resistant to decay. Also, gestures can favor embodiment of abstract words by creating it from scratch. Thus, we propose the use of gesture as a facilitating educational tool that integrates body and mind.

Keywords: education; embodiment; foreign language learning; gesture; memory

1. Introduction

When people speak, they spontaneously gesture. They do this to illustrate or to emphasize what they say (Hostetter 2011). When children acquire language, they also gesture. In particular, pointing has been described as a precursor of spoken language (Goldin-Meadow 2007; Tomasello *et al*. 2007). People trying to express themselves in a foreign language make use of gestures. The gestures help to convey meaning and to compensate for speech difficulties (Goldin-Meadow 2003; Gullberg 2008). Learners of a foreign language also express their provenience in intercultural settings through the gestures they use (Gullberg & McCafferty 2008; McCafferty 2008; McCafferty & Stam 2008). Foreign language teachers use gestures as a tool which favors and enhances the language acquisition process (for reviews, see Kusanagi 2005; Taleghani-Nikazm 2008).

However, gestures can do even more: If they are performed during learning of words and phrases, they enhance memory compared to pure verbal encoding (Zimmer 2001a). Also, gestures accompanying foreign language items enhance their memorability (Quinn-Allen 1995; Macedonia 2003; Tellier 2008) and delay their forgetting. Why this happens is the question we will discuss in this paper.

We thank Bob Bach for comments on previous versions of the manuscript. Contract grant sponsor: Cogito-Foundation, Wollerau, Switzerland.

ISSN 1450–341
Biolinguistics 6.3–4: 393–416, 2012
http://www.biolinguistics.eu

2. The Effect of Gestures and Verbal Memory: A Brief Historical Overview

Over the past three decades, laboratory research has shown that action words or phrases such as *cut the bread* are memorized better if learners perform or pantomime the action during learning than if they only hear and/or read the words (Engelkamp & Krumnacker 1980; Saltz & Donnenwerthnolan 1981). Different research groups working on this topic gave the effect of gestures on verbal information different names: 'enactment effect' (Engelkamp & Krumnacker 1980) and 'self-performed task-effect' (Cohen 1981). Many experiments using various materials (verbs, phrases, actions with real objects, common, and bizarre actions), tests (recognition, free, and cued recall) and populations (children, students, elderly subjects, people with memory impairments have independently replicated this effect, for reviews, see Engelkamp 1998; Nilsson 2000; Zimmer 2001b). Interestingly, not only healthy subjects showed a benefit in retrieval of enacted information (Rusted 2003); likewise, mentally impaired subjects (Cohen & Bean 1983) and patients suffering from memory impairments such as mild to moderate dementia (Hutton *et al.* 1996) profited. Also, it was demonstrated that during stroke rehabilitation, patients can enhance their memory performance through enactment (Nadar & McDowd 2008). More recent studies with children have also reported positive effects on learning for action/object phrases (Mecklenbräuker *et al.* 2011).

Besides enhancing the quantity of memorized items and prolonging their longevity, enactment also improves the accessibility of the learned words. In free recall tests, Zimmer *et al.* (2000) observed that enacted items pop out of the mind effortlessly. In recognition tasks, reaction time is better for enacted encoding (Masumoto *et al.* 2006) and this occurs independently of the subjects' age (Freeman & Ellis 2003). Also, recent experiments have demonstrated better accessibility of enacted action phrases through immediate and delayed free recall tests on younger and older adults (Spranger *et al.* 2008; Schatz *et al.* 2011). Overall, compared to pure verbal learning, enactment has proven to be more effective in enhancing verbal memory.

3. The Body as a Learning Tool in Foreign Language Instruction

There have been attempts to integrate the body as a learning device in foreign language learning. The first was by Asher in the late 60s. His teaching method, the Total Physical Response (TPR), required students to respond with actions to commands that were given as imperative sentences by the teacher (Asher & Price 1967). TPR was intended to support not only the understanding, but also the memorizing, of vocabulary items that can be learned through imperatives. Also, Asher pointed out that focusing on listening and action performance and not on language production corresponds to the natural sequence of native language acquisition (Asher 1977). Krashen & Terrell, well known among language teachers for their influential Natural Approach (Krashen & Terrell 1983), supported TPR as a learning technique for beginners because it is capable of involving learners in realistic language activities. However, despite its potential, TPR did not succeed in developing into an everyday learning tool for second language instruction. There are at least two reasons for this. First, Asher did not conduct empirical studies: He could not demonstrate that action has a greater impact on the acquisition of verbal information compared to audiovisual strategies. Second, when Asher developed his TPR, theories based on a universal grammar (Chomsky 1959) considered language learning to be an innate process (Fodor *et al*. 1974; Chomsky 1975). Accordingly, like mother tongue acquisition, foreign language was thought to emerge by mere listening and without tools of instruction because it results from innate processes (Feyten 1991; Krashen 2000). Explicit explanation and vocabulary teaching by any means, and therefore, also by action, were considered superfluous. Although there were other opinions in the field sustaining that child language acquisition and adult foreign language learning are fundamentally different (Bley-Vroman 1990), the mainstream followed the mentalistic view of a core grammar present in the learners' minds. This view implicitly ruled out the body as a possible learning device, as suggested by Asher.

The TPR used action as a teaching instrument. Note that action and gestures are not equal (Kendon 1981; McNeill 1992). In order to enact the word *to drink*, one can perform the action of drinking and drink some liquid. However, the gesture related to this word can also be simulation of drinking without glass and without liquid. Also, the word *to drink* can be illustrated by shaping a 'c' with a hand and raising it toward the mouth. In foreign language lessons, both can occur: action and gestures are used.

In the eighties and nineties, gestures came into play in foreign language instruction embedded in a broader framework of lessons involving drama (Mariani 1981; Schewe & Shaw 1993). Carels (1981) proposed the systematic use of pantomimic gestures in foreign language learning. Importantly, he suggested that these gestures should not only be performed by the teacher, but also by the learner, as a memory supporting strategy. He illustrates a two-step procedure. First, the teacher narrates the text and pantomimes vocabulary items that are unknown or difficult to understand. Thereafter, learners repeat the text and the pantomimes, in order to consolidate the acquisition of the novel words. Macedonia (1996) adopted a similar approach and described the use of iconic, metaphoric and deictic gestures in Italian lessons for German speaking university students. Particularly, she observed the beneficial effects of gestures on memory. However, these papers were merely descriptive and lacked empirical evidence for the use of gestures.

4. The Effects of Gestures on Memory for Foreign Language Words and Expressions

The first systematic study on the impact of gestures on memory for verbal information in a foreign language was conducted by Quinn-Allen (1995). She taught English-speaking students 10 French expressions (e.g., *Veux-tu quelque chose à boire?* 'Do you want something to drink?') by accompanying the expressions with illustrative, semantically related gestures typical of French culture.

For example, the gesture paired with the above sentence was performed by pointing the thumb toward the open mouth. The study showed that better results in retrieval were achieved over the short- and long-term, i.e. immediately after learning and after 11 weeks, if learners had performed the gestures when encoding the expressions.

In a 14-month longitudinal study, Macedonia (2003) worked on single word retention. She demonstrated that verbal items belonging to different word categories benefited from gesture use during learning. She trained university students to learn 36 words (9 nouns, 9 adjectives, 9 verbs and 9 prepositions) in an artificial language corpus. For 18 items, participants only listened to the word and read it. For another 18 items, participants were additionally instructed to perform the gestures proposed by the experimenter. Retrieval was assessed through cued recall tests at five different time points. The results showed significantly better retrieval in the short- and long-term for the enacted items.

In a study with 20 French children (average age 5.5) learning English, Tellier (2008) presented 8 common words (*house*, *swim*, *cry*, *snake*, *book*, *rabbit*, *scissors*, and *finger*). Four items were associated with a picture and four items were illustrated by a gesture that the children saw in a video and they thereafter performed. Enacted items were better memorized than items enriched visually by the pictures.

Kelly *et al*. (2009) trained 28 young adults on 12 Japanese verbs conveying common everyday meanings. The words were presented according to four modes: (i) speech, (ii) speech + congruent gesture, (iii) speech + incongruent gesture, and (iv) repeated speech. The results showed that participants memorized the largest number of words in the speech + congruent gesture mode, followed by the repeated speech mode, and the least number of words was memorized when they were accompanied by an incongruent gesture.

Another study by Macedonia & Knösche (2011) investigated the impact of enactment on abstract word learning. The words were learned while embedded in 32 sentences, each comprising 4 grammatical elements: subject, verb, adverb, and object. Only the nouns for the subjects were assigned concrete meanings. They indicated the actors. The remaining words were abstract. Twenty subjects participated in the study and learned according to two conditions. Words were either memorized audio-visually or enriched through a gesture. Gestures illustrating abstract words were arbitrary and had a symbolic value. Free recall and cued recall tests assessed the participants' memory performance at six time points. The overall results indicate that enactment, as a complement to audiovisual encoding, enhances memory performance not only for concrete but for also for abstract words (nouns, verbs, and adverbs). Moreover, in a transfer test, participants were asked to produce new (non-canonical) sentences with the words they had learned during training. Enacted items were recruited significantly more often than words learned audio-visually.

A study controlling for the type of gestures was conducted by Macedonia *et al*. (2011). They used a set of iconic gestures (i.e. creating a motor image and a set of meaningless gestures) providing mere sensorimotor input. Thirty-three German-speaking subjects were trained on 92 concrete nouns in a novel artificial corpus created for experimental purposes and based on Italian phonotactics.

Half of the items were encoded with iconic gestures (McNeill 1992). They depicted some aspect of each word's semantics and enriched the word with a plausible sensorimotor connotation. The other half of the items was learned with meaningless gestures. They could be small (shrugging one's shoulders) or large (stretching one's arms in front of oneself). They were randomly presented when the subjects read and heard the word and they changed at every trial. The results showed significantly better memory performance for iconic gestures than for meaningless gestures in the short- and long-term (after 60 days), indicating that enhancement does not come from pure physical activity complementing the verbal information.

The results of these studies suggest that performing a gesture when learning a novel word in a foreign language or in an artificial corpus significantly enhances the word's retrieval and delays forgetting compared to pure verbal learning. Moreover, there is evidence that gestures representing the word's semantics, or some feature of it, help to memorize better the word than meaningless gestures do.

5. Possible Mechanisms Underlying the Effects of Gestures on Verbal Memory

In the debate on the mechanisms underlying enactment, four main positions have emerged. The first position emphasizes the crucial role of the overt action performed by the learner (Engelkamp & Krumnacker 1980; Engelkamp & Zimmer 1985). According to this view, the physical enactment creates a motor trace in the memory representation of the verbal item. The second position assumes that doing things in a wider perspective (i.e. cognitive activities like spelling the word) can lead to better verbal memory (Cohen 1981, 1985). In the third position, imagery (i.e. a kinetic representation of the word's semantics created through action) is the factor leading to improved performance (Saltz & Donnenwerth-nolan 1981). According to the fourth position, the impact on memory is caused primarily by increased perceptual and attentional processes occurring during proprioception and/or when using objects to perform the action (Bäckman *et al.* 1986). Thus memory enhancement does not come from enactment itself, as the motor component is not crucial (Kormi-Nouri 1995, 2001). Rather, it is the multisensory information conveyed into a word that leads to deeper semantic processing and higher attention level (Knopf 1992; Knopf *et al.* 2005; Knudsen 2007).

Studies dealing with the beneficial use of gestures in foreign language learning explain memory enhancement in terms of depth of encoding. Quinn-Allen (1995) observed that gestures provide an elaborated context for language; this enables deep processing of the verbal items and thus durability of the information (Craik & Tulving 1975). In her study, Macedonia (2003) proposes the Connectivity Model of Semantic Processing (Klimesch 1994) to account for the high memorability of novel words learned with gestures. Accordingly, a complex code involving sensory and motor information is deep and so improves retrievability and resistance to decay. Tellier (2008) also addresses the question in terms of the depth of encoding due to multimodality; she refers to Paivio's

Dual Coding Theory (Paivio 1969, 1971; Paivio & Csapo 1969) and to a possible motor trace left by the gesture.

Kelly *et al*. (2009) argued that gesture helps to deepen the motor image and thus the memory trace of a novel word. Moreover, they theorize that gestures can convey non-arbitrary meaning that is grounded in our bodies, since speech and gesture are strongly interconnected systems. In the study by Kelly and colleagues, within the discussion of why gesture helps to better memorize foreign language words, the scientists overtly address the body as a tool capable of supporting memory processes.

In their study of learning words paired with meaningful iconic and meaningless gestures, Macedonia *et al*. (2011) find empirical evidence for the existence of both a motor trace and a sensory motor image connected with a novel word in a foreign language. More recently, Macedonia & Knösche (2011), investigated the impact of gestures on memory performance for abstract words learned in the context of sentences and proposed that performing a gesture when learning a word can fulfill two functions. First, it strengthens the connections to embodied features of the word that are contained in its semantic core representation. Second, in the case of abstract words such as adverbs, gesture constructs an arbitrary motor image from scratch that grounds abstract meaning in the learner's body.

With their variations in experimental design, the different studies have shed light on the manifold aspects of enactment. The positions above are not mutually exclusive. Gestures paired with novel words in a foreign language enhance attention compared to learning the words in less complex contexts such as bilingual lists. Also, words enriched with gestures are complex deep codes and therefore better retained than shallow codes (Wig *et al*. 2004). However, the question of whether enactment favors the retention of verbal information because of a motor representation or due to imagery processes could only be elucidated by neuroscientific experiments. In the next section, we will review research on the topic published in the last 30 years.

6. Sensorimotor Representation of Gesture in the Brain

The question of whether a motor trace is left as the representation of an enacted word (Engelkamp & Krumnacker 1980) has been investigated by using different neuroscientific methods. In an event-related potentials (ERP) study, Heil *et al*. (1999) trained participants to passively listen to or to perform accompanying actions to phrases with imaginary objects. On testing, participants' recognition of the enacted phrases scored better, and during recognition a larger fronto-central negativity was detected. The authors interpreted these results as indicating information processing in the motor cortices.

In a Positron Emission Tomography (PET) study, Nilsson *et al*. (2000) also tested the hypothesis that enacted items show more activity in motor cortices during retrieval compared to verbal encoding. They trained participants in three learning conditions. During verbal training, participants simply rehearsed the command. During enactment training, participants overtly performed the actions described by the commands. During imagery training, subjects were cued to imagine performing the described actions. The results showed that enactment significantly increased activity in the right primary motor cortex compared to verbal training. Interestingly, activity of the right motor cortex was also observed during verbal and imagery training.

Another PET study by Nyberg *et al*. (2001) examined brain activity in the motor cortices for verbally encoded, overtly enacted and covertly encoded items. Activity registered in motor and somatosensory areas during retrieval was common to enactment and covert encoding. These results provide evidence that both performing an action and imagining performing it recruit the same neural substrate.

In an experiment by Masumoto *et al*. (2006), participants learned action sentences according to three conditions: by enactment, by observation of an agent enacting them, and by observation of an object mentioned in the action sentences. After encoding, participants performed a recognition test, during which magnetoencephalography data were acquired. The experiment tested the hypothesis that enacted action elicited activity in the motor cortex. Interestingly, only the left primary motor cortex was statistically relevant (participants were all right-handed).

In order to clarify whether action itself (i.e. independently of its shape) works as a learning enhancer, Macedonia *et al*. (2011) conducted a study in which participants were cued to learn concrete substantives by accompanying them with either iconic or meaningless gestures. In the fMRI-scanner, participants performed an audiovisual recognition task of the words they had trained. In the contrast meaningless gestures versus iconic gestures, the latter produced activity in the dorsal part of the premotor cortex. This localization within the motor cortices was interpreted as being due to the fact that action performed during the training mainly involved distal movement. The dimension of activation in the left precentral gyrus was larger than in the right hemisphere (the iconic gestures were performed by right-handed subjects with their dominant limbs). However, the region of interest analysis of the premotor cortex demonstrated that recognizing words encoded through meaningless gestures also activated premotor cortices. Thus, verbal material paired with action during learning leaves a motor trace independently of the kind of gestures used and independently of the impact that the gestures have on memory.

7. Words are Connected to Images

More than three decades ago, Engelkamp & Krumnacker (1980) reasoned that the gesture accompanying a word is connected with an existing image of its semantics. Saltz & Donnenwerthnolan (1981) proposed that enactment is effective because it leads to the storage of a 'motoric image'. Recent neuroscientific research has helped to clarify the link between motor imagery and language. Experiments investigating spontaneous co-speech gestures and their representation in the brain have shown different time courses and brain activity patterns if speech is accompanied by matching or non matching gestures.

In an ERP study examining the impact of representative gestures accompanying speech, Kelly *et al*. (2004) showed participants videos of an actor speaking and gesturing. When talking, the actor produced gestures for the words tall, thin, short and wide in reference to objects present in the videos. Participants had to decide whether speech and gesture were congruent. Mismatching stimuli produced a larger right-lateralized N400, an indicator for semantic integration (Kutas & Hillyard 1980).

The sensitivity to semantic relations between gestures and words was similarly demonstrated in a priming experiment by Wu and Coulson (2007a, b). Participants had to judge whether the presented gesture-speech utterance followed by a related picture was either related to speech alone or to both speech and gesture. Here, again, the N400 component was smaller when the pictures were related to speech and gesture.

Over the years, the tight integration of speech and gesture has been documented in a number of ERP studies (Holle & Gunter 2007; Ozyurek *et al*. 2007; Bernardis *et al*. 2008). The results of these studies suggest that the link between speech and gesture is immediate and not modulated by attentional processes. Modulation by attention was recently investigated in a stroop task experiment (Kelly *et al*. 2010). Participants had to decide whether the gender of the speaker corresponded to the gender of the speaking person gesturing in a video. Even if the task to be performed was not to detect the (mis)match between gesture and language, when speech and gesture were incongruent, a larger N400 was produced and reaction times for the task to be accomplished were slower. Also, another ERP component, the P600, also called Late Positive Complex (LPC), peaking at about 600ms after stimulus onset, was observed as a component indexing the recognition of imageable words.

In their study, Klaver *et al*. (2005) presented subjects words of high and low imageability that had been previously controlled for word frequency. Behaviorally, subjects recognized concrete words better. In the ERP experiment, the main effect of imageability was indexed by a hippocampal P600. This correlate was interpreted as involvement of the hippocampus during processing of verbal information with high imageability. Other studies describe the P600 as a correlate associated with recollection of verbal information that is concrete (Scott 2004) and has high imageability (Rugg & Nagy 1989).

More recently, a study comparing timing and topographical distribution of ERP components when subject processed concrete vs. abstract words detected activity in visual association areas (BA 18 and 19) for abstract words (Adorni & Proverbio 2012).

Also, functional magnetic resonance imaging (fMRI) experiments have evidenced the existence of motor images related to verbal information. In a study by Willems *et al*. (2007) investigating the neural integration of speech and action, the authors used a mismatch paradigm. Participants were presented with sentences followed by iconic gestures that either matched or mismatched the preceding context. The conflict between language and gesture produced enhanced activity in the left inferior frontal cortex, the premotor cortex, and the left superior temporal sulcus. This activity was interpreted as an increase in the semantic load resulting from conflicting speech and action.

In a disambiguation paradigm, Holle *et al*. (2008) showed participants videos of a speaker uttering sentences (she touched the mouse) with an ambiguous word (mouse). The ambiguous part of the sentence was accompanied by either an iconic or a meaningless gesture. During sentence presentation, fMRI data were collected. Compared to meaningless gestures, the processing of iconic gestures revealed hemodynamic activity in the left posterior superior temporal sulcus (STS), in the inferior parietal lobules and in both ventral precentral sulci. Of particular interest is the response of the posterior STS. This cortical area is known to become active during multisensorial integration or when integration does not match expectations (Beauchamp 2005). The authors of the study attribute activation of the STS to the lack of meaning in the meaningless gestures.

In an experiment by Green *et al*. (2009), German speaking subjects were presented with short videos of an actor performing gestures and sentences while their brain activity was measured by means of fMRI. The accompanying gestures were either related or unrelated to the sentences, which were in German (familiar to the participants) or in Russian (unfamiliar). While speech accompanied by iconic gestures activated left occipital areas, speech with mismatching gestures engaged bilateral parietal and posterior temporal regions.

In another fMRI study, Straube *et al*. (2009) investigated memory for speech and gesture representations. Participants were presented with abstract sentences accompanied by video clips where an actor produced either meaningful metaphoric gestures, unrelated free gestures, or no gestures. After the training, participants were administered a recognition test. They performed better for sentences accompanied by meaningful metaphoric gesture. The results of the fMRI data analysis for the metaphoric gesture mode showed left-hemispheric activations in the inferior frontal gyrus, the premotor cortex, and the middle temporal gyrus. This left-lateralized activation pattern was interpreted by the authors as an indicator of semantic integration between speech and gestures. Interestingly, the metaphoric gesture mode showed significant correlations between memory performance and activity in the hippocampus. Several other studies have concentrated on the loci for integration between gesture and language (for reviews, see Willems & Hagoort 2007 and Willems *et al*. 2009). They indicated the posterior superior temporal sulcus, the middle temporal gyrus and the left inferior frontal gyrus as areas integrating information from different modalities.

Similarly to neurophysiological studies documenting larger N400 components, brain imaging experiments have revealed brain networks denoting disturbance and integration effort if words and gestures are incongruent. Although neuroscientific research up to now has mainly focused on language comprehension and not on memory effects of motor imagery, it has provided converging evidence that words do have a corresponding (motor) image in their semantic representation.

In foreign language learning, we found two experiments showing disturbance effects when words and gestures do not match. In an ERP experiment aiming to explore whether gestures create a deeper imagistic representation of words in memory, Kelly *et al*. (2009) trained participants on a Japanese word list comprising twelve common verbs, such as *to drink*. The verbs were learned with or without iconic hand gestures. The results demonstrated that words encoded with gestures were better memorized. Event-related potentials of words learned with gestures compared with words learned without gestures showed a larger LPC bilaterally, denoting recollection with high imageability.

In their study on foreign language learning, Macedonia *et al*. (2011) employed a set of iconic and a set of meaningless gestures. They were paired with the words to be learned. During the recognition phase in the fMRI experiment, words learned with iconic gestures activated premotor cortices, as described above, while meaningless gestures elicited activity in a vast brain network in both hemispheres comprising the cunei, the left posterior cingulate gyrus, the right anterior cingulate gyrus, the left inferior frontal junction area and the right rostrolateral prefrontal cortex. These regions orchestrate a network for cognitive control that possibly denotes conflict detection and effort to integrate information (Cole & Schneider 2007).

Although speculative, we reason that a single concept might comprise multiple images. They vary depending on the factors experienced by the subject. Motor images of a word like *car* represent the possible motions of cars but also pantomimes performed by a person producing some characteristics of a car (e.g., the shape). Thus, in the stroop-like experiments reviewed above, the mismatch possibly occurs between the internal image (i.e. the neural pattern created through learning), and the pattern of activity elicited through the perception of the presented stimulus.

8. Neural Representation of Words

Early theories of cognition considered concepts as amodal, symbolic entities (Fodor 1976, 1983; J.D. Fodor 1977), their meaning being referential and *somehow* connected to objects. As Meteyard *et al*. (2010) point out in their review, amodular theories of cognition had an Achilles' heel: The representation of how symbols refer to real things. The problem, although overtly recognized (Pylyshyn 1984; Fodor 1987), was never solved. The focus of these theories mainly resided in the structure of processes rather than in the content of symbols. However, without grounding (i.e. linking a symbol causally to its reference), it is hard to conceive how meaning could be established.

In the past decade, symbolic theories have been challenged by the advent of brain imaging techniques. The fact that merely listening to words like *kick, lick, or pick* (Pulvermüller 2005) or phrases like *press the piano pedal, bite the banana,* or *grasp the pen* (Aziz-Zadeh *et al*. 2006) activates brain motor areas controlling movements, respectively performed by leg, mouth, and hand, could not be accounted for in terms of symbolic theories on cognition (Simon 1981; Fodor 1983; Pylyshyn 1984). Similarly, if listening to words like *cinnamon* or *garlic* elicits activity in olfactory brain regions (González *et al*. 2006), even in the absence of real objects, it becomes clear that concepts, here expressed as words, are not amodal. Consequently, the word *garlic* must be tightly linked with sensory perception (i.e. with smell, taste, texture, color, etc). Also, action (e.g., peeling or mincing garlic, rubbing it on bread, and chewing it) can be part of the representation of the word. Thus, sensory and motor information related to *garlic* that has a representational role, constitutes the word's semantics (Gallese & Lakoff 2005). A word is not an abstract entity with a reference in the world; rather, a word is grounded in the perception and action a subject experiences (Kiefer & Spitzer 2001; Vigliocco *et al*. 2004; Barsalou 2008; Pulvermüller & Fadiga 2010).

From a neurobiological perspective, a word can be described as a network linking cell assemblies that code and process linguistic, sensory, motor (Pulvermüller 1996, 2001, 2002) and emotional (Vigliocco *et al*. 2009) features. In this view, words are represented in distributed networks with different topographies, including perisylvian areas and areas critically involved in processing perception and action. The extension and shape of networks change over time depending on the interaction the subject has with the world. During brain imaging experiments, stimulation activates cell assemblies processing stimuli in crucial cortical areas that are specialized for the task. However, as activity spreads within the network (McClelland & Rogers 2003), activity reaches assemblies that code features bound to the concept. This explains why simply hearing *garlic* activates not only auditory cortices but also olfactory areas. Similarly, hearing a sound related to finger actions elicits activity in motor areas associated with the hand (Hauk *et al*. 2006). Thus, there is clear evidence for a complex neural representation of a word that comprises sensorimotor components linked during learning.

9. Learning Words in a Foreign Language through Gestures

To our knowledge, there are no studies documenting the processes of acquisition of a novel word in a foreign language in terms of a functional network. However, on the basis of the literature reviewed, we reason that when people learn a novel word by merely listening to it or reading it, the neural representation of the novel word will be poor compared to a word in the native language. For the novel phoneme chain, most of the sensorimotor and emotional experience embodied in the corresponding word in the native language is lacking or, at best, only partially present. Pulvermüller (2002) proposes that there is activity in response to a novel word co-occurring with a known word in perisylvian regions, with extensions to extra-perisylvian cortex areas that code semantic features of the known word. We speculate that this could happen when learners acquire the word in the foreign language by listening to it and knowing its translation in the native language. Through frequent repetition, cell assemblies coding the novel phoneme chain would be active together with the sensorimotor network of the word in the native language. Finally, this correlated activity would strengthen the synaptic connections (Hebb 1949) between both words, and lead to the creation a larger network integrating the novel word within the semantic representation of the word in the mother tongue. However, this integration would be indirect as long as learners have not gain experience using the word and thus have grounded it in their body. From a neurobiological perspective, these networks represent the substrate of memory (Garagnani *et al*. 2007, Wennekers & Ay 2003). Thus, it is conceivable that if the integration is not driven by action and strong sensorimotor experience, the connections to the network in the native language are poor; hence memory is unstable.

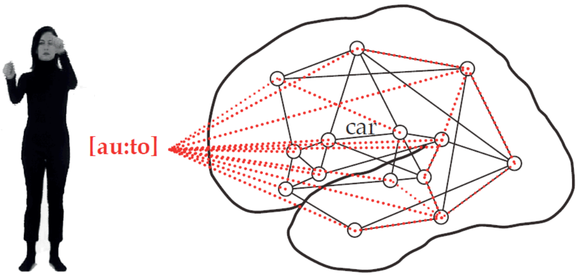


Figure 1: Schematic illustration of a possible word network representing the word car in the learners´ native language and the corresponding word in German Auto after acoustic learning. Note that the dotted lines represent indirect connections with the sensorimotor network of the native language (adapted from Pulvermüller 2002)

When learning a novel word by enacting it, the learner has a complex multimodal sensorimotor experience. It is conceivable that, depending on the kind of words and relative gestures, the process of embodiment is different. Gestures for action words like *to go* or *to give* reproduce the action itself. In this case, the novel phoneme chain possibly docks on networks representing the action itself. Thus, enactment reproduces and reinforces sensorimotor patterns created during native language acquisition. This might explain the strong effect of enactment on memory, especially for action words and phrases (Zimmer 2001b).

Gestures accompanying concrete words tend to be mostly iconic (Macedonia 2003). An iconic gesture might match internal (motor) images of the concept and create a strong connection to the novel word with a preexisting circuit that represents the concept.

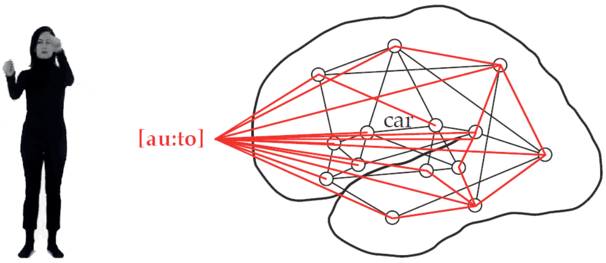


Figure 2: Schematic illustration of a possible word network representing the word car in the learners´ native language and the corresponding word in German Auto after enactment. The lines represent strong connections with the sensorimotor network of the native language (adapted from Pulvermüller 2002)

Also, the iconic gesture might support the so called ‘concreteness effect’. This effect, fully demonstrated in the literature, mirrors the easier processability (Binder *et al*. 2005) and better memorability of concrete words (Allen and Hulme 2006; Fliessbach *et al*. 2006; Romani *et al*. 2008). The concreteness effect has been accounted for in terms of the Dual Coding Theory (Paivio 2006, 1971, 1986, 1991) and the Context Availability Model (Schwanenflugel *et al*. 1988). According to Paivio, words can have two modalities of representations: one is purely verbal and the other imagistic. Whereas abstract concepts lack the imagistic component and are primarily verbally represented, concrete words are represented in both modalities. This explains their advantage in processing and retrieval. The Context Availability Model postulates only one modality of representation for concepts. However, concrete words are better processed and retrieved because of the dense associative context in which they are embedded. Both proposals (i.e. the Dual Coding Theory and the Context Availability Model) put forward the idea that the cognitive advantage is grounded in richer representation of the word´s semantics. This view is also accounted for in neurophysiological research. ERP experiments have, in fact, demonstrated that concrete words elicit a larger N400 than abstract words (West & Holcomb 2000; Levy-Drori & Henik 2006). This ERP component has been reconducted to the activation of more extended, and hence, richer semantic networks.

However, abstract words also benefit from enactment learning. In their experiment, Macedonia & Knösche (2011) cued participants to perform arbitrary gestures accompanying abstract words. For abstract nouns they represented and embodied a motor image of the word, connected somehow with the word's semantics (i.e. in a more remote way than for concrete words). For instance, for the Vimmi word *sigule* 'theory', the actress in the video simulated the opening of a book in front of her in an interested way. Here the gesture might have made the concept more concrete, and hence, have taken advantage of the concreteness effect.

For other abstract nouns, gestures addressed some emotional component present in the neural representation of the word. For instance, for the word *boruda* 'sensation' the actress performed a gesture of astonishment. Her arms and mouth were wide open. According to Kousta *et al.* (2011) abstract words differ from concrete words in terms of embodied experiential information. Whereas for concrete words sensory-motor information is preponderant in their representation, abstract words statistically contain more emotional information. Enacting a word through a gesture expressing emotion possibly reinforces the emotional content and enhances memory for the item. However, how is emotional content to be understood in terms of brain circuits? 'Canonically', emotion is processed in limbic areas. Is this always the case? In an fMRI study by Moseley *et al.* (2012) when participants passively read highly abstract emotion words, not only language areas (Broca's region, Wernicke's region, and fusiform gyrus) and limbic structures became active, but also their premotor cortices. Particularly, inferior and dorsolateral motor areas processing face and arm related movements, respectively, were involved. The authors explain the data in terms of semantic networks representing not only the intrinsic emotional content of the words, but also motor programs used to express the emotions. It seems plausible that at least facial actions are a part of a circuit for an emotional word. Thus, if a 'meaningful' facial gesture accompanying an abstract concept represents some emotional component of its semantics, it could reinforce its embodied representation and therefore support memory.

Adverbs, another category of abstract words, also are better memorized if encoded with a gesture. In this case, the gestures used are arbitrary with no representative value. Adverbs like *already* or *still* serving primarily a grammatical purpose are difficult to relate to an image and their emotional content is poor. Thus, their representation in term of functional networks is hardly connected with sensorimotor information per se. This might also represent a reason for their low memorability. For function words, Pulvermüller (1999) proposed a more localized topography restricted to perisylvian areas. Gestures accompanying adverbs are thus arbitrary and have no semantic relationship with the words' semantics. These gestures thus create a motor image from scratch. They thus enrich the original representation of the word with a sensorimotor component not present before. This might explain the better retention achieved when adverbs are paired to arbitrary gestures during learning.

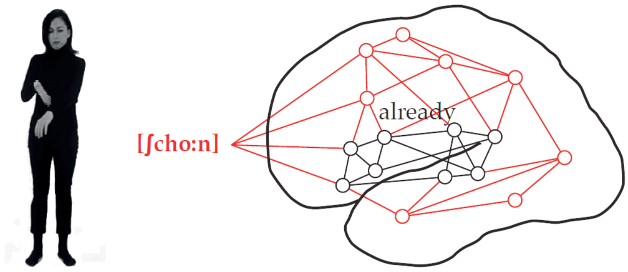


Figure 3: Schematic illustration of a possible word network representing the word already in the learners´ native language and the corresponding word in German schon after learning through enactment. Note that the original network in perisylvian areas is enriched with sensorimotor components provided by the arbitrary gesture (adapted from Pulvermüller 2002)

10. Why Gestures Enhance Memory

We have described how gestures may enrich the representation of novel phoneme chains of a foreign language and ground them in the body. We believe that learning words in a foreign language together with motor and multimodal information helps to create similar networks for the foreign language word.

Studies of foreign language word learning address the depth of processing as a factor that enhances learning (Macedonia 2003; Tellier 2008; Kelly *et al*. 2009). Pairing a gesture to a novel word makes the network more complex (i.e. deeper) by binding sensorimotor information to verbal information. According to Klimesch's Connectivist Theory on the structure of long-term memory (Klimesch 1994), it is the complexity of a code that leads to its better retrieval in memory. Thus, the factor enhancing memory when words are accompanied by gestures could be complexity. However, it should be investigated if, within the network, there are components driving the memory performance more than others. In fact, it is possible that motor information functions like a supramodal device with hub characteristics (Tomasi & Volkow 2011). As such, motor information could have hierarchical properties and affect memory more than other sensory components within the network.

Besides the neurobiological view of memory and on how words are grounded in the body, there are at least two issues that might play a role for enhanced verbal learning through gestures. The first is the synchronicity between word and gesture production. An fMRI study by Xua *et al.* (2009) has demonstrated that symbolic gestures and spoken language are processed by a common neural system mainly localized in the left hemisphere, in anterior and posterior perisylvian regions. It is possible that performing both language and gestures together when learning novel words, boosts the language system and stimulate memory structures.

Moreover, from an evolutionary perspective, language and action are tightly connected. According to a number of authors (Rizzolatti & Arbib 1998; Gentilucci & Corballis 2006; Tomasello 2008; Arbib 2009), language evolved from signs that our ancestors used in combination with vocalizations. Thus, gestures have scaffolded the emergence of a protolanguage. Because of the vast range of phenomena that have been demonstrated in neuroscientific research, particularly the mismatch effects appearing if language and gesture are incongruent, the evolutionary view has gained strong plausibility over the years. Hence, by accompanying novel words with gestures, learners assemble the two parts of an ancient communicative system. This might be beneficial for memory processes.

Second, imitation is another important issue connected to the use of gestures during encoding of foreign languages. In fact, if learners are instructed to perform a gesture they are presented with mechanisms of imitation and thus mirror neuron circuits might become active and enhance learning (Vogt *et al.* 2007; Vogt & Thomaschke 2007; Mukamel *et al.* 2010).

11. Implications for Second Language Instruction

When learning a foreign language, students usually read or listen to the verbal information they want to acquire. Traditional instruction makes wide use of listening and comprehension activities (Winitz 1981; Swain & Lapkin 1995). As homework, learners go through bilingual vocabulary lists and learn the words by reading them. Foreign language instruction is far from reconstructing the experiences we have when acquiring our native language. In fact, children make sensorimotor experiences by interacting intensively with their caregivers and their environment (Tomasello 2005; Kuhl 2010). Thus, it is no surprise that the outcome of the two learning processes is different with respect to memory. While under normal conditions it is unlikely that people forget words of their native language, adults learning a foreign idiom are plagued by forgetting what they have previously learned.

It has been demonstrated that multimodal learning helps to better memorize information (Shimojo & Shams 2001; von Kriegstein & Giraud 2006; Shams & Seitz 2008; von Kriegstein *et al.* 2008; Shams *et al.* 2011) and efforts have been made in foreign language teaching practice to enrich vocabulary with multisensory input by using flash cards (Barcroft 2009; Boers *et al.* 2009; Tonzar *et al.* 2009), videos (Sydorenko 2010), songs (Keskin 2011), and implementing them on novel technical devices such as mobile phones (Başoğlu & Akdemir 2010). However, a view linking the body and mind, considering the body in action as a learning tool, is still missing in foreign language instruction. Hence, we propose the use of gestures as a learning device that grounds foreign language in the body and thereby enhances memory.

This paper focuses on the impact of gestures memory for lexical items. In fact, the acquisition of lexical items is basic to language learning at any level. However, it is conceivable that gestures can also help to acquire morphological (Goldin-Meadow *et al*. 1995) and syntactic structures. In a German publication, Macedonia (1999) addressed both aspects when describing teaching practice in foreign language with gestures. Observations from classroom activities encourage the use of gestures for complex verbal morphology in Romance languages and for different kinds of combined clauses in syntactic contexts in Italian. Nevertheless, controlled laboratory research is lacking and is needed in order to collect empirical evidence for the use of gestures in these language domains.

References

Adorni, Roberta & Alice Mado Proverbio. 2012. The neural manifestation of the word concreteness effect: An electrical neuroimaging study. *Neuropsychologia* 50, 880–891.

Allen, R. & C. Hulme. 2006. Speech and language processing mechanisms in verbal serial recall. *Journal of Memory and Language* 55, 64–88.

Arbib, Michael A. 2009. Evolving the language-ready brain and the social mechanisms that support language. *Journal of Communication Disorders* 42, 263–271.

Asher, J. J. & B. S. Price. 1967. The learning strategy of the total physical response: Some age differences. *Child Development* 38, 1219–1227.

Asher, James J. 1977. *Learning Another Language Through Actions: The Complete Teacher's Guidebook.* San Jose, CA: Pajaro Press.

Aziz-Zadeh, L., S. M. Wilson, G. Rizzolatti & M. Iacoboni. 2006. Congruent embodied representations for visually presented actions and linguistic phrases describing actions. *Current Biology* 16, 1818–1823.

Bäckman, L., L. G. Nilsson & D. Chalom. 1986. New evidence on the nature of the encoding of action events. *Memory and Cognition* 14, 339–346.

Barcroft, J. 2009. Strategies and performance in intentional L2 vocabulary learning. *Language Awareness* 18, 74–89.

Barsalou, L. W. 2008. Grounded cognition. *Annual Review of Psychology* 59, 617–645.

Başoğlu, E. B. & O. Akdemir. 2010. A comparison of undergraduate students' English vocabulary learning: Using mobile phones and flash cards. *Turkish Online Journal of Educational Technology* 9, 1–7.

Beauchamp, M. S. 2005. See me, hear me, touch me: Multisensory integration in lateral occipital-temporal cortex. *Current Opinion in Neurobioly* 15, 145–153.

Bernardis, Paolo, Elena Salillas & Nicoletta Caramelli. 2008. Behavioural and neurophysiological evidence of semantic interaction between iconic gestures and words. *Cognitive Neuropsychology* 25, 1114–1128.

Binder, J. R., C. F. Westbury, K. A. McKiernan, E. T. Possing, & D. A Medler. 2005. Distinct brain systems for processing concrete and abstract concepts. *Journal of Cognitive Neuroscience* 17, 905–917.

Bley-Vroman, R. 1990. The logical problem of foreign language learning. *Linguistic Analysis* 20, 3–47.

Boers, F., A. M. P. Píriz, H. Stengers & J. Eyckmans. 2009. Does pictorial elucidation foster recollection of idioms? *Language Teaching Research* 13, 367–382.

Carels, P. E. 1981. Pantomime in the foreign-language classroom. *Foreign Language Annals* 14, 407–411.

Chomsky, Noam. 1959. A review of B. F. Skinner's *Verbal Behavior*. *Language* 35, 26–58.

Chomsky, Noam. 1975. *Reflections on Language*. New York: Pantheon Books.

Cohen, R. L. 1981. On the generality of some memory laws. *Scandinavian Journal of Psychology* 22, 267–281.

Cohen, R. L. & G. Bean. 1983. Memory in educable mentally retarded adults: Deficit in subject or experimenter? *Intelligence* 7, 287–298.

Cohen, R. L. 1985. On the generality of the laws of memory. In L. G. Archer & T. Nilsson (eds.), *Animal Learning and Human Memory*, 247–277. Hillsdale, NJ: Erlbaum.

Cole, Michael W. & Walter Schneider. 2007. The cognitive control network: Integrated cortical regions with dissociable functions. *NeuroImage* 37, 343–360.

Corballis, Michael C. 2009. The Evolution of Language. *Annals of the New York Academy of Sciences* 1156, 19–43.

Craik, F. I. & E. Tulving. 1975. Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General* 104, 268–294.

Engelkamp, J. & H. Krumnacker. 1980. Imaginale und motorische Prozesse beim Behalten verbalen Materials. *Zeitschrift für experimentelle und angewandte Psychologie*, 511–533.

Engelkamp, J. & H. D. Zimmer. 1985. Motor programs and their relation to semantic memory. *German Journal of Psychology*, 239–254.

Engelkamp, Johannes. 1998. *Memory for Actions*. Hove: Psychology Press.

Feyten, C. 1991. The power of listening ability: An overlooked dimension in language acquisition. *Modern Language Journal* 75, 173–180.

Fliessbach, K., S. Weis, P. Klaver, C. E. Elger & B. Weber. 2006. The effect of word concreteness on recognition memory. *NeuroImage* 32, 1413–1421.

Fodor, Janet Dean. 1977. *Semantics: Theories of Meaning in Generative Grammar*. Brighton: Harvester.

Fodor, Jerry A. 1976. *The Language of Thought*. Hassocks: Harvester Press.

Fodor, Jerry A. 1983. *The Modularity of Mind: An Essay on Faculty Psychology*. Cambridge, MA: MIT Press.

Fodor, Jerry A. 1987. *Psychosemantics: The problem of Meaning in the Philosophy of Mind*. Cambridge, MA: MIT Press.

Fodor, Jerry Alan, Thomas G. Bever & Merrill F. Garrett. 1974. *The Psychology of Language. An Introduction to Psycholinguistics and Generative Grammar.* New

York: McGraw-Hill Book Co.

Freeman, J. E. & J. A. Ellis. 2003. Aging and the accessibility of performed and to-be-performed actions. *Aging, Neuropsychology, and Cognition* 10, 298–309.

Gallese, V. & G. Lakoff. 2005. The Brain´s concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology* 22, 455–479.

Garagnani, M., T. Wennekers & F. Pulvermüller. 2007. A neuronal model of the language cortex. *Neurocomputing* 70, 1914–1919.

Gentilucci, M. & M. C. Corballis. 2006. From manual gesture to speech: A gradual transition. *Neuroscience and Biobehavioral Reviews* 30, 949–960.

Goldin-Meadow, S. 2007. Pointing sets the stage for learning language — And creating language. *Child Development* 78, 741–745.

Goldin-Meadow, Susan, Carolyn Mylander & Cynthia Butcher. 1995. The resilience of combinatorial structure at the word level: morphology in self-styled gesture systems. *Cognition* 56, 195–262.

Goldin-Meadow, Susan. 2003. *Hearing Gesture: How our Hands Help us Think*: Cambridge, MA: Harvard University Press.

González, Julio, Alfonso Barros-Loscertales, Friedemann Pulvermüller, Vanessa Meseguer, Ana Sanjuán, Vicente Belloch & César Ávila. 2006. Reading cinnamon activates olfactory brain regions. *NeuroImage* 32, 906–912.

Green, Antonia, Benjamin Straube, Susanne Weis, Andreas Jansen, Klaus Willmes, Kerstin Konrad & Tilo Kircher. 2009. Neural integration of iconic and unrelated coverbal gestures: A functional MRI study. *Human Brain Mapping* 30, 3309–3324.

Gullberg, M. 2008. Gestures and second language acquisition. In Peter Robinson (ed.), *Handbook of Cognitive Linguistics and Second Language Acquisition*, 276–305. London: Routledge.

Gullberg, M. & S. G. McCafferty. 2008. Introduction to gesture and SLA: Toward an integrated approach. *Studies in Second Language Acquisition* 30, 133–146.

Hauk, O., Y. Shtyrov & F. Pulvermüller. 2006. The sound of actions as reflected by mismatch negativity: Rapid activation of cortical sensory-motor networks by sounds associated with finger and tongue movements. *European Journal of Neuroscience* 23, 811–821.

Hebb, Donald Olding. 1949. *The Organization of Behavior. A Neuropsychological Theory*. New York: John Wiley & Sons.

Heil, M., B. Rolke, J. Engelkamp, F. Rosler, M. Ozcan & E. Enninghausen. 1999. Event-related brain potentials during recognition of ordinary and bizarre action phrases following verbal and subject-performed encoding conditions. *European Journal of Cognitive Psychology* 11, 261–280.

Holle, H. & T. C. Gunter. 2007. The role of iconic gestures in speech disambiguation: ERP evidence. *Journal of Cognitive Neuroscience* 19, 1175–1192.

Holle, Henning, Thomas C. Gunter, Shirley-Ann Rüschemeyer, Andreas Hennenlotter & Marco Iacoboni. 2008. Neural correlates of the processing of co-speech gestures. *NeuroImage* 39, 2010–2024.

Hostetter, Autumn B. 2011. When do gestures communicate? A meta-analysis. *Psychological Bulletin* 137, 297–315.

Hutton, S., L. Sheppard, J. M. Rusted & H. H. Ratner. 1996. Structuring the acquisition and retrieval environment to facilitate learning in individuals with dementia of the Alzheimer type. *Memory* 4, 113–130.

Kelly, S. D., C. Kravitz & M. Hopkins. 2004. Neural correlates of bimodal speech and gesture comprehension. *Brain and Language* 89, 253–260.

Kelly, Spencer D., Tara McDevitt & Megan Esch. 2009. Brief training with co-speech gesture lends a hand to word learning in a foreign language. *Language and Cognitive Processes* 24, 313–334.

Kelly, Spencer D., Peter Creigh & James Bartolotti. 2010. Integrating speech and iconic gestures in a stroop-like task: Evidence for automatic processing. *Journal of Cognitive Neuroscience* 22, 683–694.

Kendon, A. 1981. Gestures, their origins and distribution. *Semiotica* 37, 129–163.

Keskin, F. 2011. Using songs as audio materials in teaching Turkish as a foreign language. *Turkish Online Journal of Educational Technology* 10, 378–383.

Kiefer, M. & M. Spitzer. 2001. The limits of a distributed account of conceptual knowledge. *Trends in Cognitive Sciences* 5, 469–471.

Klaver, Peter, Jürgen Fell, Thomas Dietl, Simone Schür, Carlo Schaller, Christian E. Elger & Guillén Fernández. 2005. Word imageability affects the hippocampus in recognition memory. *Hippocampus* 15, 704–712.

Klimesch, Wolfgang. 1994. *The Dtructure of Long-Term Memory: A Connectivity Model of Semantic Processing*. Hillsdale, NJ: Erlbaum.

Knopf, M. 1992. Gedächtnis für Handlungen. Funktionsweise und Entwicklung. Dissertation, University of Heidelberg.

Knopf, M., Mack, W., Lenel, A., and Ferrante, S. 2005. Memory for action events: findings in neurological patients. *Scandinavian Journal of Psychology* 46, 11–19.

Knudsen, E. I. 2007. Fundamental components of attention. *Annual Review of Neuroscience* 30, 57–78.

Kormi-Nouri, R. 1995. The nature of memory for action events: An episodic integration view. *European Journal of Cognitive Psychology*, 337–363.

Kormi-Nouri, R. 2001. The motor component is not crucial! In H. Zimmer *et al.* (eds.), *Memory for Action: A distinct Form of Episodic Memory?*, 97–111. Oxford: Oxford University Press.

Kousta, S. T., G.Vigliocco, D. P. Vinson, M. Andrews & E. Del Campo. 2011. The representation of abstract words: Why emotion matters. *Journal of Experimental Psychology: General* 140, 14–34.

Krashen, Stephen & Tracy D. Terrell. 1983. *The Natural Approach: Language Acquisition in the Classroom*. Oxford: Pergamon.

Krashen, Stephen. 2000. What does it take to acquire language? *ESL Magazine* 3, 22–23.

Von Kriegstein, K. & A. L. Giraud. 2006. Implicit multisensory associations influence voice recognition. *PLoS Biology* 4, 1809–1820.

Von Kriegstein, K., Ö. Dogan, M. Grüter, A. L. Giraud, C. A. Kell, T. Grüter, A. Kleinschmidt & S. J. Kiebel. 2008. Simulation of talking faces in the human brain improves auditory speech recognition. *Proceedings of the National*

Academy of Sciences of the United States of America 105, 6747–6752.

Kuhl, P. K. 2010. Brain mechanisms in early language acquisition. *Neuron* 67, 713–727.

Kusanagi, Y. 2005. Analysis of research on nonverbal communication in language education. Paper presented at *JALT2004 Conference Proceedings*, Tokyo.

Kutas, M. & S. A. Hillyard. 1980. Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science* 207, 203–205.

Levy-Drori, S. & A. Henik. 2006. Concreteness and context availability in lexical decision tasks. *American Journal of Psychology* 119, 45–65.

Macedonia, M. 1996. Lautliche Bewegungsbilder. *Festschrift Universität Linz* 1.

Macedonia, M. 1999. *Sinn-voll Fremdsprachen unterrichten*. Linz: Veritas-Cornelsen.

Macedonia, Manuela. 2003. Sensorimotor enhancing of verbal memory through 'Voice Movement Icons' during encoding of foreign language (German: Voice Movement Icons. Sensomotorische Encodierungsstrategie zur Steigerung der quantitativen und qualitativen Lerneffizienz bei Fremdsprachen) University of Salzburg: PhD Thesis.

Macedonia, Manuela & Thomas R. Knösche. 2011. Body in mind: How gestures empower foreign language learning. *Mind, Brain, and Education* 5, 196–211.

Macedonia, Manuela, Karsten Müller & Angela D. Friederici. 2011. The impact of iconic gestures on foreign language word learning and its neural substrate. *Human Brain Mapping* 32, 982–998.

Mariani, L. 1981. English through mime. *English Teaching Forum* 19, 36–37.

Masumoto, K., M. Yamaguchi, K. Sutani, S. Tsuneto, A. Fujita & M. Tonoike. 2006. Reactivation of physical motor information in the memory of action events. *Brain Research* 1101, 102–109.

McCafferty, Steven G. 2008. Mimesis and second language acquisition: A sociocultural perspective. *Studies in Second Language Acquisition* 30, 147–167.

McCafferty, Steven G. & Gale Stam. 2008. *Gesture: Second Language Acquisition and Classroom Research*. London: Routledge.

McClelland, J. L. & T. T. Rogers. 2003. The parallel distributed processing approach to semantic cognition. *Nature Reviews Neuroscience* 4, 310–322.

McNeill, David. 1992. *Hand and Mind: What Gestures Reveal about Thought*. Chicago, IL & London: University of Chicago Press.

Mecklenbräuker, S., M. C. Steffens, P. Jelenec & N. K. Goergens. 2011. Interactive context integration in children? Evidence from an action memory study. *Journal of Experimental Child Psychology* 108, 747–761.

Meteyard, L., S. R. Cuadrado, B. Bahrami & G. Vigliocco. 2012. Coming of age: A review of embodiment and the neuroscience of semantics. *Cortex* 48, 788-804.

Moseley, R., F. Carota, O. Hauk, B. Mohr & F. Pulvermüller. 2012. A role for the motor system in binding abstract emotional meaning. *Cerebral Cortex* 22, 1634–1647.

Mukamel, Roy, Arne D. Ekstrom, Jonas Kaplan, Marco Iacoboni & Itzhak Fried. 2010. Single-neuron responses in humans during execution and observation of actions. *Current Biology* 20, 750–756.

Nadar, M. S. & J. McDowd. 2008. 'Show me, don't tell me'; is this a good approach for rehabilitation? *Clinical Rehabilitation* 22, 847–855.

Nilsson, L. G. 2000. Remembering actions and words. In F. I. M. Craik & E. Tulving (eds.), *The Oxford Handbook of Memory*, 137–148. Oxford: Oxford University Press.

Nilsson, L. G., L. Nyberg, T. Klingberg, C. Aberg, J. Persson & P. E. Roland. 2000. Activity in motor areas while remembering action events. *Neuroreport* 11, 2199–2201.

Nyberg, L., K. M. Petersson, L. G. Nilsson, J. Sandblom, C. Aberg & M. Ingvar. 2001. Reactivation of motor brain areas during explicit memory for actions. *Neuroimage* 14, 521–528.

Ozyurek, A., R. M. Willems, S. Kita, & P. Hagoort. 2007. On-line integration of semantic information from speech and gesture: Insights from event-related brain potentials. *Journal of Cognitive Neuroscience* 19, 605–616.

Paivio, A. 1969. Mental imagery in associative learning and memory. *Psychological Review* 76, 241–263.

Paivio, A. & K. Csapo. 1969. Concrete image and verbal memory codes. *Journal of Experimental Psychology* 80, 279-285.

Paivio, A. 2006. Dual coding theory and education. In S. Neuman (ed.), *Pathways to Literacy Achievement for High Poverty Children*. The University of Michigan School of Education.

Paivio, Allan. 1971. *Imagery and Verbal Processes*. New York: Holt.

Paivio, Allan. 1986. *Mental Representations: A Dual Coding Approach*. New York & Oxford: Oxford University Press & Clarendon Press.

Paivio, Allan. 1991. *Images in Mind: The Evolution of a Theory*. New York & London: Harvester Wheatsheaf.

Pulvermüller, F. 1996. *Neurobiologie der Sprache. Gehirntheoretische Überlegungen und empirische Befunde zur Sprachverarbeitung*, vol. 1 — Psychologia Universalis. Lengerich: Pabst Science Publishers.

Pulvermüller, F. 1999. Words in the brain's language. *Behavioral and Brain Sciences* 22, 253–336.

Pulvermüller, F. 2001. Brain reflections of words and their meaning. *Trends in Cognitive Sciences* 5, 517–524.

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

Pulvermüller, F. 2002. *The Neuroscience of Language: On Brain Circuits of Words and Serial Order*. Cambridge, MA & New York: Cambridge University Press.

Pulvermüller, F. & L. Fadiga. 2010. Active perception: Sensorimotor circuits as a cortical basis for language. *Nature Reviews Neuroscience* 11, 351–360.

Pylyshyn, Zenon W. 1984. *Computation and Cognition: Toward a Foundation for Cognitive Science*. Cambridge, MA: MIT Press.

Quinn-Allen, Linda 1995. The effects of emblematic gestures on the development and access of mental representations of French expressions. *The Modern Language Journal* 79, 521–529.

Rizzolatti, G. & M. A. Arbib. 1998. Language within our grasp. *Trends in*

Neurosciences 21, 188–194.

Romani, C., S. McAlpine & R. C. Martin. 2008. Concreteness effects in different tasks: Implications for models of short-term memory. *Quarterly Journal of Experimental Psychology* 61, 292–323.

Rugg, M. D. & M. E. Nagy. 1989. Event-related potentials and recognition memory for words. *Electroencephalography and Clinical Neurophysiology* 72, 395–406.

Rusted, J. M. 2003. Action based memory in Alzheimer's disease. *Research and Practice in Alzheimer's Disease* 7, 103–107.

Saltz, E. & S. Donnenwerthnolan. 1981. Does motoric imagery facilitate memory for sentences — a selective interference test. *Journal of Verbal Learning and Verbal Behavior* 20, 322–332.

Schatz, T. R., T. Spranger, V. Kubik & M. Knopf, 2011. Exploring the enactment effect from an information processing view: What can we learn from serial position analyses? *Scandinavian Journal of Psychology* 52, 509-515.

Schewe, Manfred & Peter Shaw. 1993. *Towards Drama as a Method in the Foreign Language Classroom*. Frankfurt am Main: Peter Lang.

Schwanenflugel, P. J., K. K. Harnishfeger & R. W. Stowe. 1988. Context availability and lexical decisions for abstract and concrete words. *Journal of Memory and Language* 27, 499–520.

Scott, S. K. 2004. The neural representation of concrete nouns: What's right and what's left?. *Trends in Cognitive Science* 8, 151–153.

Shams, Ladan & Aaron R. Seitz. 2008. Benefits of multisensory learning. *Trends in Cognitive Sciences* 12, 411–417.

Shams, Ladan, David Wozny, R. Kim, S. Robyn & Aaron Seitz. 2011. Influences of multisensory experience on subsequent unisensory processing. *Frontiers in Psychology* 2, 264.

Shimojo, Shinsuke & Ladan Shams. 2001. Sensory modalities are not separate modalities: Plasticity and interactions. *Current Opinion in Neurobiology* 11, 505–509.

Simon, Herbert A. 1981. Information-Processing models of cognition. *Journal of the American Society for Information Science* 32, 364–377.

Spranger, Tina, Tanja R. Schatz & Monika Knopf. 2008. Does action make you faster? A retrieval-based approach to investigating the origins of the enactment effect. *Scandinavian Journal of Psychology* 49, 487–495.

Straube, B., A. Green, S. Weis, A. Chatterjee & T. Kircher. 2009. Memory effects of speech and gesture binding: Cortical and hippocampal activation in relation to subsequent memory performance. *Journal of Cognitive Neuroscience* 21, 821–836.

Swain, M. & S. Lapkin. 1995. Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics* 16, 371–391.

Sydorenko, T. 2010. Modality of input and vocabulary acquisition. *Language Learning and Technology* 14, 50–73.

Taleghani-Nikazm, C. 2008. Gestures in foreign language classrooms: An empirical analysis of their organization and function. In Melissa Bowles et al. (eds.), *Selected Proceedings of the 2007 Second Language Research Forum,* 229–238. Somerville, MA: Cascadilla Proceedings Project.

Tellier, M. 2008. The effect of gestures on second language memorisation by young children. *Gesture* 8, 219–235.

Tomasello, M., M. Carpenter, & U. Liszkowski, 2007. A new look at infant pointing. *Child Development* 78, 705–722.

Tomasello, Michael. 2005. *Constructing a Language: A Usage-based Theory of Language Acquisition*. Cambridge, MA & London: Harvard University Press.

Tomasello, Michael. 2008. *Origins of Human Communication*: *The Jean Nicod Lectures*. Cambridge, MA & London: MIT Press.

Tomasi, Dardo & Nora D. Volkow. 2011. Functional connectivity hubs in the human brain. *NeuroImage* 57, 908–917.

Tonzar, C., L. Lotto & R. Job. 2009. L2 vocabulary acquisition in children: Effects of learning method and cognate status. *Language Learning* 59, 623–646.

Vigliocco, G., D. P. Vinson, W. Lewis & M. F. Garrett. 2004. Representing the meanings of object and action words: The featural and unitary semantic space hypothesis. *Cognitive Psychology* 48, 422–488.

Vigliocco, G., L. Meteyard, M. Andrews & S. Kousta. 2009. Toward a theory of semantic representation. *Language and Cognition* 1–2, 219–247.

Vogt, Stefan, Giovanni Buccino, Afra M Wohlschläger, Nicola Canessa, N. Jon Shah, Karl Zilles, Simon B. Eickhoff, Hans-Joachim Freund, Giacomo Rizzolatti & Gereon R Fink. 2007. Prefrontal involvement in imitation learning of hand actions: Effects of practice and expertise. *NeuroImage* 37, 1371–1383.

Vogt, Stefan & Roland Thomaschke. 2007. From visuo-motor interactions to imitation learning: Behavioural and brain imaging studies. *Journal of Sports Sciences* 25, 497–517.

Wennekers, T. & N. Ay. 2003. Spatial and temporal stochastic interaction in neuronal assemblies. *Theory in Biosciences* 122, 5–18.

West, W. C. & P. J. Holcomb. 2000. Imaginal, semantic, and surface-level processing of concrete and abstract words: An electrophysiological investigation. *Journal of Cognitive Neuroscience* 12, 1024–1037.

Wig, G. S., M. B. Miller, A. Kingstone & W. M. Kelley. 2004. Separable routes to human memory formation: Dissociating task and material contributions in the prefrontal cortex. *Journal of Cognitive Neuroscience* 16, 139–148.

Willems, Roel M. & Peter Hagoort. 2007. Neural evidence for the interplay between language, gesture, and action: A review. *Brain and Language* 101, 278–289.

Willems, Roel M., Aslı Özyürek & Peter Hagoort. 2007. When language meets action: The neural integration of gesture and speech. *Cerebral Cortex* 17, 2322–2333.

Willems, Roel M., Aslı Özyürek & Peter Hagoort. 2009. Differential roles for left inferior frontal and superior temporal cortex in multimodal integration of action and language. *NeuroImage* 47, 1992–2004.

Winitz, Harris (ed.). 1981. *The Comprehension Approach to Foreign Language Instruction*. Rowley, MA: Newbury House.

Wu, Ying Choon & Seana Coulson. 2007a. How iconic gestures enhance communication: An ERP study. *Brain and Language* 101, 234–245.

Wu, Ying Choon & Seana Coulson. 2007b. Iconic gestures prime related concepts: An ERP study. *Psychonomic Bulletin 38; Review* 14, 57–63.

Xua, J., P. J. Gannon, K. Emmorey, J. F. Smith & A. R. Braun. 2009. Symbolic gestures and spoken language are processed by a common neural system. *Proceedings of the National Academy of Sciences of the United States of America* 106, 20664–20669.

Zimmer, H. 2001a. Why do Actions speak louder than words. Action memory as a variant of encodingmanipulations or the result of a specific memory system? In H. D. Zimmer, R. Cohen, J. M. J. Guynn, R. Engelkamp & M. A. Foley (eds.), *Memory for Action: A Distinct form of episodicmemory*, 151–198. New York: Oxford University Press.

Zimmer, H. D., T. Helstrup & J. Engelkamp. 2000. Pop-out into memory: A retrieval mechanism that is enhanced with the recall of subject-performed tasks. *Journal of Experimental Psychology: Learning Memory and Cognition* 26, 658–670.

Zimmer, H. D. 2001b. *Memory for Action: A Distinct Form of Episodic Memory?* Oxford: Oxford University Press.

Manuela Macedonia and Katharina von Kriegstein
Max Planck Institute for Human Cognitive and Brain Sciences
Stephanstr. 1A
04103 Leipzig
Germany
macedonia@cbs.mpg.de and kriegstein@cbs.mpg.de

Bidirectional Influences of Emotion and Action in Evaluation of Emotionally-Connoted Words

Audrey Milhau, Thibaut Brouillet, Loïc Heurley
& Denis Brouillet

The goal of this review is to present the embodied character of emotionally-connoted language through the study of the mutual influences of affective language and motor action. After a brief definition of the embodied approach of cognition, the activity of language understanding is presented as an off-line embodied process implying sensory-motor resonance. Then the bidirectional character of influences between language and action will be addressed in both behavioral and neuropsychological studies, illustrated by the specific case of emotionally-connoted language. These reciprocal effects are grounded on the motor correspondence between action and the motor dimension of language, emerging from a diversity of source such as adaptive motivation, past experiences, body specificities, or motor fluency.

Keywords: action; bidirectional influences; embodied cognition; emotion; motor fluency

1. Introduction

The evolution during the 20th century of the concept of cognition and of the relationship between the individual and his environment has led most recent approaches in this area to ask the question of the incarnation of cognition. Indeed, the body appears as the first media of interaction with the environment, in the sense that it influences and constrains the cognitive processes: An object, a situation, only makes sense because it implies a number of actions for the cognitive subject. Cognition, therefore, cannot be considered as an independent entity, divorced from its context or from its surroundings, either it is internal (non-cognitive process, motricity, bodily states) or external (places, situations, events, etc.). Instead, human cognition can only be considered as the coordination of cognitive and non-cognitive processes, in a situated and integrated way (Barsalou 2008). By replacing the cognitive activity in context, the involvement of the body as a medium of the self/environment link emerged as crucial to the study and understanding of cognition.

If reading and understanding activities are based on embodied processes,

ISSN 1450–341

then the emotional content of language must also be embodied in order to be understood. Furthermore, the emergence of a new conception of emotion (with ancient roots in philosophy: William James 1894) defined as the consequence of bodily phenomena contribute to address the question of the embodiment of emotional language. In this article, our goal is to present the embodied character of emotionally-connoted language, and the mutual influences between affective language and motor action. The first section is focused on a brief definition of the embodied approach of cognition. We define the mechanisms linked to the incarnation of mental processes, and present language as an off-line embodied activity, therefore implying specific motor resonance. In the second section, we address the question of the bidirectionality of the influences between language understanding and action. Behavioral and neuropsychological studies support these mutual effects. Finally, the third section extends the bidirectional influences to the study of emotionally-connoted language, linked to motivational and non-motivational movements. The compatibility effects between affective language and motor activity are discussed in terms of motor fluency, as a novel and encouraging explanation for the matching of emotion and action.

2. Embodiment in Cognition and Language

Sensory-motor processes used to simply be considered by classic theories of cognition as input or output to cognitive systems. Nowadays, those processes are entirely integrated in recent approaches of mental functioning, known as embodied theories of cognition. Those conceptions are funded on a rejection of the segregation between high- and low-level mental processes, and generally consider mental activity in a unifying perspective. One major contribution of embodied approaches is the redefinition of memory as a memory of processes and no longer a memory of content. According to this view, the main function of memory is to mediate the interaction with the world, by matching current activity and traces left by previous experiences (Glenberg 1997). Barsalou (2008) defined this phenomenon as the "re-enactment of perceptual, motor and introspective states acquired during the interaction with the environment, the body and the mind" (p. 618). Concretely, the interaction with an object, either directly (via perception) or indirectly (via thought or imagination) consists in the mental simulation of the multimodal states linked to brain systems allocated to perception and action that were running during precedent interactions (Barsalou 1999).

Perceptual and motor processes are central in this approach, to such an extent that some theories propose a similarity of nature between perception and action: Theory of Event Coding (Hommel *et al.* 2001; Hommel 2004), Mirror Neuron System theory (Rizzolatti *et al.* 1996; Gallese *et al.* 1996; Gallese 2008). According to these approaches, the same kind of representations is used to match motor and perceptual states, explaining the resonance between perception and motor planning or execution in situation of simulation (Bach *et al.* 2010). Mental simulation, defined as the re-enactment of sensori-motor states usually linked to an object when interacting with it, can be explained in terms of an automatic and direct link between perception and action (*affordances*; Gibson 1979; Tucker &

Ellis 1998), or in terms of motor resonance.

According to various authors (Zwaan 2004, 2009; Buccino *et al.* 2005), observation and simulation of action are based on the individual own motor repertory. When observing or imagining an action that one is able to perform, motor areas in the brain are activated as if one was actually executing it (Gallese 2008), which is called motor resonance. On the other hand, when the activity is not in the individual repertoire (such as a technical dance movement, or a movement done by an animal), there is no motor resonance in the brain, because the body has no traces of such a movement.

We focused until here on simulation and resonance in situations with the object of mental activity being present, and available from direct perception, that is to say in on-line situations. However, embodied theory of cognition extends those processes to off-line situations, namely cases where the object of mental activity is absent, not available from direct perception (Niedenthal *et al.* 2005).

Language is a key example of off-line cognition: By this means, people make reference to objects that are absent from their environment, so they are able to interact with it via linguistic production and understanding. As for perception or visual imagination of action, language understanding is defined as a sensory-motor activation functioning by the generation of motor resonance between the linguistic content and the experiential repertoire of the individual (Barsalou 1999; Gibbs 2003; Zwaan 2004, 2009; for a review on the role of motor processes on language understanding, see Willems & Hagoort 2007; Fisher & Zwaan 2008; Gallese 2008). Glenberg & Robertson (1999) describe three stages in language understanding: Firstly language content is matched with the referred object, then the cognitive system simulates the action possibilities implied by this object, and finally a coherent action pattern is created, which leads to understanding. Different actions can be involved in those simulations: simulation of the action of communication (lips and tongue movements, i.e. articulatory system) and simulation of language content (for examples and reviews see Galantucci *et al.* 2006; Willems & Hagoort 2007; Gallese 2008). In this section and the followings, we will focus on the second type of motor resonance: language content simulation.

Concerning language content, if simulation theory clearly explains the embodiment of concrete language, the question of the embodiment of abstract language has been raised (for a review, see Pecher *et al.* 2011). One major account in favor of the embodiment of abstract language is based on the notion of linguistic metaphor (Lakoff & Johnson 1980, 1999; Gibbs 2003; Meier & Robinson 2004). Conceptualization of abstract notion is supposed to process by analogy, with a generalization of action patterns initially associated to a more concrete concept. For example, an abstract concept such as time is frequently associated to the more concrete concept of space, leading to a conception of time as a spatial dimension. Past is figured by the space behind oneself, and future is represented by the space in front of oneself (Boroditsky 2000).

Since language is considered as an embodied activity, no one should be surprised of the ties between language and action, and of the influence that these processes have on each other.

3. Bidirectional Influence of Language and Action

The study of language and action and of the mutual effects between them has inspired many researches, both in behavioral and neuropsychological approaches. This presentation concerns minimal linguistic units (isolated words) as well than sentences processing (however, see Willems & Hagoort 2007; Fisher & Zwaan 2008; Gallese 2008 for reviews). Moreover, the term action is used here to describe both action planning, activation of motor areas in the brain or actual execution of movements.

3.1. Effect of Language Understanding on Action

One of the first works which can be interpreted with an embodied approach of language is the study published by Bargh *et al.* (1996). The authors demonstrated that the exposition to words relative to the elderly (but with no explicit mention of this common point) have made the participants' walk slower, with their back bent, in a way reminding the way old people moves.

This study is an example of what has been called the Action Compatibility Effect (ACE; Glenberg & Kaschak 2002). This effect shows that perceive linguistic content relative to a specific action leads to behave in a way congruent with this action. When reading a sentence such as "Andy delivers the pizza to you", the realization of a congruent movement (hand moving toward the body) is easier than an opposite movement (hand moving away from the body). This compatibility effect also manifests itself with more abstract sentences like "Liz told you the story", where the notion of transfer is less evident. Many studies have been inspired by the ACE. Some works have demonstrated an effect of words on the planning of action (Gentilucci *et al.* 2000; Tucker & Ellis 2004), others have shown a facilitation effect for the activity of specific effectors when reading sentences describing movements of the same effectors (Scorolli & Borghi 2007; Borghi & Scorolli 2009). Furthermore, it has been demonstrated that the mere description of a specific concept activates the realization of ocular movements used to interact with it in everyday life (e.g., look up when hearing about a building, look down when hearing about a canyon, Spivey *et al.* 2000). Finally, studies in our laboratory have shown that even verbal responses considered disembodied as "yes" and "no" have a motor component, manifesting specific associations with movements of response (Brouillet *et al.* 2010). In the neuropsychological field, the study of premotor and motor areas of the brain has shown specific activation depending on the linguistic content. Various studies using imagery techniques such as functional MRI have demonstrated that reading action words associated to specific effectors (e.g., *pick*/*lick*/*kick*; Hauk *et al.* 2004) led to a somatotopic activation of motor and premotor areas in the brain (see also Buccino *et al.* 2005; Pulvermüller *et al.* 2005b; Tettamanti *et al.* 2005; for a review, see Jirak *et al.* 2010).

Both the behavioral and neuropsychological data support the idea of an influence of language processing on the planning and the realization of action. But if a motor resonance can occur between the motor component activated by language understanding, and actual motor execution, embodied theories of cognition predict that the reverse association is also possible. The next section considers the influence of motor activity on the comprehension of language.

3.2. *Effect of Action on Language Understanding*

Few researches have focused on this reverse link between action and language. Nevertheless, this question is of importance to validate the simulation account of language understanding. Indeed, one major critic against the simulation approach of language comprehension is that the motor activation linked to language only occurs after comprehension, and is not a part of the processing of linguistic content (Mahon & Caramazza 2008). One way to answer to this critic is to show that a simple motor activation is sufficient to enhance language understanding. If action facilitates comprehension, it would validate the idea of a critical motor activation in language understanding (Rueschemeyer *et al.* 2010). Some experimental works indeed demonstrate an effect of action on action language comprehension.

In 1989, Klatzky *et al.* had already shown that the presentation of manual postures primed the comprehension of word pairs describing a congruent action. More recently, Lindemann *et al.* (2006) demonstrated that action preparation (e.g. "drink from a cup") activates the semantic dimension linked to the goal of the movement prepared. This semantic activation functions as a prime and facilitates the performance in a subsequent lexical decision task (e.g., "Is *mouth* a word?"). Helbig *et al.* (2006) have shown a priming effect in a denomination task, when succeeding pictures consisted in objects usually used in the same way (e.g., nut-cracker and pliers; see Myung *et al.* 2006 for similar results using words related to the same actions). But one critic to this kind of studies is that the effect observed is more a visual effect than a motor one: It can be the perceptual similarities between primes and targets that allow facilitation. To reject this point, Rueschemeyer *et al.* (2010) designed a study in which words were primed by an actual motor execution, with no linguistic or visual input: Participants had to judge if words denoted functionally manipulable objects or not, while realizing an intentional movement, a passive movement or no movement at all. Results showed that participants were faster to recognize words denoting manipulable objects when they were engaged in an intentional action rather than in a passive action, or no action. This study illustrates that the activation of the neural motor system facilitates the processing of words with a motor component.

Various neuropsychological studies also support the effect of action on language understanding (for reviews see Willems & Hagoort 2007; Borghi & Pecher 2011). In particular, one criticism made to an embodied approach of language understanding was to consider that the motor activations observed in cortex and pre-cortex were nothing but by-product of activations linked to imagery of language content. Pulvermüller *et al.* (2005a) argued against this claim, showing with magnetoencephalography (MEG) that activity in the primary motor cortex linked to leg activity occurs 150 ms after the onset of action-word describing leg movements. This very early activation in the comprehension process rules out the exclusion of motor areas in language understanding.

All these works are in favor of a bidirectional link between action and language, thanks to the mutual activation of the motor component of both linguistic content, and of planned or executed movement. We will now address the question of a particular type of language that is emotionally-connoted language, and its bidirectional links with action.

4. The Case for Emotion

4.1. An Alternative Definition of Emotion

A classical conception of emotion, in accordance with common sense, would define emotion as an evaluative cognitive mechanism caused by the confrontation with an emotionally-connoted object, and responsible for bodily modifications at an internal (e.g., cardiac rhythm) or external level (e.g., approach behavior). For example, when one perceives an object evoking fear (e.g., a spider, a snake), this perception activates the emotion of fear, which implies bodily symptoms like shaking, sweating, escape reflex, etc.

But in the late nineteenth century, American philosopher William James has proposed an alternative definition of emotion, inversing this conception: Emotion might be the very consequence of the bodily changes. In a communication from 1884, James explains that bodily changes follow directly the perception of the exciting fact. Concretely, emotion is the feeling of these changes. According to this conception, if during the emotional experience, everything physical is taken back, there is nothing left of the emotion. In summary, there is no such thing as a disembodied emotion. Consequently, the simple fact of voluntarily activate bodily symptoms might be sufficient to create the emotion.

Contemporary authors have integrated this conception of emotion, and agree on the idea that comprehension of emotionally-connoted language needs at least a partial simulation of the same neuronal and bodily mechanism than actual emotional experience (Glenberg *et al.* 2005; Havas *et al.* 2007; Winkielman *et al.* 2008). The point is that bodily expression and interpretation of emotion share a reciprocal relation (Niedenthal 2007), such as when the body is already in a state linked to emotion, there is a facilitation to feel the same emotion. Various manipulations of emotional expressions have been used to illustrate this effect (see section 4.3.1. for a detailed presentation). Conversely, understanding emotionally-connoted language implies the simulation of the bodily state associated. Vermeulen *et al.* (2007) have demonstrated a switching cost in a property verification task: Participants were faster to recover the affective dimension of concepts after a trial into which they had to simulate an emotional feature rather than a sensorial one (e.g., visual, auditory, etc). This result indicates that to access the emotional dimension of concepts, a simulation of the state associated to this emotion is necessary, and that this simulation enters more easily in resonance with another affective simulation than with a sensorial one. This effect supports the idea of the necessity of an embodied activation to understand affective concepts.

This approach of emotion has since been illustrated by an increasing amount of studies, confronting emotionally-connoted language and various bodily behaviors, either directly or indirectly associated to emotion.

4.2. *Emotionally-Connoted Language and Motivational Movements*

In accordance with the conception presented in the precedent section, the role of emotion can be defined as a signal to indicate to the organism the presence of a relevant element in the environment, in terms of its potential consequences. The emotion is therefore the result of an evaluation concluding that the presence of either a negative element represents a danger or a positive one represents a potential benefit (Neumann *et al.* 2003). This evaluation has to be related to an important function of cognition which is motivation: Once the situation has been defined as positive or negative, the organism has to plan his behavior, to be as adapted as possible (Lang *et al.* 1990). Depending on the evaluation, two kinds of motivation might arise. Elliot (2006: 112) explains:

> Positively evaluated stimuli are inherently associated with an approach orientation to bring or keep the stimuli close to the organism (literally or figuratively), whereas negatively evaluated stimuli are inherently associated with an avoidance orientation to push or keep the stimuli away from the organism (literally or figuratively).

The mutual influence of emotional processes and approach and avoidance behaviors has inspired many works since the 1990's (Cacioppo *et al.* 1993; Förster & Strack 1997; Chen & Bargh 1999; Neumann & Strack 2000; Freina *et al.* 2009; Brouillet *et al.* 2010). One methodology to study those associations proposes to use arm flexion/extension movements as approach and avoidance behaviors. Cacioppo *et al.* (1993), with neutral items (unknown Chinese ideographs), and Neumann & Strack (2000) with emotionally connoted items, showed that flexion (approach behavior) allowed more positive evaluations than extension (avoidance behavior), while the extension elicited more negative responses than flexion. These works therefore illustrate the influence of motor actions of approach/avoidance on the evaluation of objects. With another procedure, Chen & Bargh (1999) demonstrated the reverse effect (i.e. the influence of emotional processes on motivational behaviors), showing that the perception of emotional words allows faster movements of response (push or pull a lever) when it corresponds to the enabled motivation (approach/pulling for the positive words, and avoidance/pushing for the negative words) than when this movement is contrary to the motor program activated. This effect of emotionally-connoted language on action has to be distinguished from pure ACE. The Action Compatibility Effect (Glenberg & Kaschak 2002) results from a direct motor resonance between the language content (either concrete or abstract) and the action planified or realized. Concerning emotional linguistic items, the matching between the stimuli and the actions is not direct, but is due to the conjunction of in one hand, the adaptive associations between emotion and motivational behaviors, and in the other hand the actual realization of such movements.

Other authors have pointed out the contextual nature of approach and avoidance behaviors: If approach consists in reducing the distance between oneself and an object, one can either move closer to the object, or move the object closer to oneself. Similarly, to avoid an object, one can move away from it, but can also push it. Freina *et al.* (2009) showed that Chen & Bargh's results (1999), obtained when subjects had a lever in the hand, could be reversed when their hands were empty. It appears that when individuals take the lever, they act as if this object was the emotional item and therefore push it when it is negative, and pull it when it is positive. On the contrary, when the hand is empty, it approaches positive items on the screen, and get away from the negative items. Similarly, Cretenet & Dru (2004, 2008; see also Dru & Cretenet 2005) specified the conclusions of Cacioppo *et al.* (1993) by distinguishing the responses from the left or right arm. Movements compatible with positive words (i.e. allowing the most positive evaluations) are not simply flexion ones, but right arm flexion and/or left arm extension. By contrast, negative words are compatible with the flexion of the left arm or the extension of the right arm ("motor congruence effect"; Cretenet & Dru 2004, 2008; Dru & Cretenet 2005, 2008). The significance, the context and the laterality of the movement are essential for the study of associations between emotionally-connoted and motivational behaviors.

To summarize, all these studies are in accordance with the idea of a bidirectional nature of the links between affective language and motivational behaviors (Neumann & Strack 2000; Neumann *et al.* 2003; Centerbar & Clore 2006).

4.3. *Compatibility between Emotionally-Connoted Language and Other Kinds of Action*

4.3.1. *A Diversity of Bodily Expressions*

All the works presented previously are based on a total motor resonance between the movement orientations evoked by the linguistic material and the actual responses executed by the participants: All the behaviors considered in these studies are standard approach and avoidance behaviors. But a wide range of motor behaviors were found to match with the motor component of emotionally-connoted language.

One famous paradigm was conceived by Strack *et al.* (1988), and transferred to the study of language by Glenberg *et al.* (2005) and Havas *et al.* (2007). This manipulation provokes or inhibits smile because the participant has to hold a pen either between his teeth (provoking a muscular activation close to a smile), or between his lips (implying a muscular activation incompatible with a smile). While holding the pen, participants had to evaluate emotionally-connoted sentences. Results showed a compatibility effect such as smiling enhanced the positive evaluation of positive sentences, while pout implied a more negative evaluation of negative sentences.

Compatibility effects between emotional expressions and the understanding of emotionally-connoted language were also studied with the use of head movements. Head movements are, at least in occidental cultures, strongly associated with intention to agree or refuse. Wells & Petty (1980) have first used these bodily manipulations to study effect of emotion expression on attitudes. In this study, participants had to express their degree of agreement with school reforms. They were listening to an auditory message explaining this reform, and while listening, they had to move their head either horizontally or vertically, under the cover story of judging the quality of headphones. Results showed that participants who had shaken their head vertically were more convinced by the message than the ones who had shaken their head horizontally. The authors interpreted this effect as an influence of the expression of agreement on the formation of a positive attitude. In a same way, and with linguistic content, ask participants to nod either as if they were agreeing or refusing something during the presentation of emotionally-connoted words allowed a better restitution of the words presented in a compatible condition, namely agreeing in front of positive words, and refusing in front of negative words (Förster & Strack 1996).

These researches demonstrate that the context of a specific movement is sufficient to transfer its affective component to the evaluation of object such as linguistic stimuli.

4.3.2. Emotionally-Connoted Language and Lateral Movements

Recent works have shown that even movements apparently unrelated to motivation or emotion could be associated to valence. The question of laterality was raised in the study of embodied language, with the work of Cretenet and Dru (2004, 2008; Dru & Cretenet 2005, 2008) which has introduced the question of the hand used in approach/avoidance compatibility effect (see section 4.2. above), but also in neurological studies. Willems *et al.* (2010) have shown that verbs describing manual and non-manual actions led to different brain activation, with manual action-verbs implying an activation of the premotor cortex, dependant of the manual dominance of the participants (left premotor cortex for right-handers, and right premotor cortex for left-handers).

In the specific field of emotion, Casasanto (2009) showed that when subjects were asked to place two items (one positive and one negative) in two separate areas situated on the left and the right part of space, they tended to relate the positive one to the side of their dominant hand, and the negative one to the other side. The author interprets this effect with the "body-specificity hypothesis": Subjects with different bodies and who use it in a different way (right-handers and left-handers) conceive the world differently. Similar to the precisions made to the studies made by Cacioppo *et al.* (1993) or Chen & Bargh (1999), these associations might be dependent on the action of the subject. Casasanto & Chrysikou (2011) asked their right-handed participants to first perform a motor task, by manipulating little objects with their two hands, while their right-hand was disabled by wearing a bulky ski glove. This manipulation was designed to make them feel and act as left-handers. The second task was the same than in Casasanto (2009), and showed that momentarily disabling the right-hand made right-handers manifest the same valence/laterality effect than left-handers, namely to associate positive to the left, and negative to the right. It indicates that associations related to dominant preference can be reversed by short-term changes in individual behavior. Similarly, work in our laboratory showed that right-handers acting as left-handers in a valence judgment task expressed the same compatibility effects between left and positive than left-handers (Milhau *et* *al.* 2012). Precisely, participants had to judge the valence of positive and negative word by pressing two keys with respectively a rightward and leftward arm movement (the matching valence/laterality was counterbalanced). They were all right-handed but had to use either their right or their left hand to answer. Results showed that when using their dominant hand, right-handers manifested the associations exposed by Casasanto (2009): They were faster to recognize the positive word when pressing the right key than the left key. But the main result is that when using their left hand to answer, right-handed participants seems to become real left-handers, since they were faster in judging positive word when answering with a leftward movement than with a rightward action. However, this effect was limited to positive stimuli and did not extend to negative words (an interpretation of this result is proposed in the next section).

4.3.3. Motor Fluency as an Explicative Approach of Compatibility Effects

When the compatibility effect between motion perception and expression is not based on the similarity of action orientation (i.e. approach/avoidance behaviors), another kind of processes must be responsible of this matching phenomenon. One interpretation, proposed by Casasanto (2009), is that the tendency to associate positive concept to the side of the dominant hand was due to the fact that participants interact with the world more frequently with this hand, and are therefore more efficient (both faster and more accurate) when executing actions on the corresponding side. Another explanation, purely biomechanical, can be invoked to explain those effects. The majority of studies about the associations between emotionally-connoted language and non-emotional movements, like lateral movements, implied the activity of the arms. Anatomical data informs us that arm movements executed across the body (adduction, e.g., right arm directed to the left) are much more costly in terms of energy than movements to the same side (abduction, Laude *et al.* 1978).

Both of these propositions focus on the dimension of facility of one movement over another one: Recent studies have advanced the notion of "motor fluency" to qualify the ease linked to the realization of specific movements, due either to habit or to a low energy cost (Beilock & Holt 2007; Casasanto & Chrysikou 2011). This concept refers to a definition of fluency as an indicator of the continuous and effortless flow by which a process is realized (Alter & Oppenheimer 2009). This experience of fluency seems to emerge of a comparison, a shift between the cost initially anticipated to realize an activity, and the cost of the actual execution of this activity. Therefore, and in accordance with ideomotor theories of cognition (Hommel *et al.* 2001), motor fluency implies that action is conceived in terms of its perceptual consequences.

In addition, analogous to the fact that perceptual (Reber *et al.* 1998; Winkielman *et al.* 2002) or conceptual fluency (Whittlesea 1993, Experiment 5) are emotionally marked, motor fluency seems to be associated with a positive emotional tagging. Since we consider fluency as monitoring the quality of one's own functioning, this positive marking would thus act as a reward system, a positive reinforcement for a successful motor activity (Winkielman & Cacioppo 2001; Cannon *et al.* 2010; Brouillet *et al.* 2011). This positive marking might be responsible for the fact that motor fluency impacts specifically positive stimuli (Milhau *et al.* 2012, see also Brouillet *et al.* 2010). The fluent movements implied a positive feeling that entered in resonance with the emotional dimension of the positive stimuli, allowing the compatibility effect highlighted in this study. Congruent with this positive marking of motor fluency, we ran in our laboratory an experiment demonstrating that the emotional dimension linked to lateral fluent movements could be sufficient to allow an emotional evaluation of neutral linguistic stimuli. The participants' task was first to execute a lateral arm movement in response to a visual signal while reading a neutral word. Then they had to evaluate the valence of the word on a scale. All the participants were right-handers and used their right hand to respond. Our results demonstrated that words judged after a rightward movement were evaluated as more positive than after a leftward movement. The motor fluency of an arm movement directed to the dominant side for the participants led them to project the positive marking of this movement onto the evaluation of neutral words (Milhau *et al.* 2012). Conversely, the absence of fluency has never been associated to negativity (Winkielman & Cacioppo 2001; Brouillet *et al.* 2011), probably because it does not question the quality of the system's functioning, and therefore it does not need any negative reinforcement in order to adapt a behavior.

To summarize, compatibility effects between non-motivational behaviors and emotionally-connoted stimuli do not exist *per se*, but depend on the context of the movement and its energy cost, namely its motor fluency.

5. Conclusion

Our goal in this article was to emphasize the bidirectional nature of the links between emotion and action, in the particular field of language understanding.

The works reviewed here lead us to two main conclusions. First, the processing of emotionally-connoted words acts as an external signal for the cognitive system that enables the compatible behavior. Subsequently to the presentation of affective language, the realization of a compatible motor response is facilitated, either in terms of motivation with approach and avoidance behaviors, or in terms of motor fluency with lateral movements. Secondly, action (understood as either action planning, motor and pre-motor cortex activation or actual motor execution) can influence the understanding of emotionally-connoted language, or even provide an emotional connotation to initially neutral linguistic material, thanks to the positive marking of motor fluency.

Both of these findings resonate with William James' (1884) claim: Emotion is not a cause but a consequence of cognitive activity. This emotional experience emerges from the successful matching (or not) between a cognitive activity, and the signification assigned to it. In the specific field of emotionally-connoted language, this matching occurs because of the resonance between a motor process, and the motor dimension of language. The wide variety of researches demonstrates that the motor correspondence between these elements can come from a diversity of sources, such as adaptive motivation, past experiences, body specificities or fluency.

References

Alter, Adam. L. & Daniel M. Oppenheimer. 2009. Suppressing secrecy through metacognitive ease: Cognitive fluency encourages self-disclosure. *Psychological Science* 20, 1414–1420.

Bach, Patric, Debra Griffiths, Matthias Weigelt & Steven P. Tipper. 2010. Gesturing meaning: non-action words activate the motor system. *Frontiers in Human Neuroscience* 4, 1–12.

Bargh, John A., Mark Chen & Lara Burrows. 1996. Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology* 71, 230–244.

Barsalou, Lawrence W. 1999. Perceptual Symbol Systems. *Behavioral and Brain Sciences* 22, 577–660.

Barsalou, Lawrence W. 2008. Grounded cognition. *Annual Review of Psychology* 59, 617–645.

Beilock, Sian L. & Lauren E. Holt. 2007. Embodied preference judgments: Can likeability be driven by the motor system? *Psychological Science* 18, 51–57.

Borghi, Anna M. & Diane Pecher. 2011. Introduction to the special topic Embodied and Grounded Cognition. *Frontiers in Psychology* 2, 187.

Borghi, Anna M. & Claudia Scorolli. 2009. Language comprehension and hand motion simulation. *Human Movement Science* 28, 12–27.

Boroditsky, Lera. 2000. Metaphoric Structuring: Understanding time through spatial metaphors. *Cognition* 75, 1–28.

Brouillet, Thibaut, Loïc Heurley, Sophie Martin & Denis Brouillet. 2010. The embodied cognition theory and the motor component of "yes" and "no" verbal responses. *Acta Psychologica* 134, 310–318.

Brouillet, Thibaut, Laurent P. Ferrier, Anne Grosselin & Denis Brouillet. 2011. Action compatibility effects are hedonically marked and have incidental consequences on affective judgment. *Emotion* 11, 1202–1205.

Buccino, Giovanni., Lucia Riggio, Giorgia Melli, Ferdinand Binkofski, Vittorio Gallese & Giacomo Rizzolatti. 2005. Listening to action related sentences modulates the activity of the motor system: A combined TMS and behavioral study. *Cognitive Brain Research* 24, 355–363.

Cacioppo, John T., Joseph R. Priester & Gary C. Berntson. 1993. Rudimentary determinants of attitudes II: Arm flexion and extension have differential effects on attitudes. *Journal of Personality and Social Psychology* 65, 5–17.

Cannon, Peter R., Amy E. Hayes & Steven P. Tipper. 2010. Sensorimotor fluency influences affect: Evidence from electromyography. *Cognition & Emotion* 24, 681–691.

Casasanto, Daniel. 2009. Embodiment of abstract concepts: Good and bad in right- and left-handers. *Journal of Experimental Psychology: General* 138, 351–367.

Casasanto, Daniel & Evangelia G. Chrysikou. 2011. When left is 'right': Motor fluency shapes abstract concepts. *Psychological Science* 22, 419–422.

Centerbar David B. & Gerald L. Clore. 2006. Do approach-avoidance actions create attitudes? *Psychological Science* 17, 22–29.

Chen, Mark & John Bargh. 1999. Consequences of automatic evaluation: Immediate behavioural predispositions to approach or avoid the stimulus. *Personality & Social Psychology Bulletin* 25, 215–224.

Cretenet, Joël & Vincent Dru. 2004. The influence of unilateral and bilateral arm flexion vs. extension on judgments: An exploratory case of a motor congruence hypothesis. *Emotion* 4, 282–294.

Cretenet, Joël & Vincent Dru. 2008. A neurobehavioural investigation into judgmental processes: Effects of bilateral motor behaviors. *Brain and Cognition* 68, 81–91.

Dru, Vincent & Joël Cretenet. 2005. Influence of bilateral motor action on judgments: Additional evidence for the motor congruence hypothesis. *Laterality* 10, 389–398.

Dru, Vincent & Joël Cretenet. 2008. Influence of unilateral motor behaviors on the judgment of valenced stimuli. *Cortex* 44, 717–727.

Elliot, Andrew J. 2006. The hierarchical model of approach-avoidance motivation. *Motivation and Emotion* 30, 111–116.

Fischer, Martin H. & Rolf A. Zwaan. 2008. Embodied language — A review of the role of the motor system in language comprehension. *Quarterly Journal of Experimental Psychology* 61, 825–850.

Förster, Jens & Fritz Strack. 1996. Influence of overt head movements on memory for valenced words: A case of conceptual-motor compatibility. *Journal of Personality and Social Psychology* 71, 421–430.

Förster, Jens & Fritz Strack. 1997. Motor actions in retrieval of valenced information: A motor congruence effect. *Perceptual and Motor Skills* 85, 1419–1427.

Freina, Laura, Giulia Baroni, Anna Borghi & Roberto Nicoletti. 2009. Emotive concepts-nouns and motor responses: Attraction or repulsion? *Memory and Cognition* 37, 493–499.

Galantucci, Bruno, Carol Fowler & M. T. Turvey. 2006. The motor theory of speech perception reviewed. *Pyschonomic Bulletin & Review* 13, 361–377.

Gallese, Vittorio. 2008. Mirror neurons and the social nature of language: The neural exploitation hypothesis. *Social Neuroscience* 3, 317–333.

Gallese, Vittorio, Luciano Fadiga, Leonardo Fogassi & Giacomo Rizzolatti. 1996. Action recognition in the premotor cortex. *Brain* 119, 593–609.

Gentilucci, Maurizio, Francesca Benuzzi, Luca Bertolani, Elena Daprati & Massimo Gangitano. 2000. Language and motor control. *Experimental Brain Research* 133, 468–490.

Gibbs, Raymond W. 2003. Embodied experience and linguistic meaning. *Brain and Language* 84, 1–15.

Gibson, James J. 1979. *The Ecological Approach to Visual Perception*. Boston, MA: Houghton Mifflin.

Glenberg, Arthur M. 1997. What memory is for? *Behavioral and Brain Sciences* 20, 1–55.

Glenberg, Arthur M. & David A. Robertson. 1999. Indexical understanding of instructions. *Discourse Processes* 28, 1–26.

Glenberg, Arthur. M. & Michael P. Kaschak. 2002. Grounding language in action. *Psychonomic Bulletin & Review* 9, 558–565.

Glenberg, Arthur M., David Havas, Raymond Becker & Mike Rinck. 2005.

Grounding language in bodily states: The case for emotion. In Diane Pecher & Rolf Zwaan (eds.), *The Grounding of Cognition: The Role of Perception and Action in Memory, Language and Thinking*. Cambridge, MA: MIT Press.

Hauk, Olaf, Ingrid Johnsrude & Friedemann Pulvermüller. 2004. Somatotopic representation of action words in human motor and premotor cortex. *Neuron* 41, 301–307.

Havas, David A., Arthur M. Glenberg & Mike Rinck. 2007. Emotion simulation during language comprehension. *Psychonomic Bulletin & Review* 14, 436–441.

Helbig, Hannah, Markus Graf & Markus Kiefer. 2006. The role of action representations in visual object recognition. *Experimental Brain Research* 174, 221–228.

Hommel, Bernhard. 2004. Event files: Feature binding in and across perception and action. *Trends in Cognitive Sciences* 8, 494–500.

Hommel, Bernhard, Jochen Müsseler, Gisa Aschersleben & Wolfgang Prinz. 2001. The theory of event coding TEC: A framework for perception and action. *Behavioral and Brain Sciences* 24, 849–937.

James, William. 1884. What is an emotion? *Mind* 9, 188–205.

Jirak, Doreen, Mareike M. Menz, Giovanni Buccino, Anna M. Borghi & Ferdinand Binkofski. 2010. Grasping language-a short story on embodiment. *Consciousness and Cognition* 19, 711–720.

Klatzky, Roberta L., James W. Pellegrino & Brian McCloskey. 1989. Can you squeeze a tomato? The role of motor representations in semantic sensibility judgments. *Journal of Memory & Language* 28, 56–77.

Lakoff, George, & Mark Johnson. 1980. *Metaphors We Live by.* Chicago, IL: University of Chicago Press.

Lakoff, George, & Mark Johnson. 1999. *Philosophy in the Flesh*. New York: Basic Books.

Lang, Peter. J., Margaret M. Bradley & Bruce N. Cuthbert. 1990. Emotion, attention, and the startle reflex. *Psychological review* 97, 377–395.

Laude, Maurice, Claude Kénési, Didier Patte & Edwin Riggs. 1978. L'abduction du bras. *Surgical and Radiologic Anatomy* 1, 65–72.

Lindemann, Oliver, Prisca Stenneken, Hein T. Van Schie & Harold Bekkering. 2006. Semantic activation in action planning. *Journal of Experimental Psychology: Human Perception and Performance* 32, 633–643.

Mahon, Bradford, & Alfonso Caramazza. 2008. A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology* 102, 59–70.

Meier, Brian P. & Michael D. Robinson. 2004. Why the sunny side is up: Associations between affect and vertical position. *Psychological Science* 15, 243–247

Milhau, Audrey, Thibaut Brouillet, Loïc Heurley & Denis Brouillet. 2012. When emotion influences action: Compatibility effect between emotional words and arm movements. Unpublished manuscript.

Milhau, Audrey, Thibaut Brouillet & Denis Brouillet. 2012. Emotion and motor action: biases in evaluation of neutral objects due to compatibility effect. Unpublished manuscript.

Myung, Jong-yoon, Sheila Blumstein & Julie Sedivy. 2006. Playing on the typewriter, typing on the piano: Manipulation knowledge of objects. *Cognition*

98, 223–243.

Neumann, Roland & Fritz Strack. 2000. Approach and avoidance: The influence of proprioceptive and exteroceptive cues on encoding of affective information. *Journal of Personality and Social Psychology* 79, 39–48.

Neumann, Roland, Jens Förster & Fritz Strack. 2003. Motor compatibility: The bidirectional link between behavior and evaluation. In Jochen Musch & Karl C. Klauer (eds.), *The Psychology of Evaluation. Affective Processes in Cognition and Emotion,* 371–391. Mahwah, NJ: Lawrence Erlbaum Associates.

Niedenthal, Paula M. 2007. Embodying emotion. *Science* 316, 1002–1005.

Niedenthal, Paula M., Lawrence W. Barsalou, Piotr Winkielman, Silvia Krauth-Gruber & François Ric. 2005. Embodiment in attitudes, social perception, and emotion. *Personality and Social Psychology Review* 9, 184–211.

Pecher, Diane, Inge Boot & Saskia Van Dantzig. 2011. Abstract concepts: Sensory-motor grounding, metaphors, and beyond. In Brian H. Ross (ed.), *The Psychology of Learning and Motivation*, 217–248. Burlington: Academic Press.

Pulvermüller, Friedemann, Olaf Hauk, Vadim V. Nikulin & Risto J. Ilmoniemi. 2005a. Functional links between motor and language systems. *European Journal of Neuroscience* 21, 793–797.

Pulvermüller, Friedemann, Yury Shtyrov & Rosti J. Ilmoniemi. 2005b. Brain signatures of meaning access in action word recognition. *Journal of Cognitive Neuroscience* 17, 884–892.

Reber, Rolf, Piotr Winkielman & Norbert Schwarz. 1998. Effects of perceptual fluency on affective judgments. *Psychological Science* 9, 45–48.

Rizzolatti, Giacomo, Luciano Fadiga, Vittorio Gallese & Leonardo Fogassi. 1996. Premotor cortex and the recognition of motor actions. *Cognitive Brain Research* 3, 131–141.

Rueschemeyer, Shirley-Ann, Oliver Lindemann, Daan van Rooij, Wessel van Dam & Harold Bekkering. 2010. Effects of intentional motor actions on embodied language processing. *Experimental Psychology* 574, 260–266.

Scorolli, Claudia & Anna M. Borghi. 2007. Sentence comprehension and action: Effector specific modulation of the motor system. *Brain Research* 1130, 119–124.

Spivey, Michael, Melinda Tyler, Daniel Richardson & Ezekiel Young. 2000. Eye movements during comprehension of spoken scene descriptions. In *Proceedings of the 22nd Annual Conference of the Cognitive Science Society,* 487–492. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Strack, Fritz, Leonard L. Martin & Sabine Stepper. 1988. Inhibiting and facilitating condition of facial expressions: A non-obtrusive test of the facial feedback hypothesis. *Journal of Personality and Social Psychology* 54, 768–777.

Tettamanti, Marco, Giovanni Buccino, Maria Saccuman, Vittorio Gallese, Massimo Danna, Paola Scifo, *et al.* 2005. Listening to action-related sentences activates fronto-parietal motor circuits. *Journal of Cognitive Neuroscience* 172, 273–281.

Tucker, Mike & Rob Ellis. 1998. On the relations between seen objects and components of potential actions. *Journal of Experimental Psychology: Human Perception and Performance* 24, 830–846.

Tucker, Mike & Rob Ellis. 2004. Action priming by briefly presented objects. *Acta*

Psychologica 116, 185–203.

Vermeulen, Nicholas, Paula M. Niedenthal & Olivier Luminet. 2007. Switching between sensory and affective systems incurs processing costs. *Cognitive Science* 31, 183–192.

Wells, Gary L. & Richard E. Petty. 1980. The effects of overt head movements on persuasion: Compatibility and incompatibility of responses. *Basic and Applied Social Psychology* 1, 219–230.

Whittlesea, Bruce W. A. 1993. Illusions of familiarity. *Journal of Experimental Psychology: Learning, Memory, & Cognition* 19, 1235–1253.

Willems, Roel M., Peter Hagoort & Daniel Casasanto. 2010. Body-specific representations of action verbs: Neural evidence from right- and left-handers. *Psychological Science* 21, 67–74.

Winkielman, Piotr & John T. Cacioppo. 2001. Mind at ease puts a smile on the face: Psychophysiological evidence that processing facilitation increases positive affect. *Journal of Personality and Social Psychology* 81, 989–1000.

Winkielman, Piotr, Norbert Schwarz & Andrzej Nowak. 2002. Affect and processing dynamics: Perceptual fluency enhances evaluations. In Simon C. Moore & Mike Oaksford (eds.), *Advances in Consciousness Research* 44, 111–135. Amsterdam: John Benjamins.

Winkielman, Piotr, Paula M. Niedenthal & Lindsay Oberman. 2008. The embodied emotional mind. In Gün R. Semin & Eliot R. Smith (eds.), *Embodied Grounding: Social, Cognitive, Affective, and Neuroscientific Approaches*, 263–288. New York: Cambridge University Press.

Zwaan, Rolf A. 2004. The immersed experiencer: Toward an embodied theory of language comprehension. In Brian H. Ross (ed.), *The Psychology of Learning and Motivation*, vol. 44, pp. 35–62. New York: Academic Press.

Zwaan, Rolf A. 2009. Mental simulation in language comprehension and social cognition. *European Journal of Social Psychology* 7, 1142–1150.

Audrey Milhau, Loïc Heurley, Thibaut Brouillet, and Denis Brouillet
Université Paul Valéry
Epsylon Laboratory
Site Saint Charles
34000 Montpellier
France

audrey.milhau@gmail.com, heurleyloic@yahoo.fr, thibaut.brouillet@gmail.com, and d.brouillet@yahoo.fr

The Human-Fostered Gorilla Koko Shows Breath Control in Play with Wind Instruments

Marcus Perlman, Francine G. Patterson & Ronald H. Cohn

Breath control is critical to the production of spoken language and commonly postulated as a unique human adaptation specifically for this function. In contrast, non-human primates are often assumed to lack volitional control over their vocalizations, and implicitly, their breath. This paper takes an embodied perspective on the development of breath control in a human-fostered gorilla, examining her sound play with musical wind instruments. The subject Koko was video recorded in her play with plastic recorders, harmonicas and whistles. The results show that Koko exercises volitional control over her breath during instrument play. More generally, the findings suggest that all great apes share the potential to develop breath control, and that the original adaptive value of breath control was its flexible development for the service of behaviors that happened to be useful within particular sociocultural and physical environments.

Keywords: breath control; embodiment; human-fostered ape; language evolution; non-human primate

1. Introduction

The notion of embodiment in cognitive science refers generally to the theory that, in order to understand the cognition of a given organism (whether human, gorilla, or octopus), we must take into account the nature of the organism's body as it is situated within its everyday physical and social environment (Gibbs 2006). Cognition, according to this idea, does not happen in a computational vacuum, and neither does it evolve or develop in one, but rather it is distributed across the boundaries of brain, body and environment, and emerges through situated activity. Thus, cognition is understood to be ecological, and its science requires that we seek to understand the adaptive interactions that are created between the biological organism of interest and its particular environment. Stemming from this idea, Johnson (2010: 588) suggests that, "[e]mbodied models [...] take cognition

---

We are grateful to Joanne Tanner, Raymond Gibbs, Nathaniel Clark, and two anonymous reviewers for their helpful comments on previous versions of this article. We would also like to express our appreciation to the Gorilla Foundation caregiving staff, present and past. And finally, we are forever indebted to the gorillas Koko, Ndume, and Michael.

biolinguistics

ISSN 1450–341
Biolinguistics 6.3–4: 433–444, 2012
http://www.biolinguistics.eu

as developing through engagement with the world, and so motivate ontogenetic analyses of what bodies do under varying environmental constraints". The purpose of this paper is to apply such an embodied perspective to a case examination of breath control by a western lowland gorilla.

The ability to voluntarily and skillfully employ one's breath is critical to human speech, and is often considered to be an innate biological adaptation that humans acquired specifically for functions related to language and possibly song (MacLarnon & Hewitt 1999; Fitch 2010). Yet, when humans engage with their typical world, they immerse themselves in a variety of activities related to the instrumental control of their breath and vocal apparatus that extend well beyond the nonetheless pervasive activities of talking and singing. In many North American families, for example, children grow up blowing out birthday candles, learning to whistle, learning breathing patterns for swimming, karate, or relaxation, holding their breath underwater, spitting, blowing spit balls, blowing up balloons and paper bags, imitating animals, trains, and flatulence, and so on.

In comparison, outside of their vocal behavior, we have little direct empirical knowledge of the breath-related abilities of great apes and other non-human primates. What we do know about their ability to control their breath is largely inferred from statements about their vocal behavior, which is often believed to be subcortical, involuntary and inflexible, with little capacity for learning (Pinker 1994; MacLarnon & Hewitt 1999; Corballis 2003; Premack 2004; Call & Tomasello 2007; Pollick & De Waal 2007; Arbib *et al*. 2008; Tomasello 2008). Often implicit in this perspective is the idea that the ability of an animal to control its vocalization and breathing is innate, either present or absent in a species, independent of environmental conditions. Humans have cortical, voluntary, and flexible control over these functions; great apes do not. However, as an empirical claim, this position does not account for regular environmental differences in their respectively typical developmental circumstances. This qualification is especially true with respect to breath control. Humans, in their usual physical and social environments, are afforded a variety of motivated opportunities to learn and flexibly use novel breathing-related behaviors. Yet we know little about the ability of a great ape to flexibly control its breath in comparable environments when such behavior is similarly motivated and reinforced.

One of the few empirical studies to directly investigate breath control in non-human apes focused primarily on the morphology of skeletons, rather than the behavior of living animals. MacLarnon & Hewitt (1999) compared the sizes of thoracic vertebral canals of humans and other modern apes, along with fossils of extinct hominids. The thoracic canal is of particular interest to the evolution of breath control, as its size poses a constraint on innervation to the intercostal and abdominal muscles that humans use to control breathing during speech. The authors conclude from their comparison that 'enhanced' breath control was not possible for ancestral hominids until at least 1.6 million years ago, around the time of *Homo ergaster* or early *Homo erectus*. They suggest that, "[i]t is not known whether non-human primates could be trained to take on a more human pattern of exhalations, but there is no evidence from their habitual vocalizations that they would be capable of doing so" (MacLarnon & Hewitt 1999: 358). So what *do* ape vocalizations reveal about their ability to control their breath?

Historically it was believed that ape vocal behavior — and hence the breathing that underlies it — was emotional and stimulus-bound, and not under volitional control. Even Jane Goodall, after many years observing the chimpanzees at the Gombe Reserve, came to the conclusion that, "the production of sound in the absence of the appropriate emotional state seems to be an almost impossible task for a chimpanzee" (Goodall 1986: 125). At one time, in fact, ape vocal behavior was believed to be so intractable that it was not subject to contingent control by operant conditioning (Skinner 1957). Today, however, Skinner's claim is recognized to be false, with counter evidence from a wide variety of primate taxa (see Pierce 1985 and Owren *et al*. 2011 for reviews). Increasing evidence also shows that apes are naturally flexible in their deployment of vocalizations — for example, suppressing vocalizations in certain situations when it is advantageous (Goodall 1986) or producing them only around particular audiences (Laporte & Zuberbühler 2010). In addition, a growing number of reports describe the use of novel and learned vocalizations and sounds by apes in different environmental circumstances, both in natural environments and in captive environments with and without human enculturation.

One example of learned vocal behavior in natural environments is reported in orangutans, which sometimes produce a lip sputter during an evening nest building routine (van Schaik *et al*. 2006). Critically, only certain populations produce the sound, and it is produced differently between those that do, being used at different phases of the nest building procedure. Thus the orangutan lip sputter appears to be acquired through social transmission and may be considered as a cultural tradition. Another oral sound tradition documented in free-ranging orangutans relates to a sound called a kiss-squeak, which is created by a sharp intake of air through pursed lips (although it is not clear the extent to which this sound reveals control over breathing per se versus the oral manipulation of air). Although this particular sound appears to be produced as a nearly universal part of the orangutan repertoire, in some populations, it is modified in particular ways — covering the mouth with either hands or leaves—that are maintained as cultural traditions. Owren *et al*. (2011) assess novel behaviors like the orangutan's lip sputter and kiss squeak as indicating "rudimentary volitional control over airflow".

Great apes in captivity have also been observed to produce learned vocalizations, even in cases without human enculturation. One study observed the spread of a lip sputter (also known as a 'raspberry' or 'Bronx cheer') element into the pant hoot sequence of a group of captive chimpanzees, originally produced by one member of the group, and eventually spreading to six (Marshall *et al*. 1999). Another set of studies found that chimpanzees in captivity sometimes use novel vocalizations and sounds as attention-getters with human caregivers, including a lip sputter and an extended modification of a voiced food grunt (Hopkins *et al*. 2007; Leavens *et al*. 2010).

Of particular relevance to the ontogenetic perspective taken here, there is also evidence that in environments with extensive human contact, apes can learn new vocal and breath-related behaviors from their human caregivers and companions. One such case concerns the zoo-born orangutan Bonnie, who acquired the ability to whistle from observing human caregivers (Wich *et al*. 2007). Bonnie, who was 30 years old at the time of the report, has been documented from the age of 13 years to whistle regularly, often apparently for her own amusement. Formal study showed that Bonnie would whistle readily on request, as well as spontaneously on her own, and was also able to imitate the whistle duration and number of repetitions produced by a human model. There are also several anecdotal reports of apes learning to smoke cigarettes. The apes in these cases appear to enjoy the activity and show a tendency to develop a habit (Witmer 1909; Kearton 1925; and see a recently publicized example of a smoking orangutan in a zoo in Malaysia at http://www.youtube.com/watch?v=B-h_JlpdRJQ).

In a few cases, great apes have developed notable breath control as a result of efforts to teach them to speak. A prominent example is the young human-fostered chimpanzee Viki, who underwent an intensive training regimen conducted by her 'foster parents', the psychologists C. and K. Hayes (Hayes 1951). From a traditional point of view, the project is remarkable for what is considered to be its abject failure. As critics like Pinker (1994) point out, Viki never succeeded in pronouncing more than four words: *mama*, *papa*, *cup*, and *up*, spoken with an unvoiced vowel described as a harsh stage whisper. Nevertheless Viki did succeed in achieving some degree of flexible control over her breath and vocal tract. Beginning at the age of 4 months, Viki was subject to an operant conditioning procedure, which usually took place at the dinner table and used food as reinforcement. By 15 months, Viki had learned to articulate the word *mama* on her own, and once acquiring this initial skill to vocalize at will, she was able to learn new speech sounds and words more easily by imitation. In addition to the speech sounds that comprised her 'words', /m/, /p/, /k/ and her vowel, Viki also learned to produce various other word-like sounds (e.g. *blook*, *bloo*), a Bronx cheer, and could blow a whistle as part of an imitation task.

Somewhat less accomplished but comparable results were achieved with a chimpanzee trained by Garner (1900), a chimpanzee described by Witmer (1909), an orangutan trained by Furness (1916), and, more recently, an orangutan taught by Laidler (1980). And although not involving explicit training to speak, the learned vocal repertoire of Kanzi, a bonobo raised from infancy with extensive human enculturation and immersion in symbolic communication, offers a more recent example (Hopkins & Savage-Rumbaugh 1991; Taglialatela *et al*. 2003). Contextual and acoustic analysis shows that Kanzi regularly made use of four novel peep vocalizations, each used to communicate in distinct semantic contexts.

Finally, the subject of the present study, Koko, a human-fostered gorilla, has been video-recorded in the performance of a variety of behaviors that appear to involve voluntary breath control (Perlman *et al*. in preparation). Video records show her huffing and grunting into a telephone, huffing on the lenses of eyeglasses, performing a fake cough, blowing her nose, performing her version of a 'raspberry' (folding her tongue lengthwise and blowing air through it), and blowing into her hand as a communicative gesture. When agitated at a visitor or caregiver, Koko sometimes performs a breath-related gesture known as *you blew it*, in which she expels air forcefully towards the face of the transgressor. More frequently, Koko performs a more welcoming routine with visitors and caregivers called the *blow test*. In an interactive exchange, Koko blows gently towards her interlocutor and smells their breath as they blow back.

Altogether these various reports suggest that under certain circumstances apes are able to develop flexible control over their breathing, and that they can acquire this control towards the performance of instrumental behaviors such as producing attention-getting vocalizations and socially transmitted, learned behaviors like whistling. Although Viki in particular was subjected to intensive operant conditioning without substantial progress towards actually learning to speak, the point is often neglected that she did acquire more general volitional control over her breathing and sound production.

These previous reports reveal the development of breath control in chimpanzees, a bonobo, and orangutans, yet studies have not addressed whether gorillas might acquire a similar level of volitional control. However, given their place in the great ape family, it is reasonable to expect gorillas to exhibit comparable potential for breath control, which would imply that the behavioral and neural basis to develop breath control is a general trait of the great apes. In support of this reasoning, the present paper reports on a series of video-recorded observations of the human-fostered gorilla Koko as she exercises breath control during play with musical wind instruments.

2. Method

2.1. *Subject*

The subject of the study is Koko, a female western lowland gorilla (*Gorilla gorilla gorilla*) who was 37 to 39 years of age during the reported observations. Koko was born in 1971 at the San Francisco Zoo, but became ill at six months and was moved from the zoo's gorilla enclosure to be cared for by humans and nursed back to health (Patterson & Linden 1981). At the age of one year, she came under the care of the second and third authors (FGP, RHC). At this time, Koko began lifelong tutelage in a sign system derived from American Sign Language, as well as immersion in spoken English. Over the course of her life, play with musical instruments has been a common interactive activity between Koko and her caregivers, and as such, one that has been encouraged and rewarded. In general, Koko's novel breath-related and vocal behaviors have been subject to demonstration, molding and various forms of reinforcement including food and verbal praise, but have not been specifically trained by operant conditioning procedures.

2.2. *Data Recording*

The data come from video recordings made during regular daily sessions involving Koko, FGP and RHC. In most of the recordings, Koko interacted most directly with FGP while RHC operated the camera. On a few occasions, family or visitors are present, although these occasions did not happen to be part of the present report. Similar video recorded sessions have been a more-or-less regular event since the project's beginning in 1972 and are an ordinary part of Koko's daily routine. In general, the video was taken without a specific research project in mind, but does reflect a broader effort to document Koko's notable abilities, communications, and life landmarks. The present study focuses on video recordings from July 2007 through December 2010.

2.3. *Analysis*

The first author (MP) viewed the video corpus and identified all instances of Koko interacting with a musical wind instrument. It was known in advance of this search that Koko commonly participated in play with wind instruments, and that numerous instances had been recorded in the corpus. Audio .wav files were extracted from the positively identified video clips, and Praat phonetic analysis software (Boersma 2001) was used to determine the onset and offset of each audible manipulation with an instrument. These times were used to compute the duration of each sound and the intervals between them.

Koko was observed to perform three relevant types of behaviors during her instrument play, and each instance was classified accordingly. The majority of the time she blew into the instrument with the result that it produced a tone, an event referred to as a toot. On some other occasions, she produced blows, instances in which she blew into or on the instrument, but without producing a tone. And finally, she sometimes produced markedly short tones that appeared to be produced by oral manipulation (e.g., sucking) but without any clear breath control. These cases were operationalized as any tone produced with a duration less than 0.26 seconds, and were excluded from further analysis. This operating definition was supported by a bimodal distribution of toot durations, with none occurring within the inclusive range of 0.26 to 0.30 seconds (i.e. orally performed toots were generally shorter than 0.26s while true toots were longer than 0.30s; Table 1). We generally did not see evidence that Koko selected particular notes in her play.

Koko's instrument play was divided into bouts, which were defined as a series of toots that were separated by no more than 30 minutes from another toot. Bouts were further divided into sequences, defined as a series of toots separated by less than 4.23 seconds from another toot. This cutoff point was selected as what appeared to be a natural dividing point for Koko: a local minimum in a bimodal distribution of inter-toot intervals (see Table 1).

Each bout was coded for the circumstances of its initiation — either self-initiated by Koko or encouraged by FGP and RHC — and its consequences, resulting in a verbal response, food, or no interactive response at all. During each bout, RHC was generally operating the video camera and directing it at Koko, but this act in itself was not considered in coding initiation or consequence. Initiations were determined by examining the clip for 30 seconds immediately preceding the bout, and consequences were assessed during the play bout and extending for 30 seconds immediately following it.

Toot Duration(s)	≤ 0.33	0.33–0.67	0.67–1.00	1.00–1.33	1.33–2.00	/0.26–0.30/
Number of Toots	16	70	37	21	6	4/0/3
Inter-Toot Interval(s)	≤ 1.5	1.5–3.0	3.0–4.5	4.5–7.5	≥ 7.5	/3.08–4.23/
Number of Intervals	113	11	2	6	12	8/0/4

Table 1: Toot Durations and Inter-toot Intervals. Note that the times are in seconds. The final column shows the range in which the local minimum was established: 0 for the specified range, along with the count for the preceding and subsequent ranges of the same size (used to distinguish oral manipulations and sequences)

3. Results

In total, Koko was observed to perform 137 individual toots (84% of total manipulations) over the course of 38 sequences and 17 different bouts, with only two bouts occurring on the same day (See Table 2). In addition, she produced 12 (7%) blows and 14 (9%) excluded instances of oral manipulations. Koko most commonly played with one of several plastic recorders, but also used other instruments including harmonicas and party favor whistles.

Instrument	Bouts	Sequences	Toots
Recorder	12	29	111
Harmonica	2	6	20
Party Favor	3	3	6
Total	17	38	137

Table 2: Number of Bouts, Sequences, and Toots by Instrument Type

Table 3 shows the counts for how each bout was initiated and its consequence. Of the 17 bouts, only 11 provided the 30 seconds of preceding video footage necessary to determine its initiation (specifically to verify that it was self-initiated). Of these 11, 6 were self-initiated compared to 5 that were encouraged. 13 of the 17 bouts were videotaped with the necessary 30 seconds of subsequent footage, and of these, 8 resulted in some kind of verbal response, 2 in food, and 3 received no apparent acknowledgment.

Initiation			Consequence		
Self	Encouraged	Unknown	Verbal	Food	None
6	5	6	12 (8)	2	3

Table 3: Number of Bouts by Initiation and Consequence. Note that food also implies a verbal consequence. For verbal consequence, the parentheses indicate the number of instances verified for a full 30 seconds after the bout

Table 4 shows the characteristics of Koko's musical bouts, sequences, and toots. On average, a bout lasted for 26.7 seconds (SD = 40.0), and individual sequences lasted for 6.3 seconds (SD = 5.64), with an average of 3.61 toots per sequence (SD = 2.68). Toot durations lasted on average 0.72 seconds (SD = 0.21) with a mean inter-toot interval of 1.06 seconds (SD = 0.18).

In addition, we were interested in Koko's rate of toots and blows for each sequence. Since sequence durations were calculated from the onset of Koko's initial toot, they failed to include the initial inhalation phase. Thus an adjusted sequence duration was calculated by subtracting the duration of the initial toot from the total duration of the sequence, thereby including only complete inhalation-exhalation breath cycles. (Consequently, the six single-blow sequences were not included in this calculation.) In sequences interrupted by oral manipulations, the longest continuous portion of toots and blows from the sequence was used. From this adjusted duration, an extrapolated rate of blow cycles per minute was calculated for each sequence.

On average, Koko blew into the instrument at a rate of 36.14 cycles per minute (SD = 8.71). A critical question is whether this rate varies from Koko's usual breathing rate, as this would indicate that her breathing was under voluntary control. In general, Koko was stationary during her instrument play and did not exceed a mild level of physical activity. For comparison, we estimated Koko's breathing rate just after mild to moderate activity on eight occasions in September and October 2011, and found the range to be from 16 to 20 breaths per minute (BPM). A rate of 25 BPM was selected as a conservatively high estimate, and a single sample t-test showed that the rate of blow cycles per minute was significantly faster than this baseline breathing rate, $t(31) = 7.24$, $p < 0.0001$.

Last, we examined the intensity of Koko's 12 blows with the question of whether they showed evidence of increased forcefulness over her baseline breathing intensity. In general, the energy of the blows was clearly audible during audio-only playback, as well as visible in a spectrogram, indicating a clear elevation in sound intensity from background noise. Koko's normal breathing, in contrast, was generally not detectable under the recording conditions, either audibly or by spectrogram. Intensity measurements were taken at the onset and offset of the blow, and these were averaged together as a baseline intensity for comparison to the blow's mean and maximum intensity. The results showed an average increase in mean intensity of 5.12 dB (SD = 3.99) and an average maximum intensity of 9.79 dB over baseline (SD = 4.67).

	Recorder		Harmonica		Party Favor		Total	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Bout duration(s)	41.80	53.10	26.87	NA	4.68	2.66	26.66	39.99
Sequence duration(s)	6.28	5.40	7.19	8.16	4.68	2.66	6.29	5.64
Toots per sequence	3.83	2.76	3.33	2.88	2.00	1.73	3.61	2.68
Toot duration(s)	0.67	0.16	0.89	0.21	0.76	0.09	0.72	0.18
Inter-toot interval(s)	1.02	0.34	1.27	0.49	1.05	0.25	1.06	0.36
Blow cycles/minute	38.09	8.48	29.79	8.19	31.17	5.53	36.14	8.71

Table 4: Characteristics of Bouts, Sequences and Toots

4. Discussion

We analyzed video recordings of the human-fostered gorilla Koko performing 38 sequences of play with musical wind instruments from 17 different bouts. Presumably due to an environment more comparable to human children learning to play a musical instrument, Koko has developed the volitional ability to use her breath to produce tones with a variety of instruments, including plastic recorders, harmonicas, and party-favor style whistles. The analysis shows that when Koko blows on an instrument, she tends to adopt a pattern of breathing that is significantly faster than her normal rate. Moreover, although difficult to verify, it would seem that the air pressure necessary to sound a tone would also require Koko to produce extra forceful breaths while playing. This point is supported by the few occasions when she directed blows at the instrument that were not in the proper place or manner to produce a tone. The breaths she exerted in these instances showed a clear increase in sound energy over the background noise, whereas her normal breathing was not generally detectable. Koko thus shows the ability to exercise volitional control over the frequency and forcefulness of her breath.

It is true that instrument play is commonly modeled for Koko, and her performance is often encouraged, engendering positive social response, at times including explicit rewards of praise and food. We note, however, that this scenario of positive engagement and reinforcement is roughly similar to many environments in which human children learn to play musical instruments and perform other breathing-related behaviors. This point is critical from the ontogenetic perspective of embodiment, according to which breath control is understood to develop ecologically, in contexts in which it is useful and relevant. Furthermore, several aspects of Koko's instrument play point to an intrinsic interest she has gained in the activity, showing it to be more than a reward-contingent trick. Koko is often inclined to play when the opportunity arises (e.g., when she is handed an instrument), and the video recorded bouts show that she often sounded her instrument repeatedly without any visible expectation of reward, sometimes for more than a minute at a time. Additionally, Koko herself initiated the majority of sessions, received food rewards for only a small proportion, and in a few cases, appeared to blow on the instrument without receiving explicit social acknowledgment (aside from video recording), possibly for her own amusement. In any case, it is clear that Koko understands the different instruments' sound making affordances, and is fully capable of taking advantage of them when she wishes.

Fitch (2006) raises the question of whether increased breath control evolved specifically in the adaptive service of speech, or whether it may have evolved first for other reasons — he offers the examples of prolonged running or swimming — and thereby was available as a pre-adaptation for spoken language. Koko's instrument play is the first documented case of breath control in a gorilla, but it contributes to a growing literature that now spans reports of voluntary breath and vocal control across the great apes. The converging consensus is that the great apes, as a family, possess the potential to control their breath and that this ability develops flexibly in contexts when it is motivated and useful. Thus breath and vocal control does not arise as a monolithic skill that an organism either does or does not possess, but instead it is acquired and practiced within an instrumental, purposeful context.

From the present report and the literature reviewed above, we see a few different contexts that appear to motivate the development of breath and vocal control. With captive chimpanzees (Hopkins *et al*. 2007; Leavens *et al*. 2010), we see the instrumental use of innovative airflow and vocal sound as an attention getter with human caregivers. In a few cases with intensive human interaction, novel, more semantically specific communicative signals are developed and learned, as with the human-fostered bonobo Kanzi's untrained modifications of peep calls (Hopkins & Savage-Rumbaugh 1991; Taglialatela *et al*. 2003) and the chimpanzee Viki's similarly sized, trained vocabulary of 'words' (Hayes 1951). And a more vague, culturally determined relevance motivates the spread of several behaviors that are seemingly more arbitrary in function. These behaviors include, for example, the nest building lip sputters of wild orangutans (van Schaik *et al*. 2003) and the raspberries incorporated into chimpanzee pant hoots (Marshall *et al*. 1999). These behaviors are meaningful in their particular sociocultural context, but otherwise do not appear to serve any clear instrumental function. The case study of Bonnie the whistling orangutan (Wich *et al.* 2009) and the present study of Koko's wind instrument play may also fall into this category. However these activities also seem to provide intrinsic entertainment for Bonnie and Koko (perhaps similar to music in humans), and are performed more independently of any specific cultural routine.

Taken together, these various research findings suggest that the adaptive value of breath control among the great apes lies in its flexible development in the service of behaviors that are relevant and motivated within a particular environment. While one may argue that many observed instances take place within an 'artificial' human context, there are at least a few observations of flexible breath-related behavior in free-ranging apes. Additionally, it is important to note that the current cross section of observational time is extremely narrow in comparison to the millions of years of adaptive contexts and social traditions that may have cycled in and out of ape populations over their histories. Over the course of great ape cultures, it seems probable that a multitude of breath-related behaviors have come in and out of fashion.

Thus, from the present embodied, ecological perspective, Fitch's question of whether breath control evolved specifically to support speech versus some other specific behavior neglects the possibility that breath control is adaptive for its flexibility, rather than for any function (or set of functions) in particular. Indeed, this point is supported by the wide variety of behaviors involving breath control that are pervasive in human cultures, spanning children's play, smoking, spitting, whistling, vocal imitation, breathing patterns in activities like swimming, diving, running, and spiritual rituals, and of course, talking and making music. Clearly, humans show an extraordinary ability to hone their breath control into a dexterous and finely tuned instrument, serviceable for a number of culturally determined functions. Mounting evidence suggests that we are not fully unique in this respect, and that our great ape relatives share with us at least a rudimentary basis for this flexibility in their breathing and vocal behavior.

References

Arbib, Michael A., Katja Liebal & Simone Pika. Primate vocalization, gesture and the evolution of human language. *Current Anthropology, 49,* 1053–1076.

Boersma, Paul. 2001. Praat, a system for doing phonetics by computer. *Glot International* 5, 341–345.

Call, Josep & Michael Tomasello. 2007. Introduction: Intentional communication in nonhuman primates. In Josep Call & Michael Tomasello (Eds.), *The Gestural Communication of Apes and Monkeys,* 1–16. Mahwah, NJ: Lawrence Erlbaum.

Corballis, Michael C. 2003. From mouth to hand: Gesture, speech, and the evolution of right-handedness. *Behavioral and Brain Sciences, 26,* 199–260.

Fitch, W. Tecumseh. 2006. Production of vocalizations in mammals. In Keith Brown (ed.), *Encyclopedia of Language and Linguistics*, 115–121. Oxford: Oxford University Press.

Fitch, W. Tecumseh. 2010. *The Evolution of Language.* Cambridge: Cambridge University Press.

Furness, William H. 1916. Observations on the mentality of chimpanzees and orangutans. *Proceedings of the American Philosophical Society* 55, 281–290.

Garner, Richard. 1900. *Apes and Monkeys: Their Life and Language.* London: Ginn.

Gibbs, Raymond W., Jr. 2006. *Embodiment and Cognitive Science*. Cambridge: Cambridge University Press.

Goodall, Jane. 1986. *The Chimpanzees of Gombe: Patterns of Behavior.* Cambridge, MA: Harvard University Press.

Hayes, Cathy. 1951. *The Ape in Our House.* New York: Harper.

Hopkins, William D. & E. Sue Savage-Rumbaugh. 1991. Vocal communication as a function of differential rearing experiences in *Pan paniscus*: A preliminary report. *International Journal of Primatology* 12, 559–583.

Hopkins, William. D., Jared Taglialatela & David A. Leavens. 2007. Chimpanzees differentially produce novel vocalizations to capture the attention of a human. *Animal Behaviour,* 73, 281–286.

Johnson, Christine M. 2010. Observing cognitive complexity in primates and cetaceans. *International Journal of Comparative Psychology* 23, 587–624.

Kearton, Cherry. 1925. *My Friend Toto: The Adventures of a Chimpanzee, and the Story of His Journey from the Congo to London*. London: Arrowsmith.

Laidler, Keith. 1980. *The Talking Ape.* New York: Stein and Day.

Laporte, Marion N. C. & Klaus Zuberbühler. 2010. Vocal greeting behaviour in wild chimpanzee females. *Animal Behaviour* 80, 467–473.

Leavens, David A., Jamie L. Russell & William D. Hopkins. 2010. Multimodal communication by captive chimpanzees (*Pan troglodytes*). *Animal Cognition* 13, 33–40.

MacLarnon, Ann M. & Gwen P. Hewitt. 1999. The evolution of human speech: The role of enhanced breathing control. *American Journal of Physical Anthropology* 109, 341–363.

Marshall, Andrew J., Richard W. Wrangham & Adam Clark Arcadi. 1999. Does learning affect the structure of vocalizations in chimpanzees? *Animal Behaviour* 58, 825–830.

Owren, Michael J., R. Toby Amoss & Drew Rendall. 2011. Two organizing principles of vocal production: Implications for nonhuman and human primates. *American Journal of Primatology* 73, 530–544.

Patterson, Francine G. & Eugene Linden. 1981. *The Education of Koko.* New York: Holt, Rinhart & Winston.

Pierce, John D., Jr. 1985. A review of attempts to condition operantly alloprimate vocalizations. *Primates* 26, 202–213.

Pinker, Steven. 1994. *The Language Instinct*. New York: William Morrow and Co.

Pollick, Amy S. & Frans B. M. de Waal. 2007. *Proceedings of the National Academy of Sciences, 104*, 8184-8189.

Premack, David. 2004. Is language the key to human intelligence? *Science* 303, 318–320.

Skinner, B. F. 1957. *Verbal Behavior*. New York: Appleton-Century-Crofts.

Taglilatela, Jared P., Sue Savage-Rumbaugh & Lauren A. Baker. 2003. Vocal production by a language-competent *Pan paniscus*. *International Journal of Primatology* 24, 1–17.

Tomasello, Michael. 2008. *Origins of Human Communication*. Cambridge, MA: MIT Press.

van Schaik, Carel P., Marc Ancrenaz, Gwendolyn Borgen, Birtue Galdikas, Cheryl D. Knott, Ian Singleton, Akira Suzuki, Sri Suci Utami & Michelle Merrill. 2003. Orangutan cultures and the evolution of material culture. *Science* 299, 102–105.

Wich, Serge A., Karyl B. Swartz, Madeleine E. Hardus, Adriano R. Lameira, Erin Stromberg & Robert W. Shumaker. 2009. A case of spontaneous acquisition of a human sound by an orangutan. *Primates* 50, 56–64.

Witmer, Lightner. 1909. A monkey with a mind. *The Psychological Clinic III*, 179–205.

Marcus Perlman
University of California, Merced
Department of Information and Cognitive Sciences
5200 North Lake Rd.
Merced, California 95343
USA
mperlman@ucmerced.edu

Francine G. Patterson & Ronald H. Cohn
The Gorilla Foundation

1732 Woodside Road
Redwood City, California 95040
USA
research@koko.org

Three Ways to Bridge the Gap between Perception and Action, and Language

Jean-Luc Petit

A rallying cry in some sectors of cognitive science, the embodiment of language is understood here in the full content of meaning of phenomenological tradition to help assess the remaining distance from neuroscience to a science of language, provided that tracking down in the brain neural events correlative of verbal behavior would not be sufficient. From an eidetic standpoint, one must build the transition between perceptive, pragmatic and semantic morphologies. From the point of view of subjective experience, one must understand how it is possible that we move from our sensory and kinaesthetic experiences to verbal expressions of a sense that could be shared by others. That is why, in order to prevent neglect of any dimension of embodiment of language, we would rather plead for a threefold approach than concede that the current naturalistic mode is the only possible.

Keywords: embodiment of language; kinaesthesia as operator of constitution; meaning morphologies and morphodynamics

1. The 'Embodiment of Language': A Spectrum of Possible Meanings

The expression 'embodiment of language' refers to a new trend of research on the neural basis of language. Apart from a convenient label to bring together different research teams, this expression suggests that there is a special relationship between language and the body. The philosophical observer may wonder whether he will find in it the traditional philosophical problem of the incarnation of mind. But, as it would be foolish to project interpretations of a philosopher on empirical science, we should first get familiar with the use of 'embodiment' by the researchers themselves. If we try to consider all the contributions to the investigation of the brain bases of language without omitting any of the protagonists in the ongoing controversy (cf. Dinstein *et al*. 2008; Hickok 2008; Lotto *et al*. 2008; Mahon & Caramazza 2008; Lingnau *et al*. 2009; Scott *et al*. 2009), we will find a spectrum of, at first sight not very homogeneous, uses: Are these different uses of the term 'embodiment' mutually compatible, and can they be reduced to one unequivocal sense? That is no mere pedantic tinkering aimed at the semantic correctness of scientific discourse. My suspicion is that the fashion

biolinguistics

ISSN 1450–341
Biolinguistics 6.3–4: 445–461, 2012
http://www.biolinguistics.eu

of embodiment makes us underestimate the difficulty of naturalizing the study of language. Rooting language in the body, without worrying too much on how to do it, are we not tempted to believe abolished the distance between our material nature and the logical ideality of meaning, between empirical data and the lived experience of meaning?

1.1. From Modularity to Interaction

After the provocative revival of Gall's Phrenology by Fodor (1983), it looked as though major cognitive functions, such as object or sentence recognition, could be fully carried out by specialized peripheral systems that operate independently of each other without exchanging information, so that the organism, in spite of its anatomic unity, is like a hydra at the cognitive point of view. The contribution of brain imaging in a study previously limited to the observation of deficits due to brain lesions made possible a new approach on the foundations of language. This new approach suggests that its functions are not underpinned by specialized modules but rather by an extensive network of distinct areas of the brain that sustain a permanent dialogue with each other (cf. Wise *et al*. 1991; Démonet *et al*. 1992; Bookheimer 2002). Such research tends assuredly to some form of embodiment: the recovery of the integrative unity of the organism on its fragmentation into multiple modules. Even more interestingly, the recognition of the interdependence between widely distributed brain regions at the basis of language opened the way for revolutionary assumptions about the direction of their mutual interactions during verbal behavior.

1.2. The Motor System Not a Mere Output

The classic model of the bases of language (cf. Lichtheim 1885, Geschwind 1965) strictly subordinates the production of speech sounds by articulators to the cognitive processing of linguistic information. This model limits the contribution of the motor system to the role of slavishly executing a motor program developed elsewhere in auditory areas and in the upper levels of the hierarchy of the cognitive system. This representation of the muscular production of speech rests on the traditional prejudice concerning the body as an instrument of thought. This prejudice is shaken by the discovery of retroactive influences, sometimes modulatory and sometimes formative, performed by the articulator system on auditory reception and semantic interpretation of phonemes and expressions (cf. Gentilucci *et al*. 2001, Fadiga *et al*. 2002). This rehabilitation of the cognitive function of movement in speech amounts to an embodiment of language.

1.3. Binding Doesn't Need Abstract Supramodal Computation

The synthesis of sensory qualities of perceived objects raises, at the level of the neuron or neuronal group, the problem of binding of unimodal signals of different pathways: visual, auditory, olfactory, vestibular, proprioceptive, and visceral in a supramodal concept of the object (cf. von der Malsburg 1995). This integration function is classically delegated to a central cognitive system hierarchically superior to the various sensory systems and exerting an influence on their operation, especially through the orientation of attention (cf. Norman & Shallice 1980).This hypothesis of a purely conceptual, amodal thus disembodied, center is unlikely given the ubiquity of the mixture of influences of the various modalities that occurs at synaptic connections. The introduction of a transversal process of integration between modalities would allow us to dispense with this hypothesis, assuming that the motor system is able to preform perception through action (cf. Rizzolatti *et al*. 1994; Umilta *et al*. 1994; Skipper *et al*. 2007).

1.4. From Mentalese to 'the Language of Neurons'

In the history of cognitive science Chomskyan idea of *competence*, with its strict distinction as to *performance* and its priority over the latter for the study of language has had a founding role (cf. Chomsky 1965). This distinction, and hierarchy resulting from it, tended to assimilate the core structure of the linguistic capability of man to a language of symbolic logic and its implementation to an application of syntactic rules to strings of symbols (a calculus). Once the language of thought has been internalized in this deep structure its realization in acts of communication could only appear as a contingent coating surface structure. The mind's *Mentalese*, following Fodor (1975), limited the contribution of neuroscience to the study of language to the realization of the logical structure of competence in a brain-machine indifferent to its program. A recent alternative to this ideology, the identification of linguistic information processing with neural dynamics itself and its laws of association is yet another form of embodiment of language (cf. Pulvermüller 2002).

1.5. Broca's Area: An All-Purpose Processor of Complexity

Whether converging or diverging, the various trends expressing themselves through the theme of embodiment are represented in the debate on the interpretation of the functions of Broca's area. Traditionally regarded as a center for motor realization of speech at the end of cognitive processing, the contribution of Broca's area was found to take place earlier and to be more complex, since it is recruited at all levels of verbal conduct: for perception as well as for production, for syntactic construction as well as for semantic interpretation (cf. Nishitani *et al*. 2005; Fadiga & Craighero 2006; Tettamanti & Weniger 2006). This redefinition of the linguistic functions of Broca's area coupled with a phylogenetic hypothesis put forward by Rizzolatti & Arbib (1998) about its origins in a monkey's premotor area site of mirror neurons crosses the issue of embodiment: Should we assign this key node in the brain circuits of language to the motor system? Or should we not rather focus on the emancipation of this area from its former utilitarian functions and the acquisition by it of a capacity to process any multimodal cognitive complexity, such as syntactic dependencies between items at non rigid positions whether in the word order or in any sequence, including arbitrary symbols, musical sounds and motor acts parts of a goal-oriented action (cf. Tettamanti *et al*. 2009; Fadiga *et al*. 2010)? The conflict of interpretations on the function of Broca's area shows that concerning the relationship between language and the body empirical research has not resulted in a generally accepted doctrine which may discharge of its responsibilities the philosophical reflection on embodiment, a reflection to which it might be time to return for a much-needed guidance.

2. That Neuroscience Cannot Do Without a Phenomenology of Language

> On ne comprendra jamais ces deux idées à la fois si l'on continue d'osciller entre la notion de «motricité» et celle «d'intelligence», et si l'on ne découvre pas une troisième notion qui permette de les intégrer, une fonction, la même à tous les niveaux, qui soit à l'œuvre aussi bien dans les préparations cachées de la parole que dans les phénomènes articulaires, qui porte tout l'édifice du langage, et qui cependant se stabilise en processus relativement autonomes.[1] (Merleau-Ponty 1945: 228)

One will not fail to note in passing that the well-known promoter of the phenomenology of the body, far from being uniquely concerned with bodily experience, was fully aware of the importance of being clear about the implementation of functions and processes of language. Does recent work on the cerebral bases of language have surpassed this vacillation between motricity and representation? As everyone knows, such work is divided into two schools: Theory of Mind (cf. Premack & Woodruf 1978; Frith & Frith 1999) and Simulation (cf. Goldman 1989, 1992) or Embodiment (cf. Gallese & Goldman 1998; Gallese 2001), although some researchers are leaving the path of controversy for a search for complementarity (cf. Schippers *et al*. 2009, 2010). Is this a confirmation of Merleau-Ponty's diagnosis? His remark would shift from diagnosis to prognosis and even premonition: Should we go that far? The turn taken by empirical research barely helps to fix our ideas.

2.1. Let's Not Replace Questions of Essence with Evolutionary Narratives

The concern of biologists for the evolutionary origins of human capabilities may sometimes cause puzzlement to the philosopher. It may seem natural to think that the primary issue of any inquiry, philosophical, empirical, or otherwise, is the question "What is it?" a question concerning the essence of the thing itself and not its becoming, its origins, its cause, its effects, etc. The transition from a study of language to a study of its neural basis led to a replacement of "What is it?" by "Where did that come from?" even though the story answering the second question does not necessarily provide the definition expected in response to the first (cf. Rizzolatti & Arbib 1998; Corballis 2002, 2004, 2011; Fadiga & Craighero 2007). Moreover, blurring the differences can be detrimental to our understanding: If speaking, hearing and understanding are to be conceived henceforth as species of movement or imitation of movement, surely our concepts of movement and of speech will have to be altered. Will moving still mean moving one's body and will understanding someone still mean knowing what he means?

---

[1] 'We will never understand these two ideas at once if we continue to oscillate between the notions of "motricity" and of "intelligence", and if one does not discover a third concept that allows to integrate both of them, a function, the same at all levels, that is at work both in the hidden preparations of speech and in the articulator processes, which supports the whole structure of language, and yet stabilizes itself in relatively autonomous processes.'

2.2. *Piecemeal Solutions Conceal the Full Extent of the Dilemma*

The shortest way to account for the embodiment of language goes through observation of articulator gestures and hand gestures that accompany or replace speech, as in sign language: These gestures are what is seen in linguistic communication; their neural correlates are the natural candidates for rooting language in the body (cf. Goldin-Meadow 1999; Gallagher & Frith 2004; Corina & Knapp 2006; Montgomery *et al*. 2007; Holle *et al*. 2008; MacSweeney *et al*. 2008; Xu *et al*. 2009; Emmorey *et al*. 2010). If the embodiment of language in all its dimensions has little meaning, an obvious short circuit is to look first in the cartographic representations of the body in sensorimotor cortical areas as mapped by Penfield & Rasmussen (1950) for the potential correlates of the lexicon of action verbs or of action sentences (cf. Hauk *et al*. 2004; Buccino *et al*. 2005; Pulvermüller 2005; Aziz-Zadeh 2006; Pulvermüller & Fadiga 2010). From there, one will address the following problem: whether it is possible to extend what is true for gesture (or the motor repertory) to the entire verbal behavior in the hope that the generalization reveals the uninterrupted passage between phoneme production/perception and construction/interpretation of expressions or sentences. But if gesture is already language, basing language on gesture cannot do much to clarify the nature of language. If action verbs derive their meaning from actions they express, this is not the case in the rest of lexicon. Their metaphorical usage is a semantic innovation that might rather undermine than safeguard the link with action.

2.3. ***The Challenge: Rooting Meaning Morphologies in the Body***

The possibility for human beings to express through linguistic expressions perceived forms of the visual field and goals or affordances of the practical field depends on an underlying mediation between the categories of perception and of action presumably to be performed in the last instance in brain circuits. The abstract principle of the semantic universality of natural language in the sense of Tarski (1936) presupposes such mediation without accounting for it: "If it is possible in general to talk of anything whatsoever in a sensible way, then it is also possible to talk of that thing in everyday language" (p. 170). One must understand step by step how it is possible that any configurations, whether objects of visual attention or goals of intended actions — configurations that emerge and stabilize in a silent experience — are promoted and safeguarded in terms of their expression in linguistic forms. To clarify the transition between the morphologies of different eidetic types (not just the linguistic type) that inform the conduct of agents-observers-speakers it is not sufficient to trace courses of events in brain circuits. The problem of embodiment is not settled at the level of neurons because it is both and inextricably eidetic and psychophysiological.

3. The Phenomenological Tradition and its Deceptive Proxy

Without lapsing into an outdated imperialism, a philosopher may be surprised by what appears to be a revival of the theme of incarnation in the literature on embodiment coupled with a misunderstanding, if not a systematic attribution error, of the original sources of this theme. The fact is that the requirement of thinking together, and inter-relating, bodily experience with understanding the actions of others, action with the perception of the environment, imitation with intersubjective communication, dates back to the phenomenological movement in the early twentieth century (e.g., Lipps, Dilthey, Husserl, Scheler, Stein, Reinach, Heidegger, etc.). However, everything happens as if current neuroscience sought a basically inadequate substitute for this phenomenology in authors who hesitate between behaviorism and cognitivism, between mentalism and physicalism, between computation and simulation (e.g., Liberman, Gibson, Goldman, etc.).

3.1. Liberman, Gibson, Goldman, et alii

Instead of relying on a phenomenology, Liberman's (1985) conception of articulator gesture seeks to frame phenomena in two successive doctrines: (i) a behaviorist concern to assign phonetic units to a coarticulation resistant, recordable movement; (ii) an assignment of phonetic encoding and decoding to a peripheral system, conforming to the modularity of mind doctrine. Gibson's concept of affordance and ecological theory of perception (1977, 1979) resembles a phenomenological description of the morphological structures of *Umwelt* for a living being, but is marred by a physicalism for which the perceptual invariants are due to information actually residing in the optical flow. Recycling Goldman's simulation theory (1989, 1992) as functional interpretation of the brain system of mirror neurons linking observation and execution of actions does not provide the satisfactory alternative that Gallese & Goldman (1998) believe to the theory of the mind that subordinates recognition of others to an inference of the cognitive subject. That is because, insofar as the observer is supposed to have his own motor system objectified as a representation in mind and to use it for predicting the future behavior of an observed agent, this so-called simulation remains a solipsistic process that takes place entirely in the cognitive system of an isolated individual. I'm not specifically trying to prove the reality of these shortcomings because they only make sense and present any seriousness from the phenomenological perspective. Whether they follow the line of Liberman, Gibson, or Goldman, their disciples presumably can live without discomfort with the aspects we just emphasized of their favorite theories. The point is that none of these approaches, despite the attraction they may have for those who seek to embody the language in the body, is likely to meet the requirements of a phenomenology of embodiment of language, a phenomenology in the lack of which, I contend, one will remain stuck midway on the path to embodiment.

3.2. Merleau-Ponty, the One Acceptable Phenomenologist

Merleau-Ponty enjoys in cognitive science a favorable view that is denied to the philosopher from whom he borrows his ideas, namely Husserl, especially in his later texts on the body and intersubjectivity and the world (Husserl 2008, 1973a). This unfair attribution goes so far as to conceal the Husserlian origins of the themes of Merleau-Ponty's phenomenology of perception, despite the fact that he himself made no mystery of their provenance. To counterbalance this trend, it is noteworthy that Merleau-Ponty's assimilation of the own body (*Leib* in contradistinction to *Körper*) with the body schema of neurologists (cf. Head & Holmes 1911) is perhaps not the bridge one might think would lead to the functional somatotopic maps of current neuroscience. That is not so much because it would imply an oversight on his part of the distinction — absolutely essential in phenomenology — between the lived body and the body object of science, as any attempt at finding neurological — or neuropsychological — correlates of our bodily experience cannot but try again crossing the gap to its own risk. It is the very authenticity of the phenomenon that was the basis for his classic description of bodily experience which is at issue. In fact, although this has not received much attention and even if it is surely not enough to tarnish his reputation, Merleau-Ponty's uncritical adherence both to Goldstein's dogma of *Gestalt* and Goldstein and Gelb (1920) observations of a single case: Schneider, a probable simulator warns Goldenberg (2002), casts serious doubt on his description (cf. Petit 2010).

3.3. Husserl, the Founder and Transcendantal Scarecrow

In science philosophers are considered producers of theories to be tested experimentally. But in Husserl the issue is not of theory but of a lived experience — even if it's a thought experiment — which requires a conversion of attitude in anyone who wants to follow suit. That's what is needed to understand a paradox of his phenomenology of language pointed out by Merleau-Ponty (1960): starting from an eidetic science of ideal essences of meaning as a priori norms for any language (*Logische Untersuchungen* IV, 1901/1913), Husserl came to the truly phenomenological point of view of later texts where: «*le langage apparaît comme le corps de la pensée pour le sujet parlant qui use de sa langue comme d'un moyen de communication avec une communauté vivante*»[2] (pp. 106–107). That is because we are invited by Husserl to take part in a thought experiment: that of the gradual lifting of intellectual obstacles enabling the scholar to go beyond the rigid dichotomies of a logical understanding to reach a harmonious integration of the two dimensions of language: ideality — incarnation in a single constitutive process.

3.4. Body/Language Ambivalence of Expressions Reconsidered

Whoever approaches language through logical ideality opposes linguistic expression and bodily expression, while an embodied conception of language allows for a founding continuity leading from one to another. The following passages might spot the starting point and end point of this development:

> Zu betonen ist, dass auch die so genannten unwillkürlichen „Ausdrücke" unserer Seelenlebens, wie Mienenspiel und Geste, zur ausgeschlossenen Sphäre gehören, obwohl die gewöhnliche Rede es bei ihnen wie bei den sprachlichen Ausdrücken zu sagen gestattet, dass ihre Bedeutung verstanden ist.[3] (*Bedeutungslehre* 1908: Husserl 1987: 10)

---

[2] 'Language appears as the body of thought to the speaker who uses language as a means of communicating with a living community.'

[3] 'It should be stressed that even the so-called involuntary "expressions" of our mental life, such as facial expression and gesture belong to the excluded sphere, although the ordinary speech permitted to say for them as for the linguistic expressions, that their meaning is understood.'

> Für Sehende, für Hörende, Sprechende sind die Worte „Ausdrücke", sind die Leiber Ausdrücke, die einen für Mitteilungen an andere Menschen, die anderen als Ausdrücke vom Dasein von Personen. Wortausdrücke setzt im Ausgedrückten Menschen als ausgeredete und nicht nur redende. Der erste und einfachste Ausdruck ist der des leiblichen Aussehens als Menschenleib, er setzt natürlich „Sehende" und verstehende voraus".[4]
> (*Phänomenologie der Intersubjektivität* III, Blg. LVI, 9 Sept. 1935: Husserl 1973a: 665)

Husserl's overcoming of a prior Cartesian solipsism which posed communication as inessential to thought promoted body expression to the status of linguistic expression and his subsequent foundation of subjective experience in intersubjectivity involved the founding of expression in communication. That said, we must beware of too quickly identify intersubjectivity with empathy and, with the latter, the ability for an observer to resonate with the behavior of an agent observed, an ability based on the brain system of mirror neurons. *Einfühlung* for Husserl — following Th. Lipps (1903) — is indeed a natural mode of perceiving the body of the other as a direct expression of inner life, but let's not lose sight, Ricœur reminds us in *Soi-même comme un autre* (1990), that it is also an ethical imperative one should practice in expressing one's own inner life, and that in this respect it is not obvious to naturalize. The thought process of a philosopher through the stages of his work takes a path and this path has some continuity — including a logical one — the interpreter would like to recover. But the task ahead us is perhaps not the unlikely reconciliation in the context of a unifying theory of the positions assumed by Husserl in *Logische Untersuchungen,* where his concern with linguistic expression is dominant, with those of *Erfahrung und Urteil,* where the issue of embodiment becomes central. The well documented multi-layered character of Husserl's thinking might inspire us a way out of such predicament: Why not deploy the issue of embodiment on as many tracks as needed to defuse looming incompatibilities and why do we not engage in parallel on all these tracks at once?

4. Three Ways to Bridge the Gap between Perception and Action, and Language

4.1. *Kinaesthesia in the Constitution of Lebenswelt*

How is it possible that the chain of physical events do not unfold in me without me, but that *I have* a sensible experience? How is it that visual forms *have for me* the value of independent things in the world or that movements of this body carry *my (or the other's) intentions*? How is it that expressions heard or produced do not simply obey the rules of phonology, syntax and semantics of some language (or reflect its statistic regularities), but are *endowed with sense* for

---

[4] 'For sighted people, for hearing people, for speaking people the words are "expressions" the bodies are expressions, the former for communicating with other people, the latter as expressions of the existence of persons. Verbal expressions suppose humans who express themselves as being spoken to, and not just speaking. The first and simplest expression is that of the physical appearance as a human body, it is naturally "seeing" and understanding in advance.'

communicating subjects? To answer these questions Husserl developed his theory of constitution. The principle is that any object of experience which appears to be provided with any value for a subject (including the value of being) must be understood *transcendentally*, i.e. based on subjective conditions of possibility qua deriving its value from the subject's own capability to make sense of it. The primary domain of this constitution includes the objects of perception and goals of practical actions of an individual subject, but it extends far beyond. The constitution of *Lebenswelt* is an intersubjective foundation of the ordinary life world of personal agents in a community, which perceive, act and communicate through speech drawing on their own resources and actively mobilizing their bodily capabilities of giving meaning. The experience of one's body and its extension through the intropathic experience of the other's body are operative in giving sense to objects through the progressive recruitment of the kinaesthetic systems of the body, from ocular to manual movements, and to locomotion. Neglect of Husserl's strong reliance on the kinaesthetic system in our dealing (*hantieren*) with anything whatsoever made of Merleau-Ponty's construing of man's being in the world as a kind of *hantieren* a slightly ghostly affair: «*Le corps hante le monde, etc.*».[5] The recruitment of the powers of the body in the constitution of the world and its objects cannot fail to interest the neurosciences, at least their embodied cognition trend: How do they take up the challenge? For my part, I do not see the above considerations as a chapter in the history of philosophy. Husserl himself conceived constitution not as the expression of idiosyncratic opinions of an individual thinker, but rather as a permanent program of collaborative research: Might cognitive neuroscience represent the updated form of implementation of such program, thereby achieving a successful naturalization of phenomenology (provisionally assuming, although it goes against the grain for the community of phenomenologists, that such naturalization is possible and desirable)?

4.2. *Tracing Neural Events in Brain Circuits of the Speaker: A Unique Pathway to Bridge the Gap?*

For the first time in history of the knowledge of man we see on the basis of data of empirical research a possibility to trace the uninterrupted course of events inside the organism that goes from perception and action to communication through language. Not content with tracking correlative activity patterns in the temporal-frontal circuits of an isolated brain during speech perception (cf. Pulvermüller *et al*. 2003; Pulvermüller 2005), neuroimagery reveals synchronizations of such patterns in the brains of communicating individuals (cf. Wilson 2007; Schippers 2009, 2010; Stephens 2010). However the narrative of brain events involved is far from answering all questions, despite the reductionist appeal of such chains of events for any naturalistic explanation of human linguistic behavior and capability. At that neuronal level 'the effort after meaning', in the words of Sir Frederick Bartlett (quoted by Barlow 1985), falls far short of giving a univocal ontological genesis. It remains a sequence of mere facts that keep the

---

[5] 'The body is haunting the world.'

contingency of what is empirical despite their derivation from the history of phylogenetic evolution and ontogenetic development. As cognitive science, the work on the neural basis of language is not just to go back to pure physical events: the occurrence of a change of brain state. First, the events that they seek as sciences of nature are the regulatory ideas of a consensus expected at the end of an ongoing controversy in the community of neuroscientists: Should we say ‘modulation’ or ‘induction’ or simply ‘spreading of activation’ or even ‘unselective cortical response to the task’ (cf. Dinstein *et al*. 2008; Hickok 2008; Mahon & Caramazza 2008)? These are not perfectly objective entities that are only what they are and whose unambiguous description might be immune to ‘the conflict of interpretations’, but rather the likely signature of a behavioral task or the language capability this conduct denotes. The fact that correlative brain events cannot be taken in isolation from the verbal behavior in ‘ecological’ conditions captured by the experimental protocol restores priority to phenomenology of language. Speaking of communication *between brains* as if the dialogue *between speaking persons* were a tale for the public remains a misnomer.

4.3. The Challenge of Part-Whole Semantic Dependency

The electrophysiological discharge of a nerve cell is an individual event so riveted to the present modality of its occurrence, that it contains no reserve of being to be further determined. Nothing in common with the entity of meaning — a semantic category expressed by any *syncategorematic* expression, one which realizes its function by its completion with other expressions of which it contains (not in explicit form but only in the signifying intention of the speaker) the empty place, a place quite determined, nevertheless, since it specifies a priori the category of suitable complementary expressions in the sentence:

> Synkategorematika werden als Träger inhaltlich bestimmter Bedeutungsmomente aufgefasst, die nach einer gewissen Ergänzung verlangen, und zwar einer Ergänzung, die, obschon der Materie nach unbestimmt, doch ihrer Form nach durch den gegebenen Inhalt mitbestimmt und somit gesetzlich umschrieben ist.[6]
>
> (*Logische Untersuchungen* IV § 5 in Husserl 1901/1913: 306).

The generality of this morphological structure of *incompleteness-dependence* is especially supported by the lastly revived structural syntax of Tesnière (1965), that some view as a possible alternative to Chomskyan generative grammar theory: *«Les connexions entre les mots ne sont indiquées par rien. Mais il est indispensable qu'elles soient aperçues par l'esprit sans quoi la phrase ne serait pas intelligible»* (*Eléments* A, I, 4)[7] (cf. Petitot, 1985, 1995, 2011, 36–37; Pulvermüller 2002: 139).

How is it possible that the brain frees itself from the transient and

---

[6] ‘Synkategorematika are to be construed as supports of meaning moments of content that require a certain supplement, and one supplement that, although underdetermined as regard the matter, is codetermined in form by the given content and is thus lawfully prescribed.’

[7] ‘Connections between words are specified by nothing. But it is essential that they are perceived by the mind without which the sentence is not intelligible.’

contingent flux of instantaneous events in its neural circuitry so as to be sensitive to dependencies which reflect the regulatory power of semantic categories, giving them life and support and impact over time? The question is no longer an absolute enigma posed by the phenomenology of language for empirical sciences, since the latter were split into a more observational neurophysiology (eventually supported by brain imagery) and a theoretical neuroscience working on hypothetical models along the lines of McCulloch & Pitts (1943), and Minsky & Papert (1969): One might consider that some naturalization of eidetics (not necessarily under that name) is in progress, in so far as the modeling (e.g., in Pulvermüller 2002: 124, 139, 214) of the detection of rule-governed relations of dependence between spatiotemporally non-adjacent elements of the verbal flow by neural networks implicitly naturalizes Husserl's theory of syntactic categories (despite its explicit anti-naturalism), as an a priori determination of the possible forms of meaning (*Logische Untersuchungen IV in* Husserl 1901/1913: 295).

4.4. Eidetic Phenomenology and its Geometric Modelling

Phenomenology claims to describe the verbal or perceptual semantic forms 'without prejudice', that is to say remaining at their own level of emergence which is that of lived experience, without making assumptions about the underlying causal mechanisms of basic brain substrate. His approach may perhaps be characterized as essentialist because it treats forms as transcendent entities objects of acts of consciousness. Anyway one cannot reproach it to be static and to freeze these forms in a Platonic heaven of Ideas, because these semantic forms of expression (words, phrases) are typically driven by *Ergänzungs-bedürftigkeit*: the need or requirement of completion which leads these forms to become parts of wholes (sentences, speech). Which brings back the semantic forms as expressed in discourse to the mereological standard of perception (*Logische Untersuchungen III in* Husserl 1901/1913 and also *Ding und Raum in* Husserl 1973b), suggesting a rule-governed transition from the muteness of perceptual (or practical) forms to the expressive forms of language. Along the same lines, a geometric morphodynamics (cf. Petitot 1985, 1992, 1999, 2011; Thom 1988) undertook in the last decades to model the morphogenetic dynamism with which semantic forms emerge from the physical substrate, by stabilizing them at the phenomenal plane and then transforming each into the other with a view to structuring the sense of experience of the speakers. More specifically, the process by which objects stand out in the visual field and the process by which prepositions apply to configurations of experience may both be represented or even reconstructed by equations of differential geometry. These equations were independently developed for the purposes of image analysis in computational vision, yet they can be interpreted, according to Petitot, as dynamic redeployment of Husserl's eidetic description of the constitution of visual objects through fusion (*Verschmelzung*) and separation (*Sonderung*) of sketches (*Abschattungen*) by gaze movements. Note that the high abstraction of such theoretical program does not prevent its promoter from claiming for it the label of Embodiment, which is consistent with his assumption of the phenomenological tradition: "The opening of the conceptual structure onto the phenomenal world is also an opening onto the body. Mind is 'embodied' and semiolinguistic structures and universals are fundamentally constrained by the compatibility between language, perception and action. Hence the spectacular renewal of *phenomenological* problematics (those of the later Husserl and Merleau-Ponty)" (Petitot 2011: 17–18).

4.5. *Lebenswelt: The limit of Body Foundationalism*

But will the constitution of the speaking world of communication save all the way the corporeal rootedness of the perceived world in kinesthesia? That is a bet made by a neuroscience of language that would aspire to naturalize our phenomenological experience of meaning. The gamble is rather risky. If only because the extension carried out by *Einfühlung* of the circle of actions and intentions of the ego to actions and intentions of others is definitely limited to the current face to face interaction. It cannot but stumble on ordinary *social acts*: (1) accomplished through speech, (2) dependent on the reception by the addressee, (3) separating in time the utterance and the realization, and (4) building supratemporal and immaterial relationships. Illustrating his eidetic analysis of the social acts with a familiar example, Reinach (1913) convincingly showed that the socially basic act of promising something to someone, and keeping one's word, owes nothing to empathy. In addition to kinaesthesia and *Einfühlung* or empathy the constitution of the verbally articulate *Lebenswelt,* including idealities of the Law regulating social acts, requires the recognition of speech and on top of speech the whole 'formalism' of language as web of reciprocally constitutive (but not purely bodily) co-operations. Such transcendentally ultimate constitutive power of language in relation to the world of institutional non physical realities we are dealing with in daily life no longer depends on kinaesthesia or *Einfühlung*: Hence the skepticism one may have towards the ambitions of a 'social neuroscience', such as planned by Gallese *et al*. (2004), looking for the basis of social cognition and eventually the roots of sociality in mirror neurons or in cortical maps of the brain.

5. Conclusion

Faced with issues unresolved by the mere narrative of events in a brain, that which is currently presented as a univocal ontological genesis of an embodied meaning will predictably break out in three directions: (i) *neurophysiologic* investigation of the organic substrate of the continuous linkage between perception and action, and language; (ii) *eidetic-geometric* morphodynamics as norm a priori backing the transformation of forms/schemes in syntactic or semantic structures; and (iii) *transcendental* constitution of the *Lebenswelt* of a community of perceiving-acting personal subjects who interact by words and gestures drawing on bodily capabilities and other operations of meaning-giving. Of these lines of approach only the first unquestionably ranks in the ideological framework of a naturalistic science, while the remaining lines cannot simply be fitted into traditional metaphysical dualism. Therefore we plead for an epistemology of language that is neither monistic nor dualistic, but rather trinitarian, as an alternative to the physicalism of current neuroscience (under its embodied disguise).

I might be objected that even a trinitarian approach should at least justify the mutual compatibility of the methods it proposes to be carried out in parallel. But the cause I am pleading here seems at first sight jeopardized by the irreconcilable character of the methods in question. "Monism, as one of my reviewers justly remarks, would presumably be the outcome of physical reductionism (of cognitive neuroscience), leading to the conclusion that a neurophysiologic account will eventually suffice." How did we come to introduce a triplication of approaches? There we came in playing the game of naturalizing the phenomenology of embodiment of language by means of cognitive neuroscience not by any ideological commitment in favor of physicalist naturalism, but to push to its limits this line of research in order to check the foreseeable incompleteness of its realization. The call for an eidetics of meaning on the one hand, for a transcendental constitution of the *Lebenswelt* on the other hand, is intended to meet the requirements arisen from the recognition of that deficiency. As a phenomenologist, I am not primarily concerned with the formal correctness of expression of a theory ideally cleaned up from any logical imperfections: capturing common intuitions is more important. About the embodiment of language, I do not defend a theory of my own: I am happy to accompany opportunistically certain research programs of which I have known in my narrow limits of scientific information. No excessive modesty in this, because I am convinced — with a few others — that the philosopher has no territory of its own to defend but is doomed to squat the territory of other disciplines. Such a situation does not allow one to ask these disciplines to put in coherence their respective approaches. The only mode of compatibility worthy to be envisaged is in the philosopher's *Erlebnis:* the lived experience of an unresolved tension between ultimately possibly incompatible approaches which nonetheless impose themselves as contingent context of the quest for truth.

References

Aziz-Zadeh, L., S.M. Wilson, G. Rizzolatti & M. Iacoboni. 2006. Congruent embodied representations for visually presented actions and linguistic phrases describing actions. *Current Biology* 16, 1818–1823.

Barlow, H.B. 1985. The twelfth Bartlett memorial lecture: The role of single neurons in the psychology of perception. *The Quarterly Journal of Experimental Psychology* 37A, 121–145.

Bookheimer, S. 2002. Functional MRI of language: New approaches to understanding the cortical organization of semantic processing. *Annual Review of Neuroscience* 25, 151–188.

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

Chomsky, N. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Corballis, M.C. 2002.*From Hand to Mouth: The Origins of Language*. Princeton, NJ:

Princeton University Press.

Corballis, M.C. 2004. The origins of modernity: Was autonomous speech the critical factor? *Psychological Review* 11, 543–552.

Corballis, M.C. 2011. *The Recursive Mind: The Origins of Human Language, Thought, and Civilization*. Princeton, NJ: Princeton University Press.

Corina, D.P. & H. Knapp. 2006. Sign language processing and the mirror neuron system. *Cortex* 42, 529–539.

Démonet, J.-F., F. Chollet, S. Ramsay, D. Cardebat, J.-L. Nespoulous, R. Wise, A. Rascol & R. Frackowiak. 1992. The anatomy of phonological and semantic processing in normal subjects. *Brain* 115, 1753–1768.

Dinstein, I., C. Thomas, M. Behrmann & D.J. Heeger. 2008. A mirror up to nature. *Current Biology* 18, R13–R18.

Emmorey, K., J. Xu, P. Gannon, S. Goldin-Meadow & A. Braun. 2010. CNS activation and regional connectivity during pantomime observation: No engagement of the mirror neuron system for deaf signers. *NeuroImage* 49, 994–1005.

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

Fadiga, L. & L. Craighero. 2006. Hand actions and speech representation in Broca's area. *Cortex* 42, 486–490.

Fadiga, L. & L. Craighero. 2007. Cues on the origin of language: From electrophysiological data on mirror neurons and motor representations. In S. Braten (ed.), *On Being Moved: From Mirror Neurons to Empathy*, 101–110. Amsterdam: John Benjamins.

Fadiga, L., L. Craighero & A. D'Ausilio. 2010. Broca's area in language, action, and music. In D. Riva & C. Nijiokiktjien (eds.), *Localization of Brain Lesions and Developmental Functions*. Eastleigh: John Libbey.

Fodor, J. 1983. *The Modularity of Mind: An Essay on Faculty Psychology.* Cambridge, MA: MIT Press.

Fodor, J. 1975. *The Language of Thought.* Cambridge, MA: Harvard University Press.

Frith, C.D. & U. Frith. 1999. Interacting minds — A biological basis. *Science* 286, 1692–1695.

Gallagher, H.L. & C.D. Frith. 2004. Dissociable neural pathways for the perception and recognition of expressive and instrumental gestures. *Neuropsychologia* 42, 1725–1736.

Gallese, V., C. Keysers & G. Rizzolatti. 2004. A unifying view of the basis of social cognition. *Trends in Cognitive Sciences* 8, 396–403.

Gallese, V. 2001. The shared manifold hypothesis. *Journal of Consciousness Studies* 8, 33–50.

Gallese, V. & A. Goldman. 1998. Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences* 2, 493–501.

Gentilucci, M., F. Benuzzi, M. Gangitano & S. Grimaldi.2001. Grasp with hand and mouth: A kinematic study on healthy subjects. *Journal of Neurophysiology* 86, 1685–1699.

Geschwind, N. 1965. The organization of language and the brain. *Science* 170, 940–944.

Gibson, J.J. 1977. The theory of affordances. In R. Shaw & J. Bransford (eds.), *Perceiving, Acting and Knowing*. Hillsdale, NJ: Erlbaum.

Gibson, J.J. 1979.*The Ecological Approach to Visual Perception.* Boston, MA: Houghton Mifflin.

Goldenberg, G. 2002. Goldstein and Gelb's case Schn. — A classic case in neurology? In C. Code, C. W. Wallech, Y. Joanette & A. Roch-Lecours (eds.), *Classic Cases in Neuropsychology II*, 281–289. Hove: Psychology Press.

Goldin-Meadow, S. 1999. The role of gesture in communication and thinking. *Trends in Cognitive Sciences* 3, 419–429.

Goldman, A. 1989. Interpretation psychologized. *Mind and Language* 4, 161–185.

Goldman, A. 1992. In defense of the simulation theory. *Mind and Language* 7, 104–119.

Goldstein, K. & A. Gelb. 1920. Über den Einfluss des vollständigen Verlustes des optischen Vorstellungsvermögens auf das taktile Erkennen. In A. Gelb & K. Goldstein (eds.), *Psychologische Analysen hirnpathologischer Fälle II*, 157–250. Leipzig: Johannes Ambrosius Barth Verlag.

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

Head, H. & G. Holmes. 1911. Sensory disturbances from cerebral lesions. *Brain* 34, 102–244.

Hickok, G. 2008. Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *Journal of Cognitive Neuroscience* 21, 1229–1243.

Holle, H., T.C. Gunter, S.-A.Rüschemeyer, A. Hennenlotter & M. Iacoboni. 2008. Neural correlates of the processing of co-speech gestures. *NeuroImage* 39, 2010–2024.

Husserl, E. 2008. Die Lebenswelt. Auslegungen der Vorgegebenen Welt und ihrer Konstitution, Texte aus dem Nachlass (1916–1937). In R. Sowa (ed.), *Husserliana XXXIX*. Dordrecht: Springer.

Husserl, E. 1987. *Vorlesungen über Bedeutungslehre, Sommersemester 1908.* In U. Panzer (ed.), *Husserliana XXVI*. Dordrecht: Martinus Nijhoff.

Husserl, E. 1973a. Zur Phänomenologie der Intersubjektivität. Texte aus dem Nachlass. In I. Kern (ed.), *I 1905–1920, II: 1921–1928, III: 1929–1935*, Den Haag: Martinus Nijhoff.

Husserl, E. 1973b. Ding und Raum — Vorlesungen 1907. In U. Claesges (ed.), *Husserliana XVI*. Den Haag: Martinus Nijhoff.

Husserl, E. 1901/1913. *Logische Untersuchungen, II/1 Untersuchungen zur Phänomenologie und Theorie der Erkenntnis.* Tübingen: Max Niemeyer Verlag.

Liberman, A.M. & I.G. Mattingly. 1985. The motor theory of speech perception revised. *Cognition* 21, 1–36.

Lichtheim, L. 1885. On aphasia. *Brain* 7, 433–484.

Lingnau, A., B. Gesierich & A. Caramazza. 2009. Asymmetric fMRI adaptation reveals no evidence for mirror neurons in humans. *Proceedings of the National Academy of Science* 106, 9925–9930.

Lipps, T. 1903. Einfühlung, innere Nachahmung, und Organempfindung. *Archiv für die gesamte Psychologie* I, 185–204.

Lotto, A.J., G.S. Hickok & L.L. Holt. 2008. Reflections on mirror neurons and speech perception. *Trends in Cognitive Sciences* 13, 110–114.

MacSweeney, M., C.M. Capek, R. Campbell & B. Woll. 2008. The signing brain: The neurobiology of sign language. *Trends in Cognitive Sciences* 12, 432–440.

Mahon, B.Z. & A. Caramazza. 2008. A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology – Paris* 102, 59–70.

McCulloch, W.S. & W.H. Pitts. 1943. A logical calculus of ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics* 5, 115–133.

Merleau-Ponty, M. 1945. *Phénoménologie de la Perception.* Paris: Gallimard.

Merleau-Ponty, M. 1960. *Signes.* Paris: Gallimard.

Minsky, M. & S. Papert. 1969. *Perceptrons.* Cambridge, MA: Cambridge University Press.

Montgomery, K.J., N. Isenberg & J.V. Haxby. 2007. Communicative hand gestures and object-directed hand movements activated the mirror neuron system. *Scan* 2, 114–122.

Nishitani, N., M. Schürmann, K. Amunts & R. Hari. 2005. Broca's region: From action to language. *Physiology* 20, 60–69.

Norman, D. A. & T. Shallice. 1980. *Attention to Action: Willed and Automatic Control of Behavior (Report No. 8006).* San Diego, CA: University of California.

Penfield, W. & T. B. Rasmussen. 1950. *The Cerebral Cortex of Man. A Clinical Study of Localization of Function.* New York: MacMillan.

Petit, J.-L. 2010. Corps propre, schéma corporel et cartes somatotopiques. In A. Berthoz & B. Andrieu (eds.), *Merleau-Ponty: Le Corps en Acte*, 41–58. Nancy: Presses Universitaires de Nancy.

Petitot, J. 2011. *Cognitive Morphodynamics. Dynamical Morphological Models of Constituency in Perception and Syntax*. Bern: Peter Lang.

Petitot, J. 1999. Morphological eidetics for phenomenology of perception. In J. Petitot, F.J. Varela, J.-M. Roy & B. Pachoud (eds.), *Naturalizing Phenomenology: Issues in Contemporary Phenomenology and Cognitive Science*, 330–371. Stanford, CA: Stanford University Press.

Petitot, J. 1995. Approche morphodynamique de l'iconicité des stemmas. In F. Madray-Lesigne & J. Richard-Zapella (eds.), *Lucien Tesnière aujourd'hui,* 105–112. Louvain: Peeters.

Petitot, J. 1985.*Morphogenèse du Sens.* Paris: Presses Universitaires de France.

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

Pulvermüller, F. 2002. *The Neuroscience of Language. On Brain Circuits of Words and Serial Order*. Cambridge: Cambridge University Press.

Pulvermüller, F., Y. Shtyrov & R. J. Ilmoniemi. 2003. Spatiotemporal patterns of neural language processing: An MEG study using minimum-norm current estimates. *NeuroImage* 20, 1020–1025.

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

Pulvermüller, F. & L. Fadiga. 2010. Active perception: Sensorimotor circuits as a cortical basis for language. *Nature Reviews Neuroscience* 11, 351–360.

Reinach, A. 1913. Die apriorischen Grundlagen des bürgerlichen Rechtes. *Jahrbuch für Philosophie und phänomenologische Forschung* 1, 685–847.

Ricoeur, P. 1990. *Soi-Même comme un Autre.* Paris: Seuil.

Rizzolatti, G. & M.A. Arbib. 1998. Language within our grasp. *Trends in Neuroscience* 21, 188–194.

Rizzolatti, G., L. Riggio & B.M. Sheliga. 1994. Space and selective attention. In C. Umilta & M. Moscovitch (eds.), *Attention and Performance XV*, 232–265.

Hillsdale, NJ: Erlsbaum.

Schippers, M.B., V. Gazzola, R. Goebel & C. Keysers. 2009. Playing charades in the fMRI: Are mirror and/or mentalizing areas involved in gestural communication? *PLoS ONE* 4, e6801.

Schippers, M.B., A. Roebroeck, R. Renken, L. Nanetti & C. Keysers. 2010. Mapping the information flow from one brain to another during gestural communication. *Proceedings of the National Academy of Science* 107, 9388–9393.

Scott, S.K., C. McGettigan & F. Eisner. 2009. A little more conversation, a little less action — candidate roles for the motor cortex in speech perception. *Nature Reviews Neuroscience* 10, 295–302.

Skipper, J.I., V. van Wassenhove, H.C. Nusbaum & S.I. Small. 2007. Hearing lips and seeing voices: How cortical areas supporting speech producing mediate audiovisual speech perception. *Cerebral Cortex* 17, 2387–2399.

Stephens, G.J., L.J. Silbert & U. Hasson. 2010. Speaker-listener neural coupling underlies successful communication. *Proceedings of the National Academy of Science* 107, 14425–14430.

Tarski, A. 1936. Der Wahrheitsbegriff in den formalisierten Sprachen. *Studia Philosophica* 1, 261–405.

Tesnière, L. 1965. *Eléments de Syntaxe Structurale.* Paris: Klincksieck.

Tettamanti, M. & D. Weniger. 2006. Broca's area: A supramodal hierarchical processor? *Cortex* 42, 491–494.

Tettamanti, M., I. Rotondi, D. Perani, G. Scotti, F. Fazio, S. F. Cappa & A. Moro. 2009. Syntax without language: Neurobiological evidence for cross-domain syntactic computations.*Cortex* 45, 825–838.

Umilta, C., C. Mucignat, L. Riggio, C. Barbieri & G. Rizzolatti. 1994. Programming shifts of visual attention. *European Journal of Cognitive Psychology* 6, 23–41.

von der Malsburg, C. 1995. Binding in models of perception and brain function. *Current Opinions in Neurobiology* 5, 520–526.

Wise, R., F. Chollet, U. Hadar, K. Friston, E. Hoffner & R. Frackowiak. 1991. Distribution of cortical neural networks involved in word comprehension and word retrieval. *Brain* 114, 1803–1817.

Xu, J., P.J. Gannon, K. Emmorey, J.F. Smith & A.R. Braun. 2009. Symbolic gestures and spoken language are processed by a common neural system. *Proceedings of the National Academy of Science* 106, 20664–20669.

Jean-Luc Petit
Université de Strasbourg &
Faculté de Philosophie
7, rue de l'Université
67000 Strasbourg
France

Collège de France
Laboratoire de Physiologie de la Perception et de l'Action
11, place Marcelin Berthelot
75005 Paris
France

jean-luc.petit@college-de-france.fr

The Link between Action and Language: Recent Findings and Future Perspectives

Claudia Repetto, Barbara Colombo & Giuseppe Riva

This paper aims to present a critical review of studies focused on embodied cognition and, more specifically, on the relationship between language and action. A critical analysis of studies using methods such as TMS and fMRI will be presented, and results reported by the different studies will be discussed, both theoretically and methodologically. Then, in response to some inconsistency detected by the analysis of literature, Virtual Reality will be presented as a possible answer or enrichment for the study of this topic. Possible future research tracks and application are discussed.

Keywords: embodied language; functional magnetic resonance imaging (fMRI); transcranial magnetic stimulation (TMS); virtual reality (VR)

1. Introduction

Traditional theories of cognition are based on the idea that knowledge is represented in the brain in the form of concepts and stored in memory systems as semantic information. Concepts, from this perspective, are conceived as amodal, abstract and arbitrary (Fodor 1975), then independent from the brain's modal system of perception (e.g., vision, audition), and action (e.g., movement, proprioception). Chomsky's theory of language (Chomsky 1965) is completely aligned with this view: The theory of Universal Grammar considers language as a corpus of abstract symbols combined together according to formal syntactic rules; two properties, among others, are distinctive of human language, the generativity and compositionality.

In more recent years, nevertheless, a radically different conception of knowledge has been taken into account, that brings together data from different methodological approaches such as neurobiology, brain imaging, and neuropsychology: the theory of Embodied Cognition (Wilson 2002; Gibbs 2006). According to the embodied cognition hypothesis, concepts are not amodal and knowledge relies on body states and experiences. Therefore, there is a tight link between concepts, action, and perception, to the extent that conceptual knowledge is mapped within the sensory-motor system. The notion that cognition is grounded in action and perception is encapsulated in the term 'embodiment'.

ISSN 1450–341

Metaphorically speaking, embodied theories of cognition extended the boundaries of anatomical structures to which traditionally a specific function was assigned: The mind is no longer confined to the brain but also includes other body parts, such as hands, legs, eyes. Moreover, within the brain, the separation between primary areas, recruited for basic sensory and motor processing, and the associative areas, in which more complex processes take place is not strictly defined anymore: actually, the distinction between low and high level processes drops down in favor of a more integrated model. This new model proposes an interplay that allows the recruitment of primary areas even during cognitive processes such as language and conceptualization.

In the last decades many neuroscientists focused on the theory of embodied cognition in general, and on embodied language in particular. Embodied theories of language predict that the neural structures involved in sensory, perceptual or motor areas are also active when processing words whose meaning embeds prominent sensory (auditory and tactile features; Goldberg *et al*. 2006), perceptual (color; Martin *et al*. 1995), or faces and places; Aziz-Zadeh *et al*. 2008), or motor features (see below for a detailed review). This interest in embodied theories of cognitive processes, thus, yielded a growing corpus of data, yet still many topics are unclear and deserve further investigation. One of the most intriguing one is the link between language processing and motor system, which has been extensively investigated in recent years.

Starting from these considerations, the present paper has two main purposes: The first one is to briefly review the recent literature that addresses the relationship between motor system and language processing, distinguishing research on the base of the tool used to investigate this issue (transcranial magnetic stimulation or functional magnetic resonance). The intention is to show how and to what extent experimental protocols with different methodologies and tools lead sometimes to contrasting results; moreover a special attention will be paid to the discussion of the capabilities that each technique inherently presents. The second goal is to reflect on future perspectives. In particular, we will present a new tool for the study of the embodiment that, to our knowledge, has not been used so far for studying cognitive processes: virtual reality.

2. TMS Studies

Transcranial magnetic stimulation (TMS) proved to be an efficient and promising method to investigate the link between action and language. Thanks to its temporal and spatial resolution, TMS became one of the most used tools to study where and when the language processes are mapped within the motor system.

Most of the researchers applied single pulse TMS protocols over the primary motor cortex (M1) during a linguistic task and registered motor evoked potential (MEP) from the muscles that are supposed to respond depending on the portion of the cortex stimulated. The rational is the following: If the linguistic task engages to some extent the portion of the cortex stimulated at the time of stimulation, then it should result in a modulation of cortico-spinal excitability and thus of the MEP amplitude (compared to rest condition).

This kind of experimental design has been mostly employed to investigate the role of M1 during the processing of abstract vs. action verbs, but results are sometimes contrasting. For example, Papeo *et al.* (2009) reported an increase of MEPs recorded while participants read action verbs compared with what happened while they read verbs describing abstract concepts; in contrast, Buccino *et al.* (2005) described a reverse situation during language comprehension: MEPs recorded from hand muscles was lower while participants heard hand-related action verbs compared to foot-related action verbs, indicating an effector specific inhibition. Although these findings might seem incoherent, several different experimental features can account for them; one of these is the timing of stimulation, which is an important issue to consider when studying excitability of such dynamic systems. In fact, we can argue that stimulation of an area occurring just while the process is taking place should produce an interference effect, and hence an inhibition of that area; by opposite, a stimulation delivered shortly before the onset of the process in this given area might act as a prime and produce a sort of facilitation effect (preactivation) for that area. Papeo *et al.* (2009) evaluated the effects of TMS over M1 at different windows of time from the linguistic stimulus onset: They reported an involvement of M1 in the linguistic process only when stimulation was delivered after 500 ms post-stimulus, that is in the post-conceptual stage but not in the previous ones. This result would lead us to think that lexical-semantic processing of action verbs does not automatically activate the M1, whose activation is modulated in a top-down manner.

The second element to take into account is the specific linguistic task performed by participants. In literature we can find different researches that employed different linguistic tasks to evaluate motor activation, each of whom entailed different linguistic processes. In some cases lexical decision was required (Pulvermuller *et al.* 2005), while others used reading (Fadiga *et al.* 2002), semantic judgments (Buccino *et al.* 2005), imagery (Fourkas *et al.* 2006), transformation tasks (Oliveri *et al.* 2004). Tomasino *et al.* (2008) compared systematically the effects of different timings of stimulation during different kind of tasks (silent reading, motor imagery and frequency judgments) and found that M1 plays a role only during motor imagery, so they concluded that the recruitment of motor networks during language understanding is not required, but it occurs only when explicit motor simulation is requested. However, the effect of TMS in modulating MEPs during semantic judgments of nouns (natural vs. tools; graspable vs. ungraspable) has been reported, even without any overt motor simulation (Gough *et al.* 2012). The identification vs. distinction of the simulation/imagery processes is still open, even if imaging data seem to support the distinction hypeothesis (Willems *et al.* 2010b, see below).

Recently TMS protocols have been employed to discover the role of morpho-syntactic features on the activity of M1: Papeo and colleagues (Papeo *et al.* 2011) compared MEPs recorded during reading tasks of action vs. abstract verbs presented using the first or the third singular person (*I* vs *he/she*); they found an increase of MEPs amplitude selectively for the action verbs at the first person, deriving from these data that motor simulation is facilitated when the conceptual representation of the verb includes the self as agent. Furthermore, a sensitivity of the primary motor cortex to the polarity of sentences was highlighted: Active action-related sentences suppressed cortico-spinal reactivity compared to passive action-related sentences, and either active or passive abstract sentences (Liuzza *et al*. 2011).

Finally, TMS can be used in offline procedures, delivering repeated trains of stimulation over a period of time lasting several minutes (rTMS or TBS) in order to modify transiently the cortical excitability and investigate the role of the stimulated area in a given process. In this case experimenters are not interested in defining the exact timing of the cognitive process but rather aim to discover if the area is involved in that process. To this field of application can be ascribed the studies carried out by Gerfo *et al*. (2008) and Willems *et al*. (2011). In both studies, motor networks (primary and/or premotor cortices) are found to be functionally relevant in action-related language understanding.

Future studies are needed to investigate with offline (facilitatory and inhibitory) stimulation the role of motor areas in different linguistic tasks in order to deepen the knowledge about their function (causal or epiphenomenal?) during language processing.

3. Imaging Studies

Functional magnetic resonance imaging (fMRI) is so far the imaging technique preferred by researchers who intend to shed light on the relationship between motor areas and language processing. While TMS studies allow to establish a causal link between experimental manipulations (i.e. site of stimulation) and behavioural tasks (i.e. linguistic tasks), fMRi experiments are correlational protocols by nature, giving the possibility to identify, among all the brain areas, those engaged during a specific process and a precise window time; further, fMRI allows to track down networks of activations, reflecting the dynamic features of the process under investigation.

A first line of research aimed to determine if and where language processing recruits brain areas usually activated during motor tasks (considered in a broad sense, i.e. motor observation, preparation, execution). This topic often intercepts and includes theoretical issues that arise from studies focused on mirror neurons. In fact, it is well known that mirror neurons in monkeys are activated not only by the observation of a movement performed by others but also when the noise associated to the action is heard (Kohler *et al*. 2002). In humans, action-related auditory inputs are well implemented in language stimuli: This happens in particular when sentences describing actions are presented auditorily. Many studies have been carried out to explore the possibility that the understanding of action-related sentence relies on the same observation-execution system by means of mirror neurons (see Aziz-Zadeh & Damasio 2008 for a review). Most of these researches, relying on different linguistics tasks, reported a somatotopic activation of premotor cortex, primary motor cortex and Broca's region (Hauk *et al*. 2004; Tettamanti *et al*. 2005; Aziz-Zadeh *et al*. 2006). Interestingly, this pattern of activation is confirmed even in children (age 4–6), as described by James & Maouene (2009), indicating that the embodied nature of language makes its appearance early in child development, when the language is not wholly acquired. Nevertheless, it is noteworthy that there is not a strong consensus about a somatotopic organisation of action words meaning representations, and this fact is not astounding considering that the organization of the premotor cortex is still poorly understood. For example, Postle *et al*. (2008), combining functional MRI with cyto-architectonically defined probabilistic maps of left hemisphere primary and premotor cortices, failed to find a direct correspondence between the activations triggered by effector-specific action words meaning and those found during the real movement of the same effectors.

As it has been noticed reviewing TMS studies, even in this case the kind of task and the features of the verbal material seem to yield different results. Raposo *et al*. (2009) comparing cerebral activation when proposing different semantic contexts (isolated action-verbs, literal sentences, idiomatic sentences) found that neural response was maximum in motor areas for isolated verbs and minimum for idiomatic sentences, with literal sentences in the middle; according to authors discussion, these findings suggest that motor response during language processing is context-dependent rather than automatic and invariable. From a similar perspective, van Dam and collaborators (van Dam *et al*. 2010) examined brain activity during the semantic judgment of verbs describing actions with different degrees of kinematic details: a region within the bilateral inferior parietal lobule proved to be sensitive to the specificity of motor programs associated to the action verbs, with the BOLD signal greater for the finest-grained actions.

Finally, fMRI can contribute to refine the theory of embodied language and also to test hypotheses that, if confirmed, can add data in favor of this theoretical position. In one recent research Willems *et al*. (2010b) investigated the construct of mental simulation, which is thought to be one of the core mechanism of embodiment, but it is still unclear whether it is the equivalent to explicit imagery. In particular, the authors found that implicit simulation of actions during language understanding is neurally dissociated from explicit motor imagery, thus confirming that the two processes are distinct in nature. Furthermore, according to simulation hypothesis, as stated by Willems *et al*. (2010a), "if understanding action words involves mentally simulating one's own actions, then the neurocognitive representation of word meanings should differ for people with different kinds of bodies, who perform actions in systematically different ways" (i.e. right- vs. left-handers): This prediction has been corroborated by fMRI data which showed a preferential activation of the right premotor cortex during lexical decision on action verbs for left-handers, and the opposite pattern of activation for the right-handers.

As showed in this short excursus, fMRI studies gave an important contribute to the study of the link between language processes and perceptive brain areas, thus adding essential pixels to the big picture of embodied semantics theory; however, beside traditional neuroscience techniques, such as fMRI and TMS, other tools could demonstrate great capabilities in this field of application: The next section is dedicated to the description of one of them, virtual reality.

4. Virtual Reality: A New Frontier for Neuroscience Research

A virtual reality system (VR) is a combination of technological devices that allows users creating, exploring and interacting with 3D environments. Typically, people entering a virtual environment feels like being a part of this world and has the opportunity to interact with it almost like he would do in real world: Just turning around his head, a user can explore visually the scene, and with other user-friendly controls he/she can move through the environment, approach objects, select them, meet other people presented as avatars or video-tape. This capability is made possible by the use of input tools (trackers, gloves, mice) that send to the computer the position and the movement of the user in real time, graphic rendering that changes the environment coherently with the information acquired, and output devices (visual, aural, and haptic) that return to the user a feedback of the interaction.

However, it is the user immersion in a synthetic environment that characterizes VR as being different from interactive computer graphics or multimedia. In fact, the sense of *presence* in a virtual world elicited by immersive VR technology indicates that VR applications may differ fundamentally from those commonly associated with graphics and multimedia systems. Even if there is not yet a common agreement about what Presence is common definitions are the "sense of being there" (Steuer 1992) "the feeling of being in a world that exists outside the self" (Waterworth *et al.* 2010; Riva et al. 2011) or the "perceptual illusion of non-mediation" (Lombard & Ditton 1997). In general, scientific literature identified a set of factors that have a direct influence on the experience of presence (IJsselsteijn & Riva 2003; Riva 2006; Youngblut 2007): (a) the processing of multimodal input (visual, tactile, auditory, kinesthetic, olfactory) from the virtual experience be combined to form coherent perceptual categories — that is that the virtual experience be recognized as 'real'; (b) the processing of the multimodal input in an egocentric reference frame — that is the user feels that he or she is within the environment as opposed to observing it from a third person perspective; and (c) the ability to give a meaning to the multimodal input — that is that the virtual experience be recognized as 'meaningful' and 'relevant'.

Far from being a merely recreational tool, VR is increasingly used in research and clinical settings (Riva 2002). Traditionally, the most common application of VR in mental health is related to the treatment of anxiety disorders (Emmelkamp 2005; Parsons & Rizzo 2008): from simple phobias (Rothbaum *et al.* 2006; Krijn *et al.* 2007), to panic disorders (Vincelli *et al.* 2003; Botella *et al.* 2007), post-traumatic stress disorder (Rothbaum *et al.* 2001; Gerardi *et al.* 2008), and generalized anxiety disorder (Repetto *et al.* 2009a, 2009b; Repetto & Riva 2011). The reason for the diffusion of the VR in this field of application is its versatility for implementing exposure therapy (VRET): In fact, VRET is safer, more controllable, less embarrassing and costly than in vivo exposure, but at the same time its immersive nature provides a real-like experience that may be more emotionally engaging than imaginal exposure (Riva 2010).

Recently Bohil and colleagues (Bohil *et al.* 2011) described the advantages of using virtual environments in several domains of neuroscience, such as spatial navigation, multisensory integration, social neuroscience, pain remediation, and neuro-rehabilitation. The authors pointed out the capabilities of VR for implementing experiments that overcome traditional limitations encountered by researchers interested in understanding the functioning of central nervous system.

One of these limitations is the gap between the degree of complexity typical of the real world and that embedded into the stimuli created ad hoc for the experimental protocol. In fact, usually participants in research settings perform tasks interacting with several different devices (i.e. computer, botton boxes) none of which is designed to simulate the real experience where the process investigated occurs. Virtual reality, by opposite, allows bypassing the common criticism toward the experimental setting, that is, its poor ecological validity: Immersing participant in virtual environments one could gain ecological validity without giving up controllability and replicability.

For researchers interested in studying cognitive processes from an embodied point of view this is a great opportunity: If representations in the cognitive system are multimodal, then to investigate their properties one should recreate the multimodal experience that can trigger the process. Furthermore, with the advance of technology, the interface between subject and VR system is more and more intended to become a non-mediated process, in which the body itself is the navigation tool (without the need of control devices). For these reasons, VR could be thought as an ideal medium for investigating several cognitive domains (Riva 1998), but the capabilities are not confined to the fact that inside the virtual experience many different source of stimulation can work together to recreate a realistic environment. In fact, VR can be considered an 'embodied technology' for its effects on body perceptions (Riva 2002): It is possible the use of VR for inducing controlled changes to the experience of the body. On one side, VR has been used to improve the experience of the body in patients with eating disorders (Perpiña *et al.* 1999; Riva *et al.* 2003; Ferrer-García & Gutiérrez-Maldonado 2012) or obesity (Riva *et al.* 2006). On the other side, different authors used VR to induce illusory perceptions — e.g., a fake limb (Slater *et al.* 2009) or body transfer illusion (Slater *et al.* 2010) — by altering the normal association between touch and its visual correlate. Being an embodied technology, VR seems a promising tool for the investigation of the link between language and action. In the recent past, the discovery of mirror neurons changed the outlook of neuroscience and established a connection between language and motor system (Gallese & Lakoff 2005; Chen & Yuan 2008).

The embodiment theory of language assigns an important role to this class of motor neurons in understanding action related concepts: mirror neurons should be activated by the linguistic stimulus and hence it should result in a modulation of the primary and premotor cortex (Gallese 2008). As reviewed in previous sections, several studies confirmed that language itself triggers motor-like responses within the cerebral areas where movement is represented (Hauk *et al.* 2004; Buccino *et al.* 2005). The opposite way to understand the relationships between language and action is to investigate if and to what extent motor inputs affect language representation and acquisition. Paulus and colleagues (Paulus *et al.* 2009) asked participants to learn functional verbal knowledge of new objects while performing different motor tasks. They found the presence of motor interference when the acquisition of manual object knowledge was paired with the concurrent manual action but this wasn't true if concurrent actions with the feet were performed. Furthermore, Macedonia and colleagues (Macedonia *et al.* 2011) studied the impact of iconic gestures on foreign language words learning: If learning of novel words was coupled to iconic gestures participants retained better the verbal material over time, if compared with meaningless gestures; this behavioral data was accompanied to imaging data, that indicated an activation of premotor cortices only for words encoded with iconic gestures.

The researches that use actions for understanding the interplay between language, motor system, and mirror neurons find in VR a privileged medium where being implemented. VR gives users the opportunity to see themselves moving in the environment while being comfortably seated in a chair. Thanks to different input devices participants could virtually perform any action, even those typically not performable in an experimental setting (to jump a rope, kick a ball, or shoot something, for example). Thus, within a virtual environment, experimenters could investigate the effect on language processing of performing different actions. The fact that users are not really moving their bodies in the real space, but still have the subjective sensation of being ‘in action’, places VR in a intermediate position between the real action and mere action observation (such as in a video): It has been demonstrated that cortical excitability is modified by the observation of movements performed by others (Strafella & Paus 2000), but this modulation is greater if the orientation of the movement is compatible with the point of view of the observer (Maeda *et al*. 2002). The advantage of VR is the fact that the movement the individual does is egocentric, exactly as if he/she would act in real world.

As Cameirao has argued (Cameirao *et al*. 2010), the first-person perspective could engage stronger the mirror neurons system because this is the perspective the system is most frequently exposed to. This observation has important rebounds in the field of rehabilitation: If the enactment of verbal material facilitates learning in non pathological samples, it should be investigated if this effect is replicable in people with language deficit. Moreover, often patients with different types of aphasia have motor deficits as well, and VR could give them the opportunity to take advantage of the action-language coupling protocols even without moving at all.

Finally, VR experiments can be conducted also in association with imaging techniques, such as fMRI: Further researches, thus, using virtual environments during fMRI scans could shed light on the cortical activations triggered by virtual movements, and on the role of mirror neurons in these processes.

5. Conclusions

This contribution, starting from a theoretical reflection on the importance of the embodied cognition, aims to emphasize the relevance of this topic for the study of the relationship between language and the motor cortex.

Recently, many studies have been presented related to this topic, but a review of those studies has revealed conflicting results. Which could be the cause of these differences? A critical analysis has allowed us to hypothesize that they may be at least partly attributable to different experimental protocols, each of whom would study a specific stage of the neurocognitive processes being examined, leading them to measure different things, and reporting different results when relating the recorded values to the same functions without stressing the differences in timing along the whole examined process.

After this first clarification, it is encouraging to be able to highlight, on the basis of the review presented here, how the investigation techniques used in the presented studies, which are extremely different from each other, and aimed at investigating different aspects, setting the research protocol with a causal perspective (TMS) or a correlational one (fMRI), still revealed a strong reliable theoretical link between language and action. But how is it possible to operationalize these results, taking into account, with a critical perspective, also the differences to which different research protocols lead? In order to answer this question in this paper we suggest and discuss the theoretical and operational usefulness and relevance of VR.

Due to its functional characteristics, which are extensively described, this tool allows to test many of the theories previously investigated with other techniques, but using a more environmentally friendly (and ecological) setting and a reverse pattern (starting from real action and not from an abstract/verbal stimulus) that would allow a real enrichment of this specific area.

In addition to this, the already known and popular applications of VR in clinical settings, open up new fields of application of studies linking language and action (with particular attention given to the contribution given by studies on mirror neurons). Actually, this method not only allows an enrichment of specific knowledge on the phenomenon, but it can be considered as a promising field for applications of theoretical insights to improve the learning or relearning of language or motor skills in deficit conditions.

References

Aziz-Zadeh, L., S. M. Wilson, G. Rizzolatti & M. Iacoboni. 2006. Congruent embodied representations for visually presented actions and linguistic phrases describing actions. *Current Biology* 16, 1818–1823.

Aziz-Zadeh, L. & A. Damasio. 2008. Embodied semantics for actions: Findings from functional brain imaging. *Journal of Physiology - Paris* 102, 35–39.

Aziz-Zadeh, L., C. J. Fiebach, S. Naranayan, J. Feldman, E. Dodge & R. B. Ivry. 2008. Modulation of the FFA and PPA by language related to faces and places. *Society for Neuroscience* 3, 229–238.

Bohil, C. J., B. Alicea & F. A. Biocca. 2011. Virtual reality in neuroscience research and therapy. *Nature Reviews Neuroscience* 12, 752–762.

Botella, C., A. García-Palacios, H. Villa, R. M. Baños, S. Quero, M. Alcañiz & G. Riva. 2007. Virtual reality exposure in the treatment of panic disorder and agoraphobia: A controlled study. *Clinical Psychology & Psychotherapy* 14, 164–175.

Buccino, G., L. Riggio, G. Melli, F. Binkofski, V. Gallese & G. Rizzolatti. 2005. Listening to action-related sentences modulates the activity of the motor system: A combined TMS and behavioral study. *Brain Research Cognitive Brain*

Research 24, 355–363.

Cameirao, M. S., S. B. Badia, E. D. Oller & P. F. Verschure. 2010. Neurorehabilitation using the virtual reality based Rehabilitation Gaming System: Methodology, design, psychometrics, usability and validation. *Journal of Neuro Engineering and Rehabilitation* 7, 48.

Chen, W. & T. F. Yuan. 2008. Mirror neuron system as the joint from action to language. *Neuroscience Bulletin* 24, 259–264.

Chomsky, N. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Emmelkamp, P. M. 2005. Technological innovations in clinical assessment and psychotherapy. *Psychotherapy and Psychosomatics* 74, 336–343.

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

Ferrer-García, M. & J. Gutiérrez-Maldonado. 2012. The use of virtual reality in the study, assessment, and treatment of body image in eating disorders and nonclinical samples: A review of the literature. *Body Image* 9, 1–11.

Fodor, J. 1975. *The Language of Thought*. Cambridge, MA: Harvard University Press.

Fourkas, A. D., A. Avenanti, C. Urgesi & S. M. Aglioti. 2006. Corticospinal facilitation during first and third person imagery. *Experimental Brain Research* 168, 143–151.

Gallese, V. & G. Lakoff. 2005. The brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology* 22, 455–479.

Gallese, V. 2008. Mirror neurons and the social nature of language: The neural exploitation hypothesis. *Society for Neuroscience* 3, 317–333.

Gerardi, M., B. O. Rothbaum, K. Ressler, M. Heekin & A. Rizzo. 2008. Virtual reality exposure therapy using a virtual Iraq: Case report. *Journal of Traumatic Stress* 21, 209–213.

Gerfo, E. L., M. Oliveri, S. Torriero, S. Salerno, G. Koch & C. Caltagirone. 2008. The influence of rTMS over prefrontal and motor areas in a morphological task: grammatical vs. semantic effects. *Neuropsychologia* 46, 764–770.

Gibbs, R. W. 2006. *Embodiement and Cognitive Science*. New York: Cambridge University Press.

Goldberg, R. F., C. A. Perfetti & W. Schneider. 2006. Perceptual knowledge retrieval activates sensory brain regions. *Journal of Neuroscience* 26, 4917–4921.

Gough, P. M., L. Riggio, F. Chersi, M. Sato, L. Fogassi & G. Buccino. 2012. Nouns referring to tools and natural objects differentially modulate the motor system. *Neuropsychologia* 50, 19–25.

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

IJsselsteijn, W. A. & G. Riva. 2003. Being there: The experience of presence in mediated environments. In F. Davide, G. Riva & W. A. IJsselsteijn (eds.), *Being There: Concepts, Effects and Measurements of User Presence in Synthetic Environments*. Amsterdam: IOS Press.

James, K. H. & J. Maouene. 2009. Auditory verb perception recruits motor systems in the developing brain: An fMRI investigation. *Developmental Science* 12, F26–34.

Kohler, E., C. Keysers, M. A. Umilta, L. Fogassi, V. Gallese & G. Rizzolatti. 2002. Hearing sounds, understanding actions: Action representation in mirror neurons. *Science* 297, 846–848.

Krijn, M., P. M. Emmelkamp, R. P. Olafsson, M. Bouwman, L. J. van Gerwen, P. Spinhoven, M. J. Schuemie & C. A. van der Mast. 2007. Fear of flying treatment methods: virtual reality exposure vs. cognitive behavioral therapy. *Aviation*, *Space*, and *Environmental Medicine* 78, 121–128.

Liuzza, M. T., M. Candidi & S. M. Aglioti. 2011. Do not resonate with actions: sentence polarity modulates cortico-spinal excitability during action-related sentence reading. *PLoS One* 6, e16855.

Lombard, Matthew & Theresa Ditton. 1997. At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication* 3, 0.

Macedonia, M., K. Muller & A. D. Friederici. 2011. The impact of iconic gestures on foreign language word learning and its neural substrate. *Human Brain Mapping* 32, 982–998.

Maeda, F., G. Kleiner-Fisman & A. Pascual-Leone. 2002. Motor facilitation while observing hand actions: specificity of the effect and role of observer's orientation. *Journal of Neurophysiology* 87, 1329–1335.

Martin, A., J. V. Haxby, F. M. Lalonde, C. L., Wiggs & L. G. Ungerleider. 1995. Discrete cortical regions associated with knowledge of color and knowledge of action. *Science* 270, 102–105.

Oliveri, M., C. Finocchiaro, K. Shapiro, M. Gangitano, A. Caramazza & A. Pascual-Leone. 2004. All talk and no action: A transcranial magnetic stimulation study of motor cortex activation during action word production. *Journal of Cognitive Neuroscience* 16, 374–381.

Papeo, L., A. Vallesi, A. Isaja & R. I. Rumiati. 2009. Effects of TMS on different stages of motor and non-motor verb processing in the primary motor cortex. *PLoS One* 4, e4508.

Papeo, L., C. Corradi-Dell´Acqua & R. I. Rumiati. 2011. 'She' is not like 'I': the tie between language and action is in our imagination. *Journal of Cognitive Neuroscience* 23, 3939–3948.

Parsons, Thomas D. & Albert Rizzo. 2008. Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: A meta-analysis. *Journal of Behavior Therapy and Experimental Psychiatry* 39, 250–261.

Paulus, M., O. Lindemann & H. Bekkering. 2009. Motor simulation in verbal knowledge acquisition. *Quarterly Journal of Experimental Psychology (Hove)* 62, 2298–2305.

Perpiña, C., C. Botella, R. Baños, H. Marco, M. Alcañiz & S. Quero. 1999. Body image and virtual reality in eating disorders: Is exposure to virtual reality more effective than the classical body image treatment? *CyberPsychology & Behavior* 2, 149–155.

Postle, N., K. L. McMahon, R. Ashton, M. Meredith & G. I. de Zubicaray. 2008. Action word meaning representations in cytoarchitectonically defined primary and premotor cortices. *Neuroimage* 43, 634–644.

Pulvermuller, F., O. Hauk, V. V. Nikulin & R. J. Ilmoniemi. 2005. Functional links between motor and language systems. *European Journal of Neuroscience* 21, 793–797.

Raposo, A., H. E. Moss, E. A. Stamatakis & L. K. Tyler. 2009. Modulation of motor and premotor cortices by actions, action words and action sentences. *Neuropsychologia* 47, 388–396.

Repetto, C., A. Gaggioli, F. Pallavicini, P. Cipresso, S. Raspelli & G. Riva. 2011. Virtual reality and mobile phones in the treatment of generalized anxiety disorders: A phase-2 clinical trial. *Personal and Ubiquitous Computing*, 1–8.

Repetto, C. & G. Riva. 2011. From virtual reality to interreality in the treatment of anxiety disorders. *Neuropsychiatry* 1, 31–43.

Repetto, Claudia, Alessandra Gorini, Davide Algeri, Cinzia Vigna, Andrea Gaggioli & Giuseppe Riva. 2009a. The use of biofeedback in clinical virtual reality: The intrepid project. *Studies in Health Technology and Informatics* 144, 128–132.

Repetto, Claudia, Alessandra Gorini, Cinzia Vigna, Davide Algeri, Federica Pallavicini & Giuseppe Riva. 2009b. The use of biofeedback in clinical virtual reality: The INTREPID project. *Journal of Visualized Experiments: JoVE*, 2–3.

Riva, G. 1998. Virtual environments in neuroscience. *IEEE Transactions on Information Technology in Biomedicine* 2, 275–281.

Riva, G. 2002. Virtual reality for health care: The status of research. *Cyberpsychology & Behavior* 5, 219–225.

Riva, G., M. Bacchetta, G. Cesa, S. Conti & E. Molinari. 2003. Six-month follow-up of in-patient experiential cognitive therapy for binge eating disorders. *Cyberpsychology & Behavior* 6, 251–258.

Riva, G. 2006. Being-in-the-world-with: Presence meets social and cognitive neuroscience. In G. Riva, M. T. Anguera, B. K. Wiederhold & F. Mantovani (eds.), *From Communication to Presence: Cognition, Emotions and Culture towards the Ultimate Communicative Experience*, 47–80. Amsterdam: IOS Press.

Riva, G., M. Bacchetta, G. Cesa, S. Conti, G. Castelnuovo, F. Mantovani & E. Molinari. 2006. Is severe obesity a form of addiction? Rationale, clinical approach, and controlled clinical trial. *Cyberpsychology & Behavior* 9, 457–479.

Riva, G. 2010. Using virtual immersion therapeutically. In K. Antony, D. Merz Nagel & S. Goss (eds.), *Use of Technology in Mental Health. Applications, Ethics and Practice*, 114–123. Springfield, IL: C.C. Thomas Publisher.

Riva, G., J. A.Waterworth, E. L. Waterworth & F. Mantovani. 2011. From intention to action: The role of presence. *New Ideas in Psychology* 29, 24–37.

Rothbaum, B. O., L. F. Hodges, D. Ready, K. Graap & R. D. Alarcon. 2001. Virtual reality exposure therapy for Vietnam veterans with posttraumatic stress disorder. *Journal of Clinical Psychiatry* 62, 617–622.

Rothbaum, Barbara Olasov, Page Anderson, Elana Zimand, Larry Hodges, Delia Lang, Delia & Jeff Wilson. 2006. Virtual reality exposure therapy and standard (in vivo) exposure therapy in the treatment of fear of flying. *Behavior Therapy* 37, 80–90.

Slater, M., D. Perez-Marcos, H. H. Ehrsson & M. V. Sanchez-Vives. 2009. Inducing illusory ownership of a virtual body. *Frontiers in Neuroscience* 3, 214–220.

Slater, M., B. Spanlang, M. V. Sanchez-Vives & O. Blanke. 2010. First person experience of body transfer in virtual reality. *PLoS One* 5, e10564.

Steuer, Jonathan. 1992. Defining virtual reality: Dimensions determining telepresence. *Journal of Communication* 42, 73–93.

Strafella, A. P. & T. Paus. 2000. Modulation of cortical excitability during action observation: A transcranial magnetic stimulation study. *Neuroreport* 11, 2289–2292.

Tettamanti, M., G. Buccino, M. C. Saccuman, V. Gallese, M. Danna, P. Scifo, F. Fazio, G. Rizzolatti, S. F. Cappa & D. Perani. 2005. Listening to action-related sentences activates fronto-parietal motor circuits. *Journal of Cognitive Neuroscience* 17, 273–281.

Tomasino, B., G.R. Fink, R., Sparing, M. Dafotakis & P. H. Weiss. 2008. Action verbs and the primary motor cortex: A comparative TMS study of silent reading, frequency judgments, and motor imagery. *Neuropsychologia* 46, 1915–1926.

van Dam, W. O., S.A. Rueschemeyer & H. Bekkering. 2010. How specifically are action verbs represented in the neural motor system: An fMRI study. *Neuro Image* 53, 1318–1325.

Vincelli, F., L. Anolli, S. Bouchard, B. K. Wiederhold, V. Zurloni & G. Riva. 2003. Experiential cognitive therapy in the treatment of panic disorders with agoraphobia: A controlled study. *Cyberpsychology & Behavior* 6, 321–328.

Waterworth, John A., Eva L. Waterworth, Fabrizia Mantovani & Giuseppe Riva. 2010. On Feeling (the) Present: An evolutionary account of the sense of presence in physical and electronically-mediated environments. *Journal of Consciousness Studies* 17, 167–178.

Willems, R. M., P. Hagoort & D. Casasanto. 2010a. Body-specific representations of action verbs: Neural evidence from right- and left-handers. *Psychological Science* 21, 67–74.

Willems, R. M., I. Toni, P. Hagoort & D. Casasanto. 2010b. Neural dissociations between action verb understanding and motor imagery. *Journal of Cognitive Neuroscience* 22, 2387–2400.

Willems, R. M., L. Labruna, M. D'Esposito, R. Ivry & D. Casasanto. 2011. A functional role for the motor system in language understanding: Evidence from theta-burst transcranial magnetic stimulation. *Psychological Science* 22, 849–854.

Wilson, M. 2002. Six views of embodied cognition. *Psychonomic Bulletin & Review* 9, 625–636.

Youngblut, C. 2007. *What a Decade of Experiments Reveals about Factors that Influence the Sense of Presence: Latest Findings*. Alexandria, VA: Institute for Defense Analysis.

Claudia Repetto, Barbara Colombo, and Giuseppe Riva
Catholic University of Sacred Heart
Largo Gemelli 1
20121 Milan
Italy
claudia.repetto@unicatt.it, barbara.colombo@unicatt.it, and giuseppe.riva@unicatt.it

The Language of Colour: Neurology and the Ineffable

Nicholas Unwin

It is often claimed, following Joseph Levine, that there is an 'explanatory gap' between ordinary physical facts and the way we perceive things, so that it is impossible to explain, among other things, why colours actually look the way they do. C.L. Hardin, by contrast, argues that there are sufficient asymmetries between colours to traverse this gap. This paper argues that the terms we use to characterize colours, such as 'warm' and 'cool', are not well understood, and that we need to understand the neurological basis for such associations if we are even to understand what is fully meant by saying, for example, that red is a warm colour. This paper also speculates on how Hardin's strategy can be generalized.

Keywords: C.L. Hardin; colour; colour vocabulary; explanatory gap; perception; *qualia*

1. Introduction

A major part of the mind–body problem is to explain why mental states have the phenomenal qualities that they do. It is often held, for example, by Joseph Levine (1983, 1991), that there is an 'explanatory gap' here. The simplest such states are perceptual *qualia*, and the examples most commonly chosen are colour *qualia*. Even if we knew everything about the physics of colour, and even if we knew everything about the eye and the brain, this would not explain why colours actually *look* the way they do, why green stimuli give rise to green *qualia*, for example, as opposed to red ones. This is perhaps because colours are essentially ineffable — simple impressions, in Hume's (1955) sense, which cannot be characterized in any useful way.

However, some (such as C.L. Hardin 1987, 1988, 1997) take a more optimistic line, and argue that there are enough asymmetries within the colour circle to ensure that it is possible to explain why colours look the way they do. In particular, any possible inverted spectrum (seeing green where others see red, for

---

I am grateful to comments made during my presentation at the *Embodied Language* conference, and also to comments made at the Work in Progress seminar at Lancaster University where a version of this paper was discussed.

ISSN 1450–341

Biolinguistics 6.3–4: 475–490, 2012
http://www.biolinguistics.eu

example) can be ruled out as detectable after all. More generally, colour vision science shows there to be many useful connections between phenomenology and physiology. For example, that red is (unlike orange) a *unique* (i.e. unmixed) hue, and unlike green, a *positive*, *advancing*, and *warm* hue, can perhaps be explained physiologically. However, attention needs to be directed to exactly how these italicized predicates — the language of colour — get their meanings. It will be argued that we need to have predicates of this kind if explanations of why colours look the way they do are to be forthcoming, and we are not to surrender to the claim of simple ineffability. It will also be argued that such terms are more than just metaphorical, and that they directly concern how the brain itself works, and thus involve a kind of embodiment of language, one which challenges more traditional pictures of how language works.

I shall argue that an ideal sort of explanation of why red should look warm is that there be some appropriate neurological connections between the visual and tactile parts of the brain (currently, the issue is undecided). This will link visual and tactile warmth in a way that is too direct to be merely metaphorical, but not so simply as to yield literal synonymy. Redness is not wholly ineffable, but not straightforwardly analysable either. However, red–green inversion is not the only inversion that needs to be ruled out. Other terms are needed, and I suggest that greens and yellows have a quality that may be described as *sharp*, *fresh* and *citrusy* whereas reds, blues and purples do not. As with *warm*, this quality does more than just reflect ordinary physical associations (I think). To explain this, we need to find direct neural links between the visual and gustatory centres of the brain. This needs to be further generalized, and I shall speculate on ways in which colour language could be further extended in an explanatorily useful way.

2. The Hering Colour Circle

Hardin draws heavily on the ideas of the 19th century physiologist and founder of modern colour vision science, Ewald Hering. Conventional wisdom says that there are only three basic colours (red, green and blue) from which all others can be obtained by mixture, and that this corresponds to the fact that there are three different kinds of colour photoreceptor in the retina (sometimes known, rather misleadingly, as the red, green, and blue cones[1]) which respond to different parts of the visible spectrum. Hering, however, insisted that there are four unique hues, that is to say, colours which actually look essentially unmixed, namely red, yellow, green, and blue. These yield four binary or essentially mixed hues, namely orange (red–yellow), purple (blue–red), turquoise (green–blue), and chartreuse (yellow–green). This notion of mixture is purely phenomenal, and does not relate to how colours may be obtained by combining lights or pigments. We now know that the phenomenon to which Hering draws our attention is post-

---

[1] The terms are misleading since the cones' sensitivity curves do not peak at the red, green and blue portions of the spectrum: The 'red' cone peaks at yellow-green, and it is the narrow difference in stimulation level between it and the 'green' cone that underpins the red–green perceptual channel.

receptoral, and concerns how differences in stimulation level in the cones are transmitted to the visual cortex. Specifically, there are two retinocortical channels, the red–green channel and the yellow–blue channel (together with the achromatic white–black channel), each of which yields opponent processing. Thus when the first channel is excited, the subject perceives red or a reddish hue; when it is inhibited, the subject perceives green or a greenish hue. Likewise, when the second channel is excited, the subject perceives yellow or a yellowish hue; when it is inhibited, the subject perceives blue or a bluish hue. The results may be summarized in the following diagram:

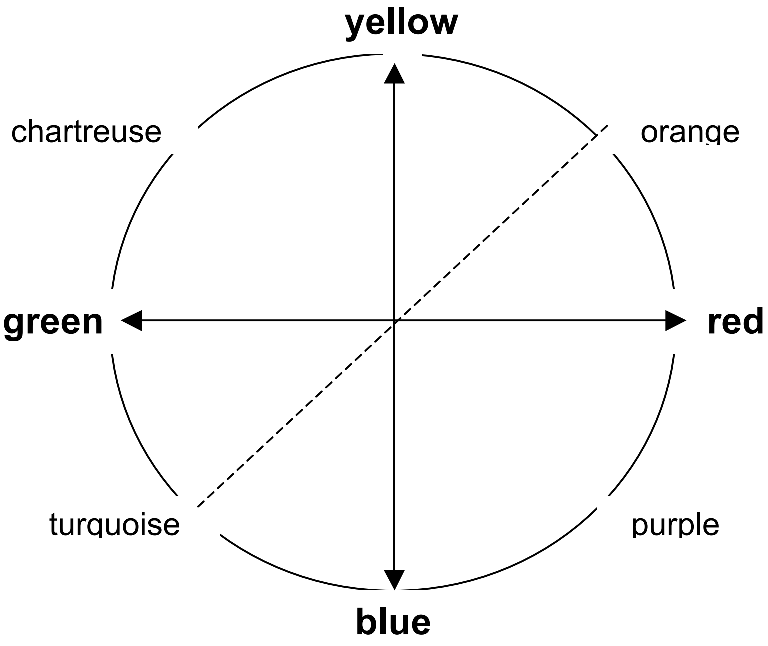


Figure 1: The Hering colour circle

Two classes of phenomenal facts are explained by this analysis. Firstly, the difference between unique and binary hues is accounted for. Colours which look essentially unique and unmixed correspond to the activation of just one retinocortical channel, whereas colours which look like mixtures of unique hues correspond to the activation of two. Secondly, we can see why some combinations of unique hues are perceptually impossible, namely red–green and yellow–blue.[2] This is because a given channel cannot simultaneously be excited and inhibited, any more than a given energy level can simultaneously increase and decrease. As Hering originally predicted, much of the phenomenology of colour perception, that is to say much of what directly presents itself to the colour-sighted percipient, is matched by the underlying physiology, and in a very straightforward sort of way.[3] The explanatory gap has not been completely traversed, to be sure, since there are many other phenomenological facts yet to be

---

[2] Except in extraordinary circumstances. On this, see Crane & Piantanida (1983), Billock *et al.* (2001), and Suarez & Nida-Rümelin (2009).

[3] The phenomenology is supported by quantitative results. See, for example, Hurvich (1997).

explained, but the idea that such a gap is untraversable in principle has been put into serious doubt.

Closely connected to the explanatory gap is the hypothesis of inverted *qualia*, namely that it is possible that you see colours differently from me (see the Appendix for an illustration of this phenomenon). Hardin argues against Levine that there are enough asymmetries in the colour circle depicted above to rule this out. In particular, it is highly implausible that you might see orange where I see red since what you would call 'red' (namely orange) is a binary hue and can be directly perceived as such. If there are to be any undetectable inversions, then at the very least, unique hues must be exchanged with other unique hues. The standard inversion scenario invokes an exchange of red and green, thus leading to a reflection of the colour circle in the vertical diameter. However, Hardin argues that such an inversion is also detectable since red and yellow are essentially *warm* hues, whereas green and blue are essentially *cool* hues; and this too is explicable physiologically. This is much more controversial, however, and this paper argues that this is largely because terms like 'warm' and 'cool', as used in this context, are not well understood.

3. Word/Colour Associations

What do we mean when we say that red is a warm colour? One problem is that ordinary physical associations are involved. As is often observed, reds and yellows are the colours of fire, whereas greens and blues are the colours of lakes. Obviously, red is a *fiery* colour, for example, but people with red–green inverted *qualia* will also agree: (What we call) green is indeed, for them, a colour of fire! Many will insist that this sort of thing is all that the warmth/coolness distinction amounts to, and that they cannot see anything more directly phenomenological involved. The fact that there is this sort of disagreement here is embarrassing, and weakens a lot of ordinary phenomenological evidence, for it undermines the assumption that we can all tell, without too much difficulty, how things look to us. At any rate, if it turns out that ordinary physical associations (together with some additional cultural conventions, perhaps) accounts entirely for the warm–cool distinction, then it cannot be used to explain why red looks like red as opposed to green. But we do not ordinarily suppose that a blue gas flame looks warm even though it feels it, and there is evidence that people really do perceive a genuine phenomenological distinction here. For example, Katra & Wooten, in a recent unpublished study (quoted in Hardin 1997), asked ten subjects to rate eight colour samples as 'warm' or 'cool' on a ten-point scale, with ten as 'very warm.' The mean results gave the lowest rating to the blue sample, and the highest rating to the orange sample. There was a high level of agreement among subjects:

> The remarkable correspondence between the obtained ratings of warmth and coolness and the activation levels in the opponent channels […] suggests that the attribution of thermal properties to colors may be linked to the low-level physiological processes involved in color perception. Higher ratings of warmth corresponded with levels of activation of the opponent

> channels in one direction, while cooler ratings corresponded with activation in the opposite direction. This suggests that a link to the activation level of the opponent channels, rather than the psychological quality of hue, drives the association of temperature with color, and that the association is more than simply a cognitive process.

Also, Ou *et al.* (2004) showed:

> [I]n a psychophysical experiment, 31 observers, including 14 British and 17 Chinese subjects assessed 20 colours on 10 colour-emotion scales: warm–cool, heavy–light, modern–classical, clean–dirty, active–passive, hard–soft, tense–relaxed, fresh–stale, masculine–feminine, and like–dislike. Experimental results show no significant difference between male and female data, whereas different results were found between British and Chinese observers for the tense–relaxed and like–dislike scales. […] Four colour-emotion models were developed, including warm–cool, heavy–light, active–passive, and hard–soft. […] The results show that for each colour emotion the models of the three studies agreed with each other, suggesting that the four colour emotions are culture-independent across countries.[4]

So, suppose that there really is a relevant sort of distinction here.[5] How will it help us? Levine argues, against Hardin, that although warmth is connected to redness, the connection is essentially shallow. The former could be subtracted from the latter to yield a residue, a 'cool red'. If this is right, then we do indeed have an explanatory problem since we have not managed to target what is essential to red. But Hardin rejects the intelligibility of 'cool red', and surely rightly so. What we call visual warmth does seem to be an essential part of redness. It is not all of redness, to be sure, otherwise yellow could not also be warm, so it is possible that a residue exists.[6] But this residue is only half a colour, not a purified red. Should the residue be combined with coolness to produce a new hue, what we would end up with would be something wholly alien and unimaginable, not anything that much resembles red. Appeals to warmth will not explain everything about why red looks the way it does, of course, but it surely can be used to explain something. In particular, if it can be shown that there are some direct links of an appropriate kind between the neurons which fire when we see red (and yellow) and those which fire when we feel warmth, then an explanation is well on its way.

Such neurological connections need to be found, of course, but their elusiveness is not the only obstacle to explanatory success. There remains another inverted *qualia* scenario that needs to be ruled out, which I call 'diagonal inversion', namely one which involves reflection in the dotted diagonal axis of the Hering colour circle depicted above, where red is exchanged with yellow and

---

[4] Quoted from the Abstract.

[5] A negative result here is that young children are less inclined to associate red with warmth, which suggests that the connection is cultural and learnt, not biological. See Morgan *et al.* (1975). The suggestion is not conclusive, however, since innate connections can take time to develop. It is nevertheless odd to suppose that children see colours differently from adults.

[6] Though even this is unobvious. Just because being coloured is a part of being red, for example, it does not follow that there is a residue, namely a quality which remains when we subtract colouredness from redness.

green with blue. Turquoise and orange stay fixed, and purple is exchanged with its complement, chartreuse.[7] How could we extend the warmth-coolness point to handle this? Many people I have asked agree that greens and yellows have a quality that may be described as ‘sharp’, ‘fresh’, and ‘citrusy’ — whereas reds, blues, and purples do not. Obviously, there is the risk, once again, that we are dealing only with familiar physical associations here, since limes are green and lemons are yellow. But blackcurrants also have a sharp and citrusy taste, but do not look to me (at any rate) sharp or citrusy. On the contrary, the pale mauve of a blackcurrant yoghurt looks like the very opposite of a sharp, fresh, or citrusy colour. By contrast, lemons and limes not only taste sharp, fresh; and citrusy — they look that way as well. At least, they do to me, though I have only rather limited evidence that they do to others as well. The very fact that the matter is hard to settle indicates that the semantics of words such as ‘warm’, ‘cool’, ‘sharp’, ‘fresh’, and ‘citrusy’ (the language of colour) needs considerable critical attention if the explanatory gap is to be traversed successfully.

There has been plenty of research done on words apart from ‘warm’ and ‘cool’ that we might associate with different colours (though they do not illuminate the case of diagonal inversion). For example, Lars Sivik (1997: 187), when developing the Swedish Natural Colour System (NCS), discovered that:

> [t]here are many words in common use to describe the character and associative meanings of colours. Besides such attributes as yellow, blue, strong, weak, deep, and saturated, colours can also have connotations like cold, joyful, depressing, sick, healthy, dirty, feminine, masculine, etc. Such colour-relevant adjectives can add up to a rather long list. It is now possible to make a semantic map from the average judgements for each of all imaginable words, or for pairs of opposites if we choose to use bi-polar scales as in the masculine–feminine example above. In our first studies (Sivik 1970), twenty-seven such antonyms were mapped out in the NCS.

These results concern all colours, including browns and greys, not just maximally saturated hues (which are what I have been considering). Whether, and to what extent, these colour/word connections yield explanations will depend on just why they come about. Unless they relate to intrinsic phenomenological facts, rather than external associations, they will probably not be significant if only

---

[7] Some rule out this possibility on the grounds that yellow is a much lighter colour then red, and also that there are more perceptual differences between red and blue than there are between yellow and green (see, e.g., Hardin 1988: 134–42, and Palmer 1999). However, these facts relate not to hue but to chroma, or saturation level, and hue and chroma are usually thought of as independent dimensions of colour, fact that is directly, visibly evident to us; so it is unclear if such asymmetries are really explanatory in any relevant way. That yellow has a very low chromatic content is easily explained by the fact that it occurs in the middle of the visible spectrum, where the light-dark sensitivity curve peaks. Green also has a lower chromatic content than red or blue. These facts might change for a ‘diagonally inverted’ percipient, who might be able to perceive a ‘supersaturated yellow’, a colour which relates to yellow as red relates to pink (and who would be unable to perceive a ‘supersaturated pink’, i.e. what we call ‘red’). Supersaturated yellow is unimaginable to us, but does not seem paradoxical in the way in which Levine’s ‘cool red’ is, since it merely involves stretching things a bit. We have no idea how many perceptible differences there are between green and supersaturated yellow. For more on this, see Unwin (2011).

because they will not differentiate between the experiences of normal percipients and those of colour-inverted percipients. Without such deep links, colours will remain essentially ineffable, and hence unexplainable. It is difficult to see what research could be done that would decide the matter, since it would have to rely very heavily on asking somewhat technical questions of naive subjects. However, the main focus of this paper concerns what we are *saying* when we say that red is a warm colour, and which perhaps we are not saying when we say that pink is a feminine colour, and this contrast is in itself rather hard to analyse.

So what sort of claim are we making when we say that red is a warm colour? It might be thought that the term 'warm' is purely metaphorical here, as in a 'warm' greeting. However, this does not seem to do justice to the force with which the warmth strikes us. Metaphors are things that we can usually take or leave, and although they can sometimes be very striking, they do not seem to relate to intrinsic character in a sufficiently robust sort of way. It is, after all, meant to be a primitive phenomenological fact that red looks warm, so primitive that if something fails to look warm then it necessarily fails to look red. It might be thought, on the other hand, that the term is simply literal: Red *is* just literally warm. This, however, is also unsatisfactory as it fails to do justice to the differences in the sensory modalities and associated secondary qualities. True, we describe chillies as 'hot' and this seems literally (i.e. not just metaphorically) right as far as appearances go, despite the fact that we are referring to flavour rather than an ordinary tactile sensation caused by a rise in temperature; but gustatory heat, or piquancy, is not strictly speaking a taste in the way in which sweetness, sourness, and so forth are tastes, since it is carried to the brain by a different set of nerves. Flavour is a complex intermodal sense, and should be distinguished from pure taste. Moreover, piquancy does relate very closely to an ordinary burning sensation on the tongue even if it is not accompanied by a rise in temperature. The reason is that similar things are happening to the tongue in each case. By contrast, warmth does not seem to relate to redness in this direct sort of way: The resemblance is not sufficiently close. What we seem to have, therefore, is something in between literal synonymy and metaphor. This in itself yields a problem, since it is unclear what that amounts to. True, we are familiar with dead metaphors, which are half way between real metaphors and literal meanings, but this again does not seem to be the sort of thing we should be looking for.

Finding generally agreed associations between colours and particular words is evidently not enough to yield interesting explanations of why colours look the way they do, if only because it needs to be shown that the meanings of the words in question attach themselves sufficiently deeply to the phenomenology. To some extent, cultural associations can be identified and used to screen the reliability of such associations. For example, we need not attach much significance to the connection between the colour pink and the word 'feminine', if only because the association is comparatively recent and does not extend across all cultures. But it may be hard to generalize this kind of screening technique, and Sivik's research in developing the NCS yields a bewildering array of terms and associations. Much work has been done in configuring semantic scales and developing the topography of associations, but the issue of explanation, of just *why* certain colours (or colour combinations) should be thought of as 'mighty' or ‘militaristic’, for example, is not easy to address.

4 The Role of Neurology

A more promising locus of explanation is the brain, but here we also have difficulties. I suggested, following Hardin (1988: 129–84), that there might be significant links between those neurons which fire when we see reds and yellows on the one hand, and those which fire when we perceive tactile warmth on the other. There is currently not much evidence for this, but the crucial point is that *if* such links were to exist, then we would have something which is genuinely explanatory; and conversely, without such links, it is hard to see how any useful explanation (i.e. one which addresses the explanatory gap) can exist. This is of philosophical importance even if it is purely speculative. I suggest that it would do more than just explain why warmth is associated with reds and yellows; it would also reinforce the claim that reds and yellows really do look intrinsically warm, and would help us to answer the doubters who claim that they cannot see any such intrinsic warmth, and that we are dealing only with physical and cultural associations. This might seem paradoxical: Surely, it may be said, the *explanandum* needs to be firmly in place before we look for the *explanans*. Specifically, we must be confident that red really does look intrinsically warm (and hence know what this means) before asking why this should be so. In response, I can say that in my own case, I was disinclined to believe that red was *intrinsically* warm until the possibility of a neurologically grounded intermodal link was suggested to me: Until then, I could not see clearly what could even be *meant* by the claim. Merely saying that seeing red and feeling warmth resemble each other is not enough: Unlike the case of piquancy, the resemblance itself is not strong enough to underpin the claim.[8]

Some qualification is needed here. I am not suggesting that *all* of what is meant by saying that red is a warm colour is that there are appropriate intermodal neural links, still less that the unknown details should form part of the meaning of what we are currently saying. Since the links are required to explain the resemblance, we would otherwise run the risk that *explanans* and *explanandum* will coalesce, thus rendering the explanation trivial: That is to say, we end up saying that the existence of certain intermodal neural links explains why there are certain intermodal neural links. Rather, the neurological claim — or at the very least the weaker and physically non-specific claim that there is an important, deep link within the internal processing mechanisms involved in seeing red and feeling warmth (without the details) — should be part of what *elucidates* the particular sort of resemblance between redness and warmth that we are trying to explain. The explanation itself then consists in filling in the details. There may remain a kind of circularity here, but it is relatively harmless. It does, however, ensure that questions of meaning — i.e. questions about what we are actually

---

[8] However, Austen Clark (1993, 1994) has argued that such resemblances across our whole quality space, together with neural links, yields a sufficiently asymmetrical system that they yield a full physicalist reduction of all our sensory qualities. For a critique of Clark, see Unwin (2011).

saying when we say that red looks warm — remain prominent.

Nevertheless, it might still be wondered whether we really need to talk about the *language* of colour at all. After all, nonhuman animals have colour vision, and presumably there can be intermodal links there as well. Perhaps red rags present a warm sort of visual sensation to bulls which blue or green rags do not do. At any rate, it is certainly possible for certain wavelengths of incoming light to be more arousing than others, as they are in humans, and we do not need to ask about the bull's colour vocabulary or cross-modal linguistic associations in order to establish this. Furthermore, it is often said, following Thomas Nagel (1979), that we have no conception of what nonhuman *qualia* can be like, and this fact is itself often used as an argument against physicalism and in favour of there being an untraversable explanatory gap. But here, as elsewhere, the point can be exaggerated. If we suppose, for the sake of argument, that bulls are aroused more by red rags than by blue ones, then it is reasonable to infer that the hue perceived in the former case more closely resembles human red than blue. This is because red light is more arousing than blue. This conclusion can only be tentative, of course, but it would be reinforced if it could be shown that there are further physiological resemblances between bulls' brains and human ones. Yet once again, the specific issue of *language* seems to have dropped out of the picture, and the suspicion may be that it should never have been in the picture in the first place.

However, this would be a mistake. We need to consider other terms besides 'warm' and 'cool' if we are to generalize this strategy, as we need to do if we are to handle diagonal inversion, for example, and the only realistic way to tell whether people perceive certain intermodal resemblances is to ask whether certain words are associated with certain colours. This, at least, must be our first line of inquiry. We then need to filter out the external or culturally-driven associations by focusing on what may be intrinsic to the sensory processing itself. What seems to follow, then, is that the terms we use to characterize the phenomenology of colours are themselves closely connected to what goes on in the brain. Colour language is thus, in a sense, embodied, and cannot be studied independently from bodily functioning, in particular, brain activity. If we neglect the language, then we lack the means to test — or even to look for — hypotheses about intersensory connections; and if we neglect the embodiment, then we fail to distinguish the intrinsic language of colour from the much looser web of associations revealed by, for example, Sivik's (1997) studies.

5. Conclusion

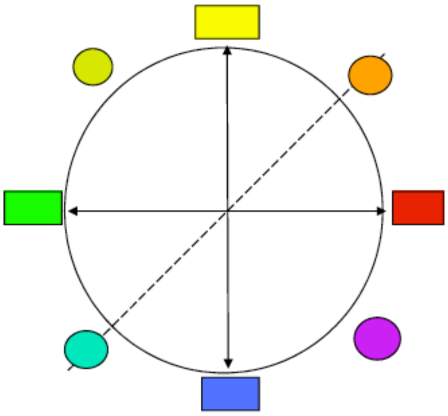
What further empirical research is needed? Ideally, we should like to hear from people who view the world through hue-inverting spectacles, especially those who have worn them from birth! If people who have always worn red–green inverting spectacles judge that (what we call) red things look warm and green things look cool, then our thesis is seriously undermined, but is confirmed if they make opposite judgements. We can likewise ask if diagonally inverted percipients associate the terms 'sharp', 'fresh', and 'citrusy' with colours in the same way as the rest of us. We can also ask people who try on any kind of hue-inverting spectacles in midlife whether after a time things start to look as they did before (compare this with studies that show that people acclimatize gradually to up–down inverting spectacles). If so, then the whole notion of a colour *quale* will be placed in jeopardy; but it would also suggest, and for just that reason, that we do not have the kind of explanatory gap that provoked the discussion in the first place. In the absence of the necessary technology, more down-to-earth studies of the kind undertaken by Sivik, Katra and Wooten, and others are desirable, where normal subjects are asked how well they think a given term is associated with a given colour. But terms may need to be restricted to phenomenal terms relating to non-visual senses (it is unclear how else to guard against irrelevant external associations). We also need to learn more about inter-sensory connections in general, and research on synaesthesia may be of help here.[9] With such research in place, we are better able to see if the brain connections mirror the sensory associations.

What of the philosophical conclusions, in particular with regard to the explanatory gap? It should be noted that no attempt has been made to close the gap completely, and it is hard to see how to do this. Explanations tend to be contrastive, and there are too many potential contrasts. That is to say, instead of simply asking questions of the form ‘Why X?’ we tend to ask questions of the form ‘Why X as opposed to Y?’, and there are too many candidates for ‘Y’ here. Even if we can explain why green looks like green as opposed to red (we can talk about warmth and coolness), that will not explain why green looks like green as opposed to blue. I have suggested ways in which we could extend a similar type of explanation here; but even if that were successful, it would not explain why green looks like green as opposed to some wholly alien hue, such as Levine’s ‘cool red’, for example. Nor does it address David Chalmers’s (1996) ‘hard problem’, namely of why green should look like green as opposed to nothing at all (more generally, why physical processes should give rise to any *qualia* of any kind in the first place). But the ‘hard problem’ is not the only problem of interest, and explanatoriness comes in degrees. Just because we have failed to explain everything, it does not follow that we have explained nothing, and we should not belittle the significance of coming to understand how and why our ordinary colour vocabulary links with other sensory words. And brain processes certainly play an explanatory role here, even if the mind–body problem remains alive and unsolved.

---

[9] It is often said that we are all weak synaesthetes, and this weak synaesthesia is evidently crucial to the Hardin strategy for traversing the explanatory gap. Full-blown synaesthesia is evidently irrelevant to such a project if only because there is such variation between subjects as to how different modalities connect. However, Marks (1978: 218–20) claims that synaesthesia is not what is at stake here.

Appendix: The world as seen through hue-inverting spectacles



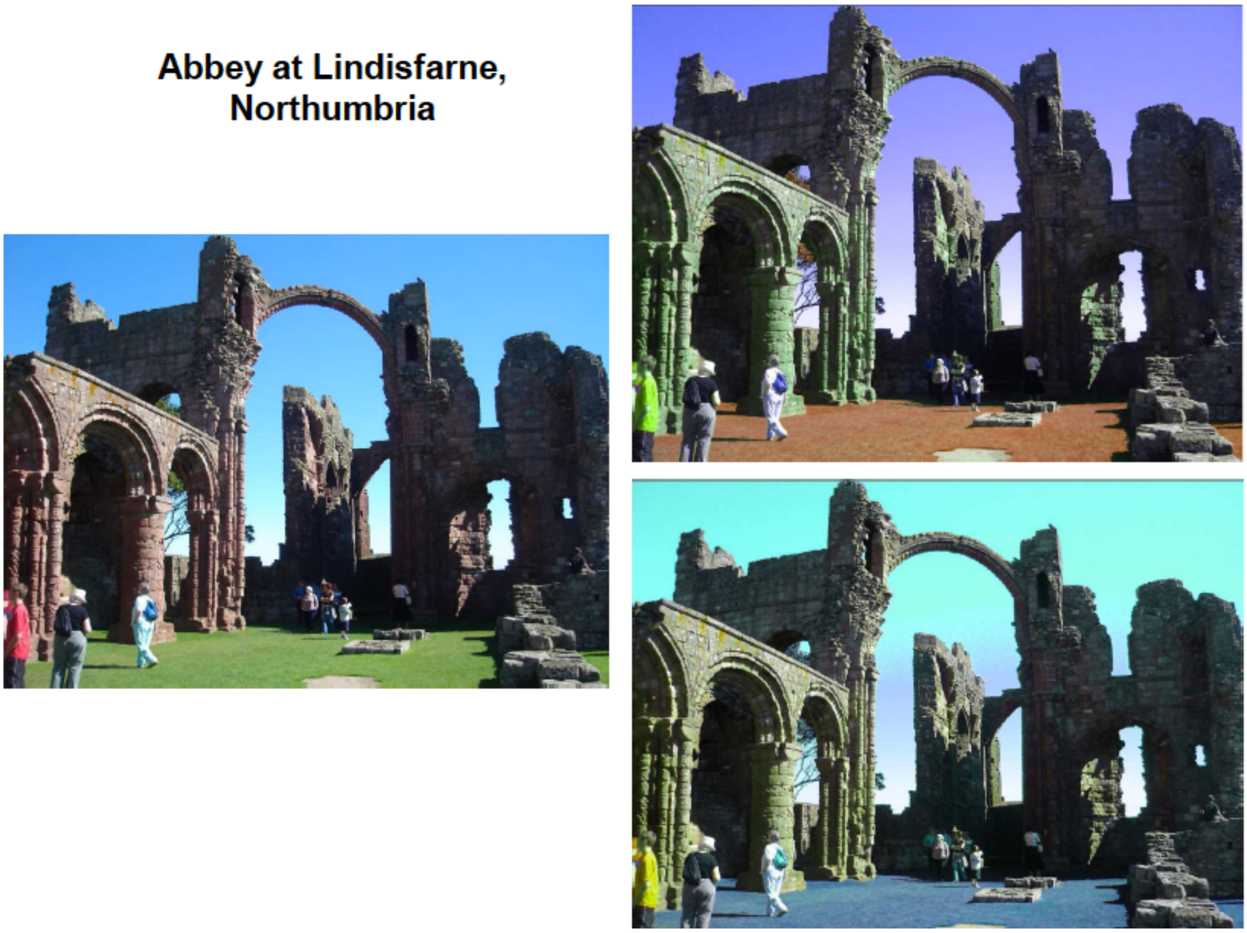
The following slides contain normal photographs together with red-green and diagonal inversions.

Red-green inversion involves reflection in the vertical axis of the colour circle.

Diagonal inversion involves reflection in the dotted axis. Red is thus exchanged with yellow, and green with blue.

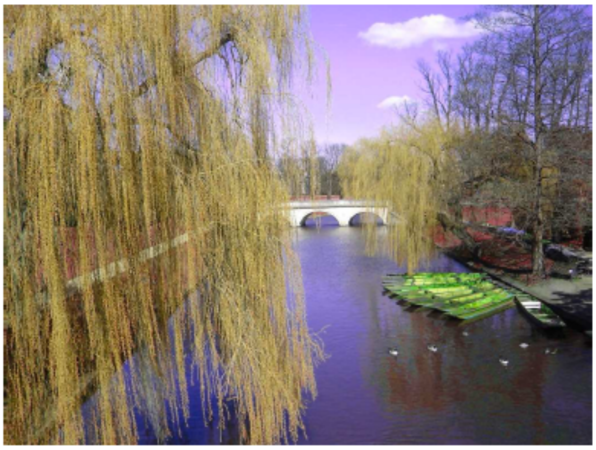
A Hawaiian sunset

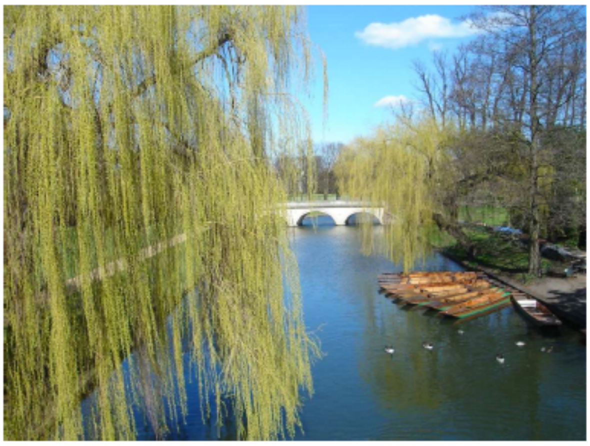
Holy Island, Northumbria



Abbey at Lindisfarne, Northumbria

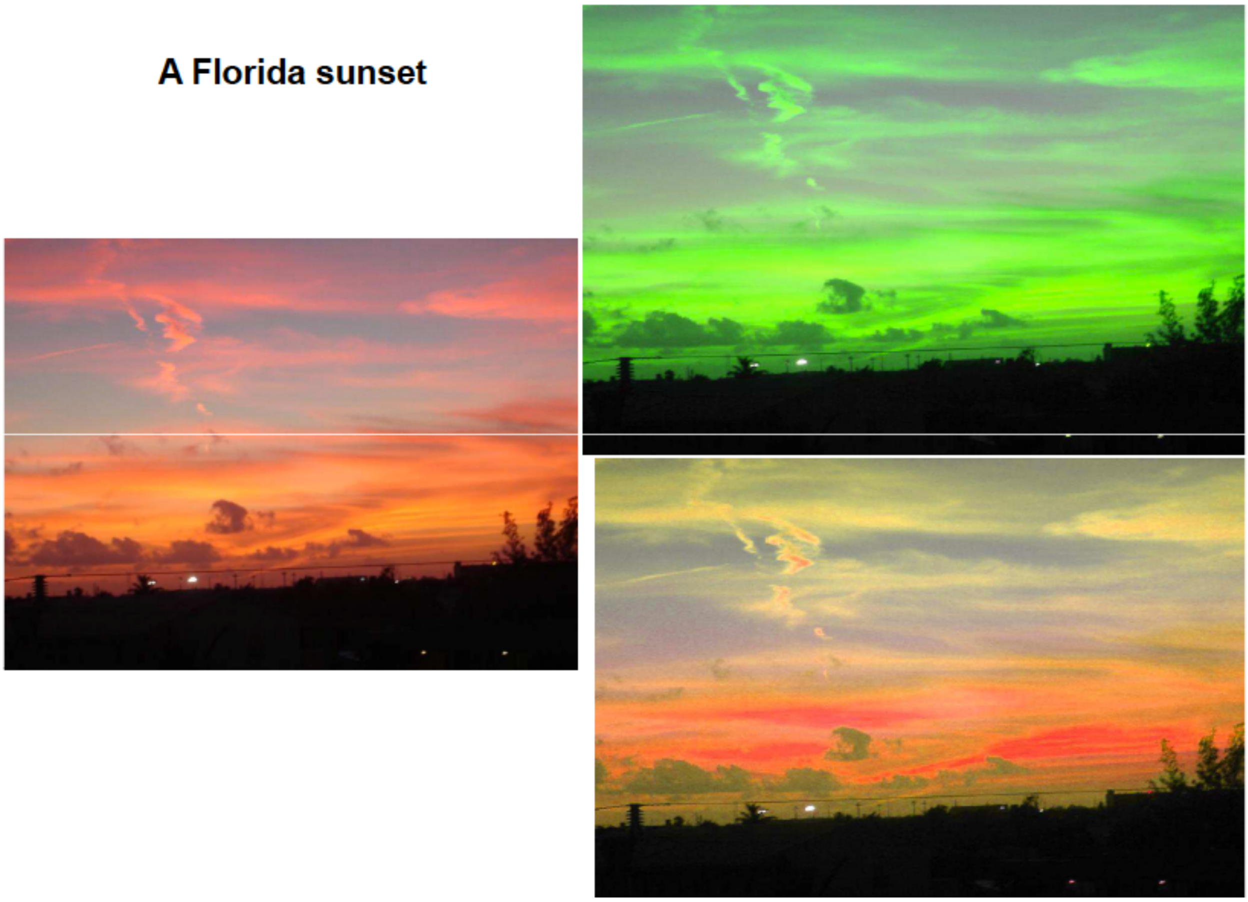
The Backs, Cambridge





Rivington reservoir, Lancashire


A London street


A Florida sunset

References

Billock, V. A., G. A. Gleason & B. H. Tsou. 2001 Perception of forbidden colors in retinally stabilized equiluminant images: An indication of softwired cortical color opponency? *Journal of the Optical Society of America A* 18, 2398–2403.

Chalmers, David. 1996. *The Conscious Mind: In Search of a Fundamental Theory*. Oxford: Oxford University Press.

Clark, Austen. 1993. *Sensory Qualities.* Oxford: Clarendon Press.

Clark, Austen. 1994. I am Joe's explanatory gap. Retrieved from http://selfpace.uconn.edu/paper/PGAP.HTM.

Crane, H. & T. P. Piantanida. 1983. On seeing reddish green and yellowish blue. *Science* 221, 1078–1080.

Hardin, C. L. 1987. Qualia and materialism: Closing the explanatory gap. *Philosophy and Phenomenological Research* 48, 281–298.

Hardin, C. L. 1988. *Color for Philosophers: Unweaving the Rainbow* (expanded edn.). Indianapolis, IN: Hackett.

Hardin, C. L. 1997. Reinverting the spectrum. In Alex Byrne & David R. Hilbert (eds.), *Readings on Color, Vol. 1: The Philosophy of Color*, 289–301. Cambridge, MA: MIT Press.

Hardin, C. L. & Luisa Maffi (eds.). 1997. *Color Categories in Thought and Language*. Cambridge: Cambridge University Press

Hume, David. 1955. *A Treatise of Human Nature*. Ed. by Selby–Bigge. Oxford: Oxford University Press.

Hurvich, Leo M. 1997. Chromatic and achromatic response functions. In Alex Byrne & David R. Hilbert (eds.), *Readings on Color, Vol. 2: The Science of Color*. Cambridge, MA: MIT Press.

Kaiser, Peter K. 1984. Physiological response to color: A critical review. *Color Research and Application* 9, 29–36.

Katra, B. & B. H. Wooten. n.d. Perceived lightness/darkness and warmth/coolness in chromatic experience. Unpublished ms.

Levine, Joseph. 1983. Materialism and Qualia: The explanatory gap. *Pacific Philosophical Quarterly* 64, 354–361.

Levine, Joseph. 1991. Cool red: A reply to Hardin. *Philosophical Psychology* 4, 27–40.

Marks, Lawrence E. 1978. *The Unity of the Senses*. New York: Academic Press.

Morgan, G. A., F. E. Goodson & T. Jones. 1975. Age differences in the associations between felt temperatures and color choices. *American Journal of Psychology* 88, 125–130.

Nagel, Thomas. 1979. *Mortal Questions*. Cambridge: Cambridge University Press.

Ou, Li-Chen, M. Ronnier Luo, Andrée Woodcock & Angela Wright. 2004. A study of colour emotion and colour preference. Part I: Colour emotions for single colours. *Color Research and Application* 29, 232–240.

Palmer, Stephen E. et al. 1999. Color, consciousness and the isomorphism constraint (plus commentaries). *Behavioral and Brain Sciences* 22, 923–989.

Sivik, Lars. 1997. Colour systems for cognitive research. In C. L. Hardin & Luisa Maffi (eds.), *Color Categories in Thought and Language*, 163–193. Cambridge: Cambridge University Press.

Suarez, Juan & Martine Nida-Rümelin. 2009. Reddish green: A challenge for modal claims about phenomenal structure. *Philosophy and Phenomenological Research* 78, 346–391.

Unwin, Nicholas. 2011. Why do colours look the way they do? *Philosophy* 86, 405–424.

Nicholas Unwin
Lancaster University
Department of Politics, Philosophy and Religion
Lancaster LA1 4YL
United Kingdom

n.unwin@lancaster.ac.uk

★ FORUM ★

Notice

We would like to use this opportunity to thank all those involved in creating the sixth volume of *Biolinguistics*. Our special gratitude goes to the reviewers that have served us throughout 2012, who are listed below (colleagues who reviewed more than one submission are suffixed by an asterisk). For everything else, we thank our supporters as well as all the members of the *Biolinguistics* Advisory Board, the *Biolinguistics* Editorial Board, and the *Biolinguistics* Task Team that are not specifically mentioned by name for active participation and feedback. We do, however, want to emphasize the excellent editorial support that Evelina Leivada provided, which made the final proofing so much easier.

Reviewers

Michael L. Arbib
Lisa Aziz-Zadeh
Marcello Barbieri
Lluís Barcelo-Coblijn
Harold Bekkering
Antonio Benítez Burraco
Iris Berent
Anna M. Borghi
Carlo Cecchetto
G. J. Dalenoort
Wessel van Dam
Anna Maria Di Sciullo
John E. Drury
Dean Falk
Vyv Evans
Marta Ferrer-García
W. Tecumseh Fitch
Ian FitzPatrick
Naama Friedman
Koji Fujita*
Alberto Greco
Vittorio Gallese
Johannes Haack
Michel A. Hofman
James R. Hurford
William Idsardi
Maria Kambanaros*
Julian Kiverstein
David A. Leavens
Pierre Livet
Terje Lohndal*
Victor M. Longa*
Christopher Macann
Pete Mandik
James McGilvray
Hiroki Narita*
Tanja Nijboer
Diane Pecher
Massimo Piattelli-Palmarini
M. Carme Picallo
Ian Roberts
Tom Roeper
Bridget Samuels
Laura J. Speed
Yoshida Teruyoshi
Robert Truswell
George Tsoulas
Kate Watkins
Jeffrey Watumull
Masaya Yoshida

We also acknowledge a four-year grant for editorial office expenses awarded by the University of Cyprus to Kleanthes Grohmann (2009–2012) and support from the European Union in the form of a Marie Curie International Reintegration Grant to Cedric Boeckx (PIRG-GA-2009-256413).

biolinguistics

ISSN 1450–341

Biolinguistics

Volume 6, Issue 3–4

Summer/Fall 2012

TABLE OF CONTENTS

★ EDITORIAL ★

ISSN 1450–3417

Biolinguistics 6.3-4: 246–491, 2012
http://www.biolinguistics.eu